David Havelock
Sonoko Kuwano
Michael Vorländer
EDITORS



HANDBOOK OF Signal Processing in Acoustics

VOLUME 1

Springer

Handbook of Signal Processing in Acoustics

Volume 1

Edited by

David Havelock

National Research Council, Canada

Sonoko Kuwano

Osaka University, Japan

Michael Vorländer

RWTH Aachen University, Germany

Handbook of Signal Processing in Acoustics

Volume 1

Springer

Editors
David Havelock
National Research Council
Institute for Microstructural Sciences
Acoustics and Signal Processing Group
1200 Montreal Road
Ottawa ON K1A 0R6
Canada

Sonoko Kuwano
Osaka University
Graduate School of Human Sciences
Department of Environmental Psychology
1-2 Yamadaok Suita
Osaka
Japan

Michael Vorländer
RWTH Aachen University
Institute of Technical Acoustics
Aachen
Germany

ISBN: 978-0-387-77698-9 e-ISBN: 978-0-387-30441-0

Library of Congress Control Number: 2008923573

© 2008 Springer Science+Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.
Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Editor Biographies

David Havelock
National Research Council
Institute for Microstructural Sciences
Acoustics and Signal Processing Group
1200 Montreal Road
Ottawa ON K1A 0R6
Canada

Sonoko Kuwano
Osaka University
Graduate School of Human Sciences
Department of Environmental Psychology
1-2 Yamadaok Suita
Osaka
Japan

Michael Vorländer
RWTH Aachen University
Institute of Technical Acoustics
Aachen
Germany

Editorial Board

Last Name	Given Name(s)	Affiliation/Address
Beauchamp	**James W.**	University of Illinois Urbana-Champaign, School of Music, Dept. of ECE, Urbana, IL, USA
Christie	**Douglas R.**	Earth Physics, Research School of Earth Sciences, The Australian National University, Canberra, AUSTRALIA
Deffenbaugh	**Max**	ExxonMobil Research and Engineering Company, Annandale, NJ USA
Elliott	**Stephen J.**	Southampton University, Institute of Sound and Vibration Research, Southampton, ENGLAND
Fastl	**Hugo**	Technische Universitaet Muenchen AG Technische Akustik, MMK Muenchen, GERMANY
Gierlich	**Hans Wilhelm**	HEAD acoustics GmbH Telecom Division, Herzogenrath, GERMANY
Guyader	**Jean-Louis**	Labratory for Vibration Acoustics Villeurbanne, FRANCE
Jacobsen	**Finn**	Technical University of Denmark, Lyngby, DENMARK
Karjalainen	**Matti**	Helsinki University of Technology, FINLAND
Kollmeier	**Birger**	Universitat Oldenburg, Oldenburg, GERMANY
O'Shaughnessy	**Douglas D.**	INRS-EMT (Telecommunications), Montreal, QC, CANADA
Riquimaroux	**Hiroshi**	Doshisha University, Department of Knowledge Engineering & Computer Sciences, Sensory & Cognitive Neuroscience Research Laboratory, Kyotanabe, Kyoto, JAPAN
Sullivan	**Edmund J.**	EJS Consultants Portsmouth, RI, USA

Last Name	Given Name(s)	Affiliation/Address
Suzuki	**Hideo**	1-2-3-S2502, Utase, Mihama-ward, Chiba-city, JAPAN 261-0013
Taroudakis	**Michael**	University of Crete Department of Mathematics, and FORTH, Institute of Applied and Computational Mathematics, Heraklion, GREECE
Ueha	**Sadayuki**	Tokyo Institute of Technology, Director, Precision and Intelligence Lab, Yokohama, JAPAN
Verrillo	**Ronald T.**	(deceased) Institute for Sensory Research, Syracuse, NY, USA
Yamada	**Ichiro**	Aviation Environment Research Center, Airport Environment Improvement Foundation, Tokyo, JAPAN
Yamasaki	**Yoshio**	Waseda University, Graduate School of Global Information and Telecommunication Studies, Saitama, JAPAN
Tohyama	**Mikio**	Waseda University, JAPAN

Preface

Acoustics has a special relationship with signal processing. Many concepts in signal processing arise naturally from our general experience with sound and vibration and, more than in many other fields, acoustics is concerned with the acquisition, analysis, and synthesis of signals. Consequently, there is a rich resource of signal processing expertise within the acoustics community.

There are many excellent reference books devoted to signal processing but the objective of the *Handbook of Signal Processing in Acoustics* is to bring together the signal processing expertise specific to acoustics and to capture the interdisciplinary nature of signal processing within acoustics. It is also hoped that the handbook will promote networking and the interchange of ideas between technical areas in acoustics.

The handbook comprises 104 *Chapters* organized into 17 *Parts*. Each Part addresses a technical area of acoustics, reflecting the general demarcations of specialization within the acoustics community. An expert with broad knowledge of signal processing within their respective technical area was invited to act as a Section Leader for each Part of the handbook. These Section Leaders contributed substantially to the handbook project by helping to define the contents and scope of each chapter, finding an appropriate contributing expert author, and managing the review and revision of material. Collectively with the Editors, they form the Editorial Board for the handbook.

Planned sections on Architectural Acoustics, Nonlinear Acoustics, and Ultrasound are unfortunately omitted from the handbook; nevertheless, the handbook otherwise provides thorough coverage of the field of acoustics and we can hope that possible future editions might include these areas.

The handbook is written from the perspective of acoustics, by acousticians with signal processing expertise. Emphasis is placed

in the description of acoustic problems and the signal processing related to their solutions. The reader is assumed to have basic knowledge of signal processing. Signal processing techniques are described but the reader is referred elsewhere for derivations and details.

The authors were not required to adhere to strict standards of style or notation, and were asked to prepare short, concise, self-sufficient chapters. This results in variations in style and notation throughout the handbook that reflects the diversity of perspectives within the acoustics community.

David Havelock

Acknowledgments

David Havelock gratefully acknowledges support from the Institute for Microstructural Sciences, of the National Research Council of Canada, for making time and resources available during the preparation of this handbook. He also thanks John C. Burgess (University of Hawaii, retired) for the encouragement to begin this project and expresses his gratitude to the co-Editors, Sonoko Kuwano and Michael Vorländer, who were a pleasure to work with throughout this project. The efforts of all of the Section Leaders are greatly appreciated and we thank each of them sincerely for their patience, perseverance, and faith that were required to see this project to completion. We thank each of the authors for their contributions.

Contents

Editor Biographies v

Editorial Board vii

Preface ix

Acknowledgments xi

Contributors xvii

PART I ACOUSTIC SIGNALS AND SYSTEMS 1

1 Signals and Systems 3

2 Acoustic Data Acquisition 17

3 Spectral Analysis and Correlation 33

4 The FFT and Tone Identification 53

5 Measuring Transfer-Functions and Impulse Responses 65

6 Digital Sequences 87

7 Filters 107

8 Adaptive Processing 125

9 Beamforming and Wavenumber Processing 131

PART II AUDITORY SYSTEM AND HEARING 145

10 Anatomy, Physiology and Function of the Auditory System 147

11 Physiological Measures of Auditory Function 159

12 Auditory Processing Models 175

13 Speech Intelligibility 197

14 Signal Processing in Hearing Aids 205

PART III PSYCHOACOUSTICS 213

15 Methods for Psychoacoustics in Relation to Long-Term Sounds 215

16 Masking and Critical Bands 229

17 Aspects of Modeling Pitch Perception 241

18 Calculation of Loudness for Normal and Hearing-Impaired Listeners 251

19 Psychoacoustical Roughness 263

PART IV MUSICAL ACOUSTICS 275

20 Automatic Music Transcription 277

21 Music Structure Analysis from Acoustic Signals 305

22 Computer Music Synthesis and Composition 333

23 Singing Voice Analysis, Synthesis, and Modeling 359

24 Instrument Modeling and Synthesis 375

25 Digital Waveguide Architectures for Virtual Musical Instruments 399

26 Modeling of Musical Instruments 419

PART V SPEECH 447

27 Display and Analysis of Speech 449

28 Estimation of Speech Intelligibility and Quality 483

29 Gaussian Models in Automatic Speech Recognition 521

30 Speech Synthesis 557

31 Speech Coders 587

PART VI AUDIO ENGINEERING 621

32 Transducer Models 623

33 Loudspeaker Design and Performance Evaluation 649

34 PA Systems for Indoor and Outdoor 669

35 Beamforming for Speech and Audio Signals 691

36 Digital Audio Recording Formats and Editing Principles 703

37 Audiovisual Interaction 731

38 Multichannel Sound Reproduction 747

39 Virtual Acoustics 761

40 Audio Restoration 773

41 Audio Effects Generation 785

42 Perceptually Based Audio Coding 797

PART VII TELECOMMUNICATIONS 819

43 Speech Communication and Telephone Networks 821

44 Methods of Determining the Communicational Quality of Speech Transmission Systems 831

45 Efficient Speech Coding and Transmission Over Noisy Channels **853**

46 Echo Cancellation **883**

47 Noise Reduction and Interference Cancellation **897**

48 Terminals and Their Influence on Communication Quality **909**

49 Networks and Their Influence on Communication Quality **915**

50 Interaction of Terminals, Networks and Network Configurations **921**

List of Important Abberviations **927**

Contributors

Tomonari Akamatsu
National Research Institute of Fisheries Engineering, Fisheries Research Agency of Japan, Kamisu, Ibaraki, Japan

Benoit Alcoverro
CEA/DASE, BP 12, 91680 Bruyères Le Châtel, France

Lydia Ayers
Computer Science Department, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, e-mail: layers@cs.ust.hk

Juha Reinhold Backman
Nokia, Espoo, Finland

Rolf Bader
Institute of Musicology, University of Hamburg, Hamburg, Germany

James W. Beauchamp
School and Music and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 2136 Music Building, 1114 West Nevada Street, Urbana, IL 61801, USA, e-mail: jwbeauch@uiuc.edu

Kim Benjamin
Naval Undersea Warfare Center, Newport, RI, USA

A.J. Berkhout
Delft University of Technology, Delft, The Netherlands

Jeff Bilmes
Department of Electrical Engineering, University of Washington, Seattle, WA, USA

Ole-Herman Bjor
Norsonic AS, Lierskogen, Norway, e-mail: ohbjor@norsonic.com

Stanley J. Bolanowski (deceased)
Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA

Baris Bozkurt
Izmir Institute of Technology (IYTE), Izmir, Turkey

Nicolas Brachet
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Thomas Brand
University of Oldenburg, Oldenburg, Germany

David J. Brown
Geoscience Australia, Canberra, Australia, (Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria), e-mail: David.Brown@CTBTO.org

John C. Burgess
Department of Mechanical Engineering, University of Hawaii, 2540 Dole Street, Honolulu, HI 968221, USA, e-mail: jcb@hawaii.edu

Paola Campus
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Yves Cansi
CEA/DASE/LDG, BP12, 91680 Bruyères-le-Châtel, France

Josef Chalupper
Siemens Audiological Engineering Group, Gebbertstrasse 125, Erlangen 91058, Germany, e-mail: Josef.Chalupper@siemens.com

N. Ross Chapman
University of Victoria, Victoria BC, Canada

Jakob Christensen-Dalsgaard
Institute of Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, e-mail: jcd@biology.sdu.dk

Peter Daniel
Brüel & Kjær, GmbH, Bremen, Germany

Roger B. Dannenberg
Carnegie Mellon University, Pittsburgh, PA, USA

Torsten Dau
Technical University of Denmark, Lyngby, Denmark, e-mail: tda@elektro.dtu.dk

Hans-Elias de Bree
Microflown Technologies, The Netherlands, USA

Max Deffenbaugh
ExxonMobil Research and Engineering Company, Annandale, NJ, USA

Thierry Dutoit
Faculte Polytechnique de Mons, Mons, Belgium

Stephen J. Elliott
Institute of Sound and Vibration Research, Southampton, UK

Paulo A.A. Esquef
Nokia Institute of Technology, Rod. Torquato Tapajós, 7200, Tarumã 69048-660 Manaus-AM, Brazil, e-mail: paulo@esquef@indt.org.br

Richard R. Fay
Parmly Hearing Institute, Loyola University Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, USA, e-mail: rfay@luc.edu

Michael Fehler
Massachusetts Institute of Technology, Cambridge, USA

Sandy Fidell
Fidell Associates, Inc., Woodland Hills, CA, USA

Petr Firbas
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria (International Atomic Energy Agency, Vienna International Centre, Vienna, Austria)

Erling Frederiksen
Brüel & Kjær, Sound and Vibration Measurement A/S, Skodsborgvej 307, 2850 Nærum, Denmark

Milton A. Garcés
Infrasound Laboratory, University of Hawaii, Manoa, HI, USA, e-mail: milton@isla.hawaii.edu

H.W. Gierlich
HEAD acoustics GmbH, Herzogenrath, Germany

Norbert Goertz
Institute for Digital Communications, School of Engineering and Electronics, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JL, UK

Masataka Goto
National Institute of Advanced Industrial Science and Technology (AIST), Tokoyo, Japan

Philippe Gournay
Université de Sherbrooke, Sherbrooke, QC, Canada

Jørgen Hald
Brüel & Kjær, Sound & Vibration Measurements A/S, Nærum, Denmark

Joe Hammond
University of Southampton, Southampton, UK

Colin H. Hansen
School of Mechanical Engineering, University of Adelaide, Adelaide, SA 5005, Australia, e-mail: chanson@mecheng.adelaide.edu.au

Uwe Hansen
Department of Physics, Indiana State University, Terre Haute, IN, USA

Sabih I. Hayek
Department of Engineering Science and Mechanics, University Park, PA 16802, USA

Ulrich Heute
Institute for Circuit and System Theory, Faculty of Engineering, Christian-Albrecht, University, Kaiserstr. 2, D-24143 Kiel, Germany

Thomas Hoffmann
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Volker Hohmann
University of Oldenburg, Oldenburg, Germany

Masaaki Honda
Waseda University, Tokyo, Japan

Andrew B. Horner
Department of Computer Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, e-mail: horner@cs.ust.hk

Adrianus J.M. Houtsma
Aircrew Protection Division, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL 36362-0577, USA, e-mail: adrian.houtsma@amedd.army.mil

Finn Jacobsen
Technical University of Denmark, Lyngby, Denmark, e-mail: fja@elektro.dtu.dk

Finn B. Jensen
NATO Undersea Research Centre, La Spezia, Italy, e-mail: jensen@nurc.nato.int

Walter Kellermann
Multimedia Communications and Signal Processing, University Erlanger-Nuremberg, Erlanger, Germany, e-mail: wk@lnt.de

Youngmoo E. Kim
Electrical & Computer Engineering, Drexel University, Philadelphia, PA, USA, e-mail: ykim@drexel.edu

Anssi Klapuri
Institute of Signal Processing, Tampere University of Technology, Korkeakoulunkatu 1, 33720 Tampere, Finland, e-mail: Anssi.Klapuri@tut.fi

Birger Kollmeier
University of Oldenburg, Oldenburg, Germany

Jan Felix Krebber
Institute of Communication Acoustics, Ruhr-University Bochum, Bochum, Germany, e-mail: jan.krebber@rub.de

Christine E. Krohn
ExxonMobil Upstream Research Company, Houston, TX, USA

Sonoko Kuwano
Osaka University, Osaka, Japan

Gerald C. Lauchle
State College, PA, USA

Walter Lauriks
Katholieke Universiteit Leuven, Heverlee, Belgium; Laboratorium voor Akoestick on Thermische Fysica, K.U. Leuven, Leuven, Belgium

Alexis Le Pichon
CEA/DASE/LDG, BP12, 91680 Bruyères-le-Châtel, France

Philippe Leclaire
Université de Bourgogne, Nevers, France; Laboratorium voor Akoestick on Thermische Fysica, K.U. Leuven, Leuven, Belgium

Roch Lefebvre
Université de Sherbrooke, Sherbrooke, QC, Canada

Cuiping Li
Department of Earth and Atmospheric Science, Purdue University, West Lafayette IN 47907, USA

Tapio Lokki
Department of Media Technology, Helsinki University of Technology, Helsinki, Finland, e-mail: Tapio.Lokki@tkk.fi

Aki Vihtori Mäkivirta
Genelec Oy, Iisalmi, Finland

Brian H. Maranda
Defence Research and Development Canada – Atlantic, Dartmouth, NS, Canada

James Mathews
PCB Piezoelectronics Inc., San Clemente, CA, USA

Manfred Mauermann
University of Oldenburg, Oldenburg, Germany

Walter Metzner
Department of Physiological Science, University of California, California, USA

Yasushi Miki
Computer Science, Faculty of Engineering, Takushoku University, Tokyo, Japan

Ben Milner
School of Computing Sciences, University of East Anglia, Norwich, Norfolk, UK

Riikka Möttönen
Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, Helsinki, Finland

Swen Müller
National Institute of Metrology, Xerém, Brazil

Seiichiro Namba
Osaka University, Osaka, Japan

Ramesh Neelamani
ExxonMobil Upstream Research Company, Houston, TX, USA

David E. Norris
BBN Technologies, 1300 N. 17th St., Arlington, VA 22209, USA

Robert L. Nowack
Department of Earth and Atmospheric Science, Purdue University, West Lafayette, IN 47907, USA

Yasuhiro Oikawa
Waseda University, Tokyo, Japan

Kazuo Okanoya
Laboratory for Biolinguistics, Brain Science Institute, Riken, 2-1 Hirosawa, Saitama 351-0198, Japan, e-mail: okanoya@brain.riken.jp

John V. Olson
Geophysical Institute, University of Alaska, Fairbanks, USA

Thorkild Find Pedersen
Brüel & Kjær, Sound & Vibration Measurements A/S, Nærum, Denmark

Ville Pulkki
Department of Signal Processing and Acoustics, Helsinki University of Technology, Helsinki, Finland, e-mail: ville.pulkki@tkk.fi

Robert B. Randall
University of New South Wales, Sydney, Australia

Helmut Riedel
University of Oldenburg, Oldenburg, Germany

Francis Rumsey
Institute of Sound Recording, University of Surrey, Guildford, UK

Masahiko Sakai
Ono Sokki Co., Ltd., 1-16-1 Hakusan, Midori-ku, Yokohama, 226-8507 Japan, e-mail: msakai@onosokki.co.jp

Rebecca Saltzer
ExxonMobil Upstream Research Company, Houston, TX, USA

Mikko Sams
Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, Helsinki, Finland

Lauri Savioja
Department of Media Technology, Helsinki University of Technology, Helsinki, Finland, e-mail: Lauri.Savioja@tkk.fi

Julius O. Smith
Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA 94305, USA, website: http://ccrma.stanford.edu/~jos/

Stefka Stefanova
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Edmund J. Sullivan
EJS Consultants, Portsmouth, RI, USA

David C. Swanson
The Applied Research Laboratory, The Pennsylvania State University, Philadelphia, PA, USA

Curt A.L. Szuberla
Geophysical Institute, University of Alaska, Fairbanks, USA

Hideki Tachibana
Institute of Industrial Science, University of Tokyo, Tokyo, Japan

Yasushi Takano
Department of Mechanical Engineering Research Laboratory, Hitachi, Ltd., Ibaraki, Japan

Ernst Terhardt
Technical University of Munich, Munich. Germany, e-mail: teirhardt@ei.tum.de

Christine Thomas
Department of Earth and Ocean Sciences, University of Liverpool, Liverpool, UK

Ippei Torigoe
Department of Mechanical Engineering and Materials, Science Faculty of Engineering Kumamoto University, Kumamoto, Japan

Stefan Uppenkamp
University of Oldenburg, Oldenburg, Germany

Ronald T. Verrillo (deceased)
Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA

Tuomas Virtanen
Institute of Signal Processing, Tampere University of Technology, Korkeakoulunkatu 1, 33720 Tampere, Finland, e-mail: Tuomas.Virtanen@tut.fi

Stephen Voran
Institute for Telecommunication Sciences, Boulder, CO, USA

Erhard Werner
Tannenweg 16, D 29693 Hademstorf, Germany

Rodney W. Whitaker
Los Alamos National Laboratory, PO Box 1663, Los Alamos, NM 87544, USA

Paul White
University of Southampton, Southampton, UK

Ru-Shan Wu
University of California, Santa Cruz, CA, USA

Xianyn Wu
ExxonMobil Upstream Research Company, Houston, TX, USA

Ning Xiang
School of Architecture and Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA, e-mail: xiangn@rpi.edu

Ichiro Yamada
Aviation Environment Research Center, Airport Environment Improvement Foundation, Tokyo, Japan

Kohei Yamamoto
Kobayasi Institute of Physical Research, Tokyo, Japan

Yoshio Yamasaki
Waseda University, Tokyo, Japan

Nobutoshi Yoshida
Ono Sokki Co. Ltd., 1-16-1 Hakusan, Midori-ku, 226-8507 Yokohama, Japan, e-mail: yoshi@onosokki.co.jp

Udo Zölzer
Helmut Schmidt University, Hamburg, Germany

Handbook of Signal Processing in Acoustics

Volume 2

Edited by

David Havelock

National Research Council, Canada

Sonoko Kuwano

Osaka University, Japan

Michael Vorländer

RWTH Aachen University, Germany

Handbook of Signal Processing in Acoustics

Volume 2

Springer

Editors
David Havelock
National Research Council
Institute for Microstructural Sciences
Acoustics and Signal Processing Group
1200 Montreal Road
Ottawa ON K1A 0R6
Canada

Sonoko Kuwano
Osaka University
Graduate School of Human Sciences
Department of Environmental Psychology
1-2 Yamadaok Suita
Osaka
Japan

Michael Vorländer
RWTH Aachen University
Institute of Technical Acoustics
Aachen
Germany

ISBN: 978-0-387-77698-9 e-ISBN: 978-0-387-30441-0

Library of Congress Control Number: 2008923573

© 2008 Springer Science+Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.
Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Editor Biographies

David Havelock
National Research Council
Institute for Microstructural Sciences
Acoustics and Signal Processing Group
1200 Montreal Road
Ottawa ON K1A 0R6
Canada

Sonoko Kuwano
Osaka University
Graduate School of Human Sciences
Department of Environmental Psychology
1-2 Yamadaok Suita
Osaka
Japan

Michael Vorländer
RWTH Aachen University
Institute of Technical Acoustics
Aachen
Germany

Editorial Board

Last Name	Given Name(s)	Affiliation/Address
Beauchamp	**James W.**	University of Illinois Urbana-Champaign, School of Music, Dept. of ECE, Urbana, IL, USA
Christie	**Douglas R.**	Earth Physics, Research School of Earth Sciences, The Australian National University, Canberra, AUSTRALIA
Deffenbaugh	**Max**	ExxonMobil Research and Engineering Company, Annandale, NJ USA
Elliott	**Stephen J.**	Southampton University, Institute of Sound and Vibration Research, Southampton, ENGLAND
Fastl	**Hugo**	Technische Universitaet Muenchen AG Technische Akustik, MMK Muenchen, GERMANY
Gierlich	**Hans Wilhelm**	HEAD acoustics GmbH Telecom Division, Herzogenrath, GERMANY
Guyader	**Jean-Louis**	Labratory for Vibration Acoustics Villeurbanne, FRANCE
Jacobsen	**Finn**	Technical University of Denmark, Lyngby, DENMARK
Karjalainen	**Matti**	Helsinki University of Technology, FINLAND
Kollmeier	**Birger**	Universitat Oldenburg, Oldenburg, GERMANY
O'Shaughnessy	**Douglas D.**	INRS-EMT (Telecommunications), Montreal, QC, CANADA
Riquimaroux	**Hiroshi**	Doshisha University, Department of Knowledge Engineering & Computer Sciences, Sensory & Cognitive Neuroscience Research Laboratory, Kyotanabe, Kyoto, JAPAN
Sullivan	**Edmund J.**	EJS Consultants Portsmouth, RI, USA

Last Name	Given Name(s)	Affiliation/Address
Suzuki	**Hideo**	1-2-3-S2502, Utase, Mihama-ward, Chiba-city, JAPAN 261-0013
Taroudakis	**Michael**	University of Crete Department of Mathematics, and FORTH, Institute of Applied and Computational Mathematics, Heraklion, GREECE
Ueha	**Sadayuki**	Tokyo Institute of Technology, Director, Precision and Intelligence Lab, Yokohama, JAPAN
Verrillo	**Ronald T.**	(deceased) Institute for Sensory Research, Syracuse, NY, USA
Yamada	**Ichiro**	Aviation Environment Research Center, Airport Environment Improvement Foundation, Tokyo, JAPAN
Yamasaki	**Yoshio**	Waseda University, Graduate School of Global Information and Telecommunication Studies, Saitama, JAPAN
Tohyama	**Mikio**	Waseda University, JAPAN

Preface

Acoustics has a special relationship with signal processing. Many concepts in signal processing arise naturally from our general experience with sound and vibration and, more than in many other fields, acoustics is concerned with the acquisition, analysis, and synthesis of signals. Consequently, there is a rich resource of signal processing expertise within the acoustics community.

There are many excellent reference books devoted to signal processing but the objective of the *Handbook of Signal Processing in Acoustics* is to bring together the signal processing expertise specific to acoustics and to capture the interdisciplinary nature of signal processing within acoustics. It is also hoped that the handbook will promote networking and the interchange of ideas between technical areas in acoustics.

The handbook comprises 104 *Chapters* organized into 17 *Parts*. Each Part addresses a technical area of acoustics, reflecting the general demarcations of specialization within the acoustics community. An expert with broad knowledge of signal processing within their respective technical area was invited to act as a Section Leader for each Part of the handbook. These Section Leaders contributed substantially to the handbook project by helping to define the contents and scope of each chapter, finding an appropriate contributing expert author, and managing the review and revision of material. Collectively with the Editors, they form the Editorial Board for the handbook.

Planned sections on Architectural Acoustics, Nonlinear Acoustics, and Ultrasound are unfortunately omitted from the handbook; nevertheless, the handbook otherwise provides thorough coverage of the field of acoustics and we can hope that possible future editions might include these areas.

The handbook is written from the perspective of acoustics, by acousticians with signal processing expertise. Emphasis is placed

in the description of acoustic problems and the signal processing related to their solutions. The reader is assumed to have basic knowledge of signal processing. Signal processing techniques are described but the reader is referred elsewhere for derivations and details.

The authors were not required to adhere to strict standards of style or notation, and were asked to prepare short, concise, self-sufficient chapters. This results in variations in style and notation throughout the handbook that reflects the diversity of perspectives within the acoustics community.

David Havelock

Acknowledgments

David Havelock gratefully acknowledges support from the Institute for Microstructural Sciences, of the National Research Council of Canada, for making time and resources available during the preparation of this handbook. He also thanks John C. Burgess (University of Hawaii, retired) for the encouragement to begin this project and expresses his gratitude to the co-Editors, Sonoko Kuwano and Michael Vorländer, who were a pleasure to work with throughout this project. The efforts of all of the Section Leaders are greatly appreciated and we thank each of them sincerely for their patience, perseverance, and faith that were required to see this project to completion. We thank each of the authors for their contributions.

Contents

Editor Biographies v
Editorial Board vii
Preface ix
Acknowledgments xi
Contributors xvii

PART VIII NOISE 931

51 Sound Levels of Noise Sources 933

52 Outdoor Measurements 949

53 Measurements of Physical Characteristics Using Sound at Audible or Lower Frequencies 983

54 Community Response to Noise 1005

PART IX ACTIVE NOISE CONTROL 1021

Introduction 1023

55 Forward Control in Single-Channel Systems 1029

56 Forward Control in Multichannel Systems 1057

57 Feedback Control Systems 1081

PART X STRUCTURAL ACOUSTICS AND VIBRATION 1107

58 Intensity Techniques 1109

59 Nearfield Acoustical Holography 1129

60 Actuator Design and Deployment 1141

61 Materials Testing 1167

PART XI BIORESPONSE TO VIBRATION 1183

62 Tactile Responses to Vibration 1185

63 Uses and Applications of Somatosensation 1215

64 Sensory Substitution 1231

PART XII TRANSDUCERS FOR ACOUSTIC MEASUREMENT 1245

65 Condenser Microphones 1247

66 Dynamic Pressure Microphones 1267

67 Dynamic Pressure Difference Microphones 1275

68 The Microflown Particle Velocity Sensor 1283

69 Microphone Calibration 1293

70 Piezoceramic Accelerometer 1313

71 Laser Doppler Velocimeter 1329

72 Capacitive Sensors 1339

PART XIII ENGINEERING ACOUSTICS 1347

73 Energy Saving Transducers and Thermoacoustics 1349

74 Sinusoidal Representation of Acoustic Signals 1365

75 Room Transfer Function 1381

76 Spatial Information of Sound Fields 1403

PART XIV INFRASONICS 1423

77 Infrasound Event Detection Using the Progressive Multi-Channel Correlation Algorithm 1425

78 Ray Tracing in an Inhomogeneous Atmosphere with Winds 1437

79 Infrasonic Data Processing at the Provisional Technical Secretariat of the Comprehensive Nuclear-Test-Ban Organization 1461

80 The Design and Performance of Infrasound Noise-Reducing Pipe Arrays 1473

81 Processing Infrasonic Array Data 1487

82 Infrasound Propagation 1497

PART XV SEISMOLOGY 1521

83 Structure and Elastic Properties of the Earth 1523

84 Seismic Wave Propagation 1535

85 Seismic Data Acquisition 1545

86 Signal Models in Seismic Processing 1559

87 Deconvolution and Optimal Filtering in Seismology 1571

88 Seismic Migration Imaging 1585

89 Geophysical Parameter Estimation 1593

90 Earthquake Localization 1627

91 Seismic Tomography 1635

92 Array Signal Processing 1655

PART XVI UNDERWATER AND OCEANOGRAPHIC ACOUSTICS 1667

93 Propagation and Signal Modeling 1669

94 Statistical Signal Processing 1695

95 Inverse Problems in Underwater Acoustics 1723

96 Active Sonar 1737

97 Passive Sonar 1757

98 Transducers for Sonar Systems 1783

99 Flow Noise 1821

PART XVII ANIMAL BIOACOUSTICS 1833

100 Bat Bioacoustics 1835

101 Fish Bioacoustics 1851

102 Amphibian Bioacoustics 1861

103 Avian Bioacoustics 1887

104 Cetacean Bioacoustics with Emphasis on Recording and Monitoring 1897

Index 1909

Cumulative Contents 1931

Contributors

Tomonari Akamatsu
National Research Institute of Fisheries Engineering, Fisheries Research Agency of Japan, Kamisu, Ibaraki, Japan

Benoit Alcoverro
CEA/DASE, BP 12, 91680 Bruyères Le Châtel, France

Lydia Ayers
Computer Science Department, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, e-mail: layers@cs.ust.hk

Juha Reinhold Backman
Nokia, Espoo, Finland

Rolf Bader
Institute of Musicology, University of Hamburg, Hamburg, Germany

James W. Beauchamp
School and Music and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 2136 Music Building, 1114 West Nevada Street, Urbana, IL 61801, USA, e-mail: jwbeauch@uiuc.edu

Kim Benjamin
Naval Undersea Warfare Center, Newport, RI, USA

A.J. Berkhout
Delft University of Technology, Delft, The Netherlands

Jeff Bilmes
Department of Electrical Engineering, University of Washington, Seattle, WA, USA

Ole-Herman Bjor
Norsonic AS, Lierskogen, Norway, e-mail: ohbjor@norsonic.com

Stanley J. Bolanowski (deceased)
Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA

Baris Bozkurt
Izmir Institute of Technology (IYTE), Izmir, Turkey

Nicolas Brachet
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Thomas Brand
University of Oldenburg, Oldenburg, Germany

David J. Brown
Geoscience Australia, Canberra, Australia, (Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria), e-mail: David.Brown@CTBTO.org

John C. Burgess
Department of Mechanical Engineering, University of Hawaii, 2540 Dole Street, Honolulu, HI 968221, USA, e-mail: jcb@hawaii.edu

Paola Campus
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Yves Cansi
CEA/DASE/LDG, BP12, 91680 Bruyères-le-Châtel, France

Josef Chalupper
Siemens Audiological Engineering Group, Gebbertstrasse 125, Erlangen 91058, Germany, e-mail: Josef.Chalupper@siemens.com

N. Ross Chapman
University of Victoria, Victoria BC, Canada

Jakob Christensen-Dalsgaard
Institute of Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, e-mail: jcd@biology.sdu.dk

Peter Daniel
Brüel & Kjær, GmbH, Bremen, Germany

Roger B. Dannenberg
Carnegie Mellon University, Pittsburgh, PA, USA

Torsten Dau
Technical University of Denmark, Lyngby, Denmark, e-mail: tda@elektro.dtu.dk

Hans-Elias de Bree
Microflown Technologies, The Netherlands, USA

Max Deffenbaugh
ExxonMobil Research and Engineering Company, Annandale, NJ, USA

Thierry Dutoit
Faculte Polytechnique de Mons, Mons, Belgium

Stephen J. Elliott
Institute of Sound and Vibration Research, Southampton, UK

Paulo A.A. Esquef
Nokia Institute of Technology, Rod. Torquato Tapajós, 7200, Tarumã 69048-660 Manaus-AM, Brazil, e-mail: paulo@esquef@indt.org.br

Richard R. Fay
Parmly Hearing Institute, Loyola University Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, USA, e-mail: rfay@luc.edu

Michael Fehler
Massachusetts Institute of Technology, Cambridge, USA

Sandy Fidell
Fidell Associates, Inc., Woodland Hills, CA, USA

Petr Firbas
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria (International Atomic Energy Agency, Vienna International Centre, Vienna, Austria)

Erling Frederiksen
Brüel & Kjær, Sound and Vibration Measurement A/S, Skodsborgvej 307, 2850 Nærum, Denmark

Milton A. Garcés
Infrasound Laboratory, University of Hawaii, Manoa, HI, USA, e-mail: milton@isla.hawaii.edu

H.W. Gierlich
HEAD acoustics GmbH, Herzogenrath, Germany

Norbert Goertz
Institute for Digital Communications, School of Engineering and Electronics, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JL, UK

Masataka Goto
National Institute of Advanced Industrial Science and Technology (AIST), Tokoyo, Japan

Philippe Gournay
Université de Sherbrooke, Sherbrooke, QC, Canada

Jørgen Hald
Brüel & Kjær, Sound & Vibration Measurements A/S, Nærum, Denmark

Joe Hammond
University of Southampton, Southampton, UK

Colin H. Hansen
School of Mechanical Engineering, University of Adelaide, Adelaide, SA 5005, Australia, e-mail: chanson@mecheng.adelaide.edu.au

Uwe Hansen
Department of Physics, Indiana State University, Terre Haute, IN, USA

Sabih I. Hayek
Department of Engineering Science and Mechanics, University Park, PA 16802, USA

Ulrich Heute
Institute for Circuit and System Theory, Faculty of Engineering, Christian-Albrecht, University, Kaiserstr. 2, D-24143 Kiel, Germany

Thomas Hoffmann
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Volker Hohmann
University of Oldenburg, Oldenburg, Germany

Masaaki Honda
Waseda University, Tokyo, Japan

Andrew B. Horner
Department of Computer Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, e-mail: horner@cs.ust.hk

Adrianus J.M. Houtsma
Aircrew Protection Division, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL 36362-0577, USA, e-mail: adrian.houtsma@amedd.army.mil

Finn Jacobsen
Technical University of Denmark, Lyngby, Denmark, e-mail: fja@elektro.dtu.dk

Finn B. Jensen
NATO Undersea Research Centre, La Spezia, Italy, e-mail: jensen@nurc.nato.int

Walter Kellermann
Multimedia Communications and Signal Processing, University Erlanger-Nuremberg, Erlanger, Germany, e-mail: wk@lnt.de

Youngmoo E. Kim
Electrical & Computer Engineering, Drexel University, Philadelphia, PA, USA, e-mail: ykim@drexel.edu

Anssi Klapuri
Institute of Signal Processing, Tampere University of Technology, Korkeakoulunkatu 1, 33720 Tampere, Finland, e-mail: Anssi.Klapuri@tut.fi

Birger Kollmeier
University of Oldenburg, Oldenburg, Germany

Jan Felix Krebber
Institute of Communication Acoustics, Ruhr-University Bochum, Bochum, Germany, e-mail: jan.krebber@rub.de

Christine E. Krohn
ExxonMobil Upstream Research Company, Houston, TX, USA

Sonoko Kuwano
Osaka University, Osaka, Japan

Gerald C. Lauchle
State College, PA, USA

Walter Lauriks
Katholieke Universiteit Leuven, Heverlee, Belgium; Laboratorium voor Akoestick on Thermische Fysica, K.U. Leuven, Leuven, Belgium

Alexis Le Pichon
CEA/DASE/LDG, BP12, 91680 Bruyères-le-Châtel, France

Philippe Leclaire
Université de Bourgogne, Nevers, France; Laboratorium voor Akoestick on Thermische Fysica, K.U. Leuven, Leuven, Belgium

Roch Lefebvre
Université de Sherbrooke, Sherbrooke, QC, Canada

Cuiping Li
Department of Earth and Atmospheric Science, Purdue University, West Lafayette IN 47907, USA

Tapio Lokki
Department of Media Technology, Helsinki University of Technology, Helsinki, Finland, e-mail: Tapio.Lokki@tkk.fi

Aki Vihtori Mäkivirta
Genelec Oy, Iisalmi, Finland

Brian H. Maranda
Defence Research and Development Canada – Atlantic, Dartmouth, NS, Canada

James Mathews
PCB Piezoelectronics Inc., San Clemente, CA, USA

Manfred Mauermann
University of Oldenburg, Oldenburg, Germany

Walter Metzner
Department of Physiological Science, University of California, California, USA

Yasushi Miki
Computer Science, Faculty of Engineering, Takushoku University, Tokyo, Japan

Ben Milner
School of Computing Sciences, University of East Anglia, Norwich, Norfolk, UK

Riikka Möttönen
Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, Helsinki, Finland

Swen Müller
National Institute of Metrology, Xerém, Brazil

Seiichiro Namba
Osaka University, Osaka, Japan

Ramesh Neelamani
ExxonMobil Upstream Research Company, Houston, TX, USA

David E. Norris
BBN Technologies, 1300 N. 17th St., Arlington, VA 22209, USA

Robert L. Nowack
Department of Earth and Atmospheric Science, Purdue University, West Lafayette, IN 47907, USA

Yasuhiro Oikawa
Waseda University, Tokyo, Japan

Kazuo Okanoya
Laboratory for Biolinguistics, Brain Science Institute, Riken, 2-1 Hirosawa, Saitama 351-0198, Japan, e-mail: okanoya@brain.riken.jp

John V. Olson
Geophysical Institute, University of Alaska, Fairbanks, USA

Thorkild Find Pedersen
Brüel & Kjær, Sound & Vibration Measurements A/S, Nærum, Denmark

Ville Pulkki
Department of Signal Processing and Acoustics, Helsinki University of Technology, Helsinki, Finland, e-mail: ville.pulkki@tkk.fi

Robert B. Randall
University of New South Wales, Sydney, Australia

Helmut Riedel
University of Oldenburg, Oldenburg, Germany

Francis Rumsey
Institute of Sound Recording, University of Surrey, Guildford, UK

Masahiko Sakai
Ono Sokki Co., Ltd., 1-16-1 Hakusan, Midori-ku, Yokohama, 226-8507 Japan, e-mail: msakai@onosokki.co.jp

Rebecca Saltzer
ExxonMobil Upstream Research Company, Houston, TX, USA

Mikko Sams
Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, Helsinki, Finland

Lauri Savioja
Department of Media Technology, Helsinki University of Technology, Helsinki, Finland, e-mail: Lauri.Savioja@tkk.fi

Julius O. Smith
Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA 94305, USA, website: http://ccrma.stanford.edu/~jos/

Stefka Stefanova
Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria

Edmund J. Sullivan
EJS Consultants, Portsmouth, RI, USA

David C. Swanson
The Applied Research Laboratory, The Pennsylvania State University, Philadelphia, PA, USA

Curt A.L. Szuberla
Geophysical Institute, University of Alaska, Fairbanks, USA

Hideki Tachibana
Institute of Industrial Science, University of Tokyo, Tokyo, Japan

Yasushi Takano
Department of Mechanical Engineering Research Laboratory, Hitachi, Ltd., Ibaraki, Japan

Ernst Terhardt
Technical University of Munich, Munich. Germany, e-mail: teirhardt@ei.tum.de

Christine Thomas
Department of Earth and Ocean Sciences, University of Liverpool, Liverpool, UK

Ippei Torigoe
Department of Mechanical Engineering and Materials, Science Faculty of Engineering Kumamoto University, Kumamoto, Japan

Stefan Uppenkamp
University of Oldenburg, Oldenburg, Germany

Ronald T. Verrillo (deceased)
Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA

Tuomas Virtanen
Institute of Signal Processing, Tampere University of Technology, Korkeakoulunkatu 1, 33720 Tampere, Finland, e-mail: Tuomas.Virtanen@tut.fi

Stephen Voran
Institute for Telecommunication Sciences, Boulder, CO, USA

Erhard Werner
Tannenweg 16, D 29693 Hademstorf, Germany

Rodney W. Whitaker
Los Alamos National Laboratory, PO Box 1663, Los Alamos, NM 87544, USA

Paul White
University of Southampton, Southampton, UK

Ru-Shan Wu
University of California, Santa Cruz, CA, USA

Xianyn Wu
ExxonMobil Upstream Research Company, Houston, TX, USA

Ning Xiang
School of Architecture and Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA, e-mail: xiangn@rpi.edu

Ichiro Yamada
Aviation Environment Research Center, Airport Environment Improvement Foundation, Tokyo, Japan

Kohei Yamamoto
Kobayasi Institute of Physical Research, Tokyo, Japan

Yoshio Yamasaki
Waseda University, Tokyo, Japan

Nobutoshi Yoshida
Ono Sokki Co. Ltd., 1-16-1 Hakusan, Midori-ku, 226-8507 Yokohama, Japan, e-mail: yoshi@onosokki.co.jp

Udo Zölzer
Helmut Schmidt University, Hamburg, Germany

PART I
ACOUSTIC SIGNALS AND SYSTEMS

Finn Jacobsen

Technical University of Denmark, Lyngby, Denmark

1 Signals and Systems

Joe Hammond and Paul White 3

Temporal Signal Classification • System Definition and Classification

2 Acoustic Data Acquisition

David C. Swanson 17

Electronic Noise • Analog-to-Digital Conversion • Anti-aliasing Filters • Basic Digital Signal Calibration • Multichannel System Calibration • Data Validation and Diagnostic Techniques Summary

3 Spectral Analysis and Correlation

Robert B. Randall 33

Introduction • The Fourier Transform and Variants • Practical FFT Analysis • Spectral Analysis Using Filters • Correlation Functions • Time–Frequency Analysis

4 The FFT and Tone Identification

John C. Burgess 53

Introduction • The FFT (and DFT) • Leakage • Windows • Estimation of Amplitude, Frequency, and Phase • Applications

5 Measuring Transfer-Functions and Impulse Responses

Swen Müller 65

Introduction • Measurement Methods • Excitation Signals

6 Digital Sequences

Ning Xiang 87

Introduction • Golay Codes • Binary Maximum-Length Sequences • Fast MLS Transform (FMT) • Gold and Kasami Sequences • Legendre Sequences • Application Remarks

7 Filters

Ole-Herman Bjor 107

Introduction • Frequency Weightings • Octave- and Fractional-Octave-Bandpass Filters • Phase Linear Filters • In-Phase/Quadrature Filters • Perceptual Masking Filters

8 Adaptive Processing

Thorkild Find Pedersen 125

Introduction • The LMS Algorithm • Adaptive Applications

9 Beamforming and Wavenumber Processing

Jørgen Hald 131

Planar Nearfield Acoustical Holography • Beamforming

1

Signals and Systems

Joe Hammond and Paul White

University of Southampton, Southampton, UK

Signal processing is the science of applying transformations to measurements, to facilitate their use by an observer or a computer, and digital signal processing (DSP) is the enabling technology for applications across all disciplines and sectors.

The study of sound and vibration is highly dependent on the use of special-purpose signal analysers or software packages. The accessibility and convenience of DSP analysis modules and procedures can sometimes create a deceptive air of simplicity in often complicated phenomena. It is important that practitioners, while availing themselves of the full range of DSP capabilities, should have a clear understanding of the fundamentals of the science of signal processing and so be fully aware of the assumptions, implications and limitations inherent in their analysis methods.

Signal processing and analysis involves the three phases of data acquisition, processing and interpretation (of the results of the processing) and, of course, all three are linked in any application. The last phase is naturally very much related to the subject under investigation, but the first two may be discussed independently of specific applications. A vast body of theory and methodology has been built up as a consequence of the problems raised by the need for data analysis; this is often referred to as "signal analysis" or "time-series analysis" depending on the context [1]. The choice of methodology is often reliant on some prior knowledge of the phenomenon being analysed. This usually relates to classifying the characteristics of the data and/or the way in which the data may be modelled owing to knowledge of (or assumptions about) the way in which the data may have been generated. This section

considers signal and system characteristics that underpin signal processing.

In acoustics the signal in question is generally the output of a pressure transducer. Such a signal is a time history that depends on the spatial location of the transducer. We denote this by $p(t,\mathbf{r})$ where t denotes the time dependence and $\mathbf{r}$ is the vector representation for its spatial location (with Cartesian co-ordinates x, y, z). Multiple signals may be available from an array of sensors as required, e.g., in beamforming. For the present it is convenient to drop the spatial dependence and consider signals as evolving with time.

1 • Temporal Signal Classification

The physical phenomenon under investigation is often translated by a transducer into an electrical equivalent, and a single signal evolving in continuous time is denoted as $x(t)$. In many cases, data are *discrete* owing to some inherent or imposed sampling procedure. In this case the data might be characterised by a sequence of numbers. When derived from the continuous time process $x(t)$ we write $x(n\Delta)$ or x_n $(n = 0, 1, 2, \ldots)$, where we have implied that the sampling interval Δ is constant (i.e. uniform sampling). The time histories that can occur are often very complex, and it is helpful to consider signals that exhibit particular characteristics. This allows us to relate appropriate analysis methods to those specific types of signal.

Figure 1 illustrates a broad categorisation of signal types.

A basic distinction is the designation of a signal as "random" or "deterministic" where by "random" we mean one that is not

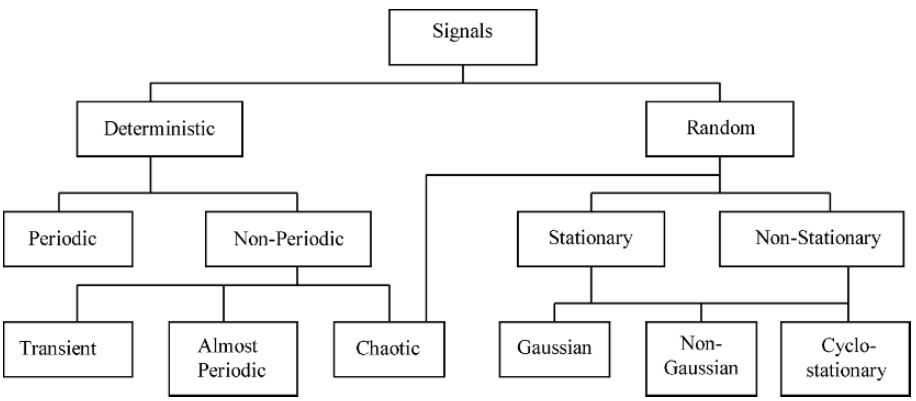


FIGURE 1 Classification of signals.

exactly predictable. Very often, processes are mixed and the demarcations shown in Figure 1 are not easily applied and consequently the analysis procedure to be used may not be apparent. We have included chaotic processes under both "deterministic" and "random" categories since such signals are generated from a deterministic non-linear phenomenon but nevertheless have an output with an unpredictable, random-like behaviour. The classification of data as being deterministic or random might be debatable in many cases and the choice must be made on the basis of the knowledge of the physical situation. Often signals may be modelled as being a mixture of both, e.g. a deterministic signal "embedded" in unwanted random disturbances (noise).

Many of the classes defined above, such as the stationary and periodic signals, are mathematical constructs to which no real-world signal can belong. However, these classes do provide one with a set of models which, in many cases, provide good approximations to measured processes and they suggest a suitable analysis framework.

A brief note for each of the categories is given below – continuous time is used throughout. An analysis method appropriate for each is also noted.

1.1 Periodic

A signal is periodic with period T_p if $x(t) = x(t + T_p)$.

Such a signal can be represented in the frequency domain as a Fourier series:

$$x(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(2\pi k f_0 t) + b_k \sin(2\pi k f_0 t) = \sum_{k=\infty}^{\infty} c_k e^{2\pi i k f_0 t}. \quad (1)$$

The feature of this is that the frequencies in representation are the fundamental, $f_0 = 1/T_p$, and multiples (harmonics) (plus a d.c. term), and the coefficients, a_k and b_k (or c_k), give the amplitude and phase of the components.

Figure 2a shows the time-series of a small section of voiced speech, the vowel /e/. The approximately periodic character of this signal is evident and suggests that it could be gainfully analysed by computing its Fourier series representation. This Fourier series is shown in Figure 2b, which has been computed based on the first period of the signal shown in Figure 2a. One should appreciate that the assumption of periodicity does not hold exactly for this example, as is commonly the case.

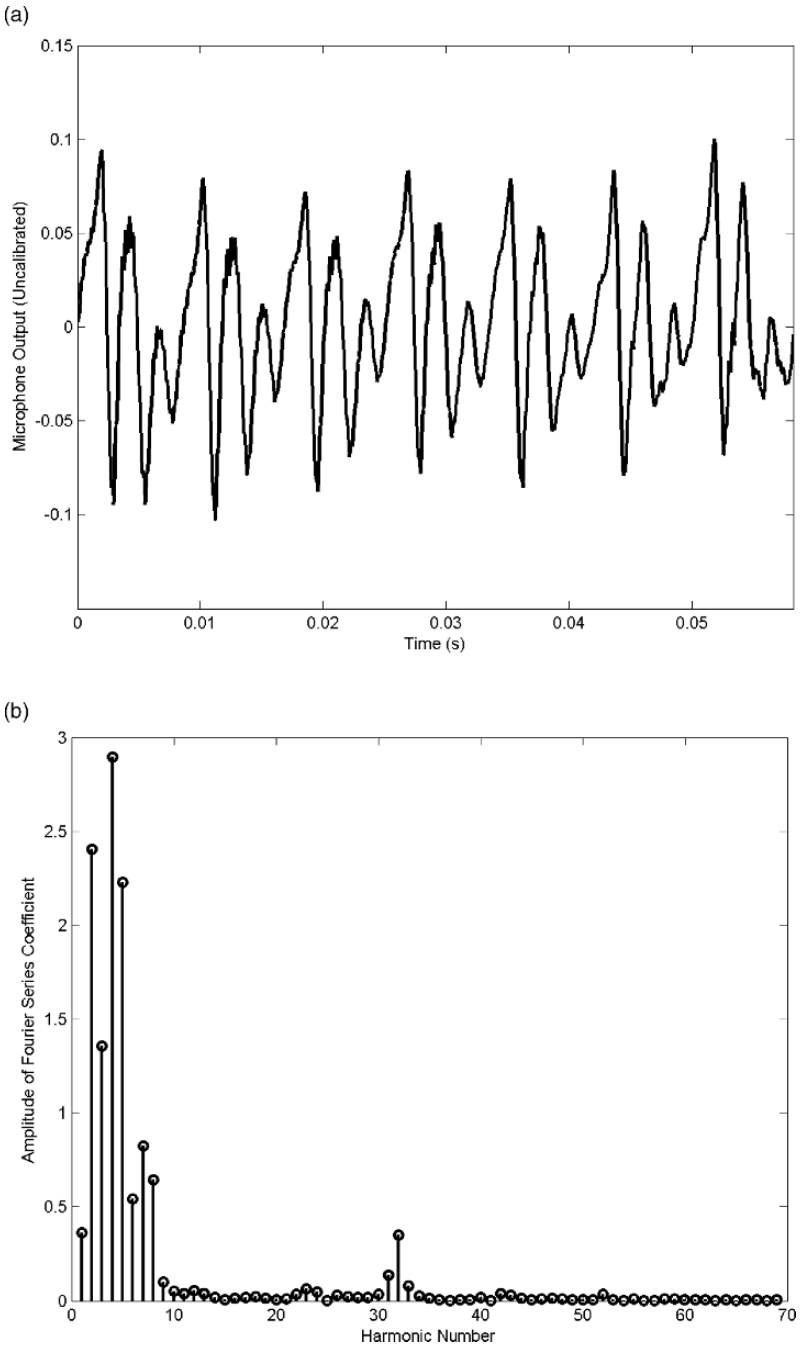


FIGURE 2 Example of voiced speech: (**a**) time-series, (**b**) Fourier coefficients.

1.2 Transient

A transient signal is one which is essentially localised, i.e. a signal that has a finite duration. Such a signal has a Fourier integral representation:

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{2\pi i f t} df. \tag{2}$$

This differs from the periodic case in that the frequency range becomes a continuum and amplitudes "in a band" are $X(f)df$ so $X(f)$ is now an amplitude density.

Figure 3a shows an example of the time-series of an unvoiced speech segment, in this case the sound /th/. The lack of a periodic structure in this signal is apparent. It can be considered as a transient signal and analysed by employing the Fourier transform defined in (2), which is shown in Figure 3b.

1.3 Almost Periodic

This could be regarded as a process where T_p in (1) varies with time. There is no natural generalisation of (1) for this.

An alternative way of constructing a simple almost periodic process is, e.g.,

$$x(t) = \sin(2\pi t) + \sin\left(2\pi\sqrt{2}t\right). \tag{3}$$

This process whilst being the sum of two periodic signals is itself not periodic because the ratio of the two frequency components is not a rational number. In this case using the Fourier integral (2) is appropriate.

Figure 4a shows the time-series of a segment of a sustained piano note. These notes contain discrete frequencies that are approximately harmonically related. The inharmonicities present in piano string vibrations mean that such signals should be regarded as almost periodic and so the use of the Fourier transform (2) represents a suitable tool for analysis; the results of such an analysis are shown in Figure 4b.

1.4 Chaotic Process

The range of phenomena that lead to chaotic dynamics is wide, but a feature is that the describing equation is deceptively simple in its deterministic form, but generates intricate signals. An example is a second-order non-linear system driven by a sinusoidal excitation. The methodology of analysis uses topological feature analysis, by which one attempts to reveal the underlying simple generation mechanism.

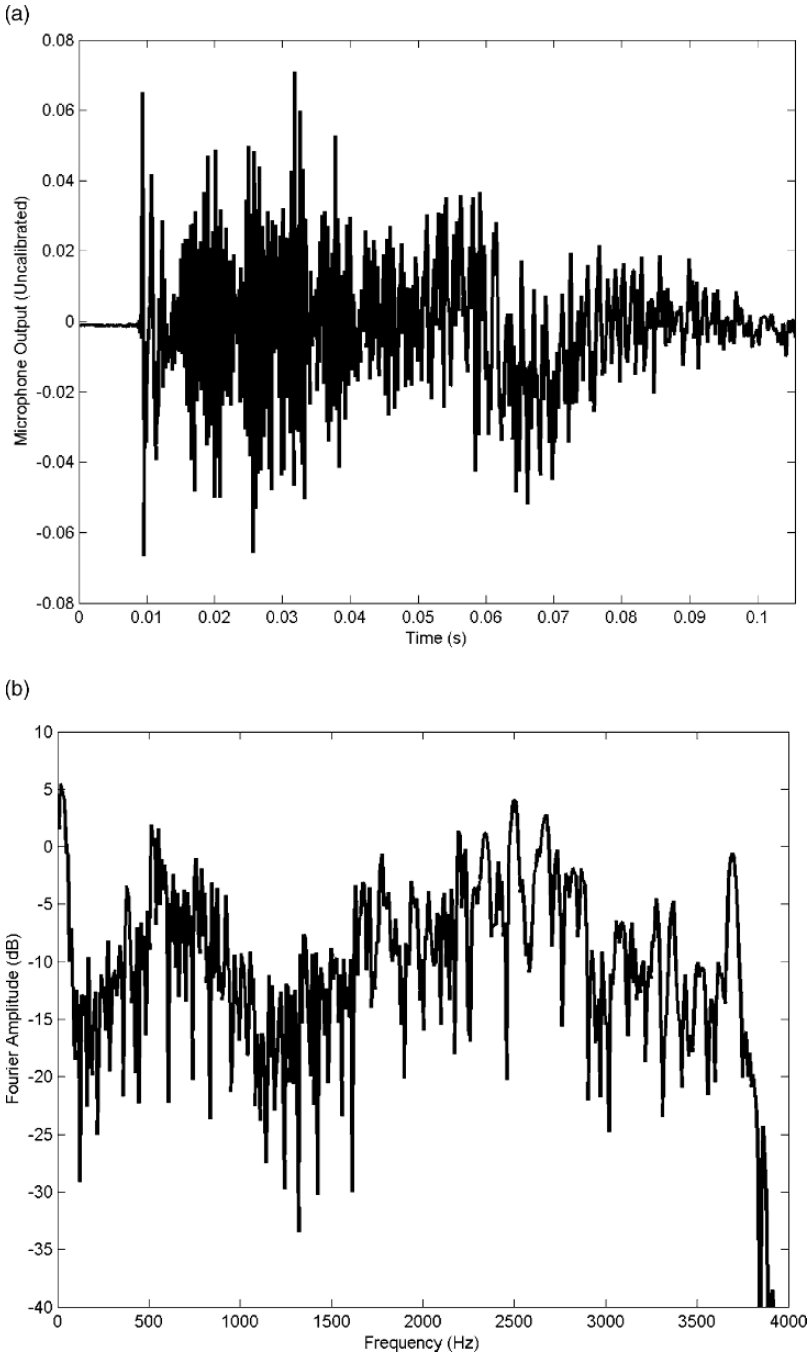


FIGURE 3 Example of unvoiced speech: (**a**) time-series, (**b**) magnitude of the Fourier transform.

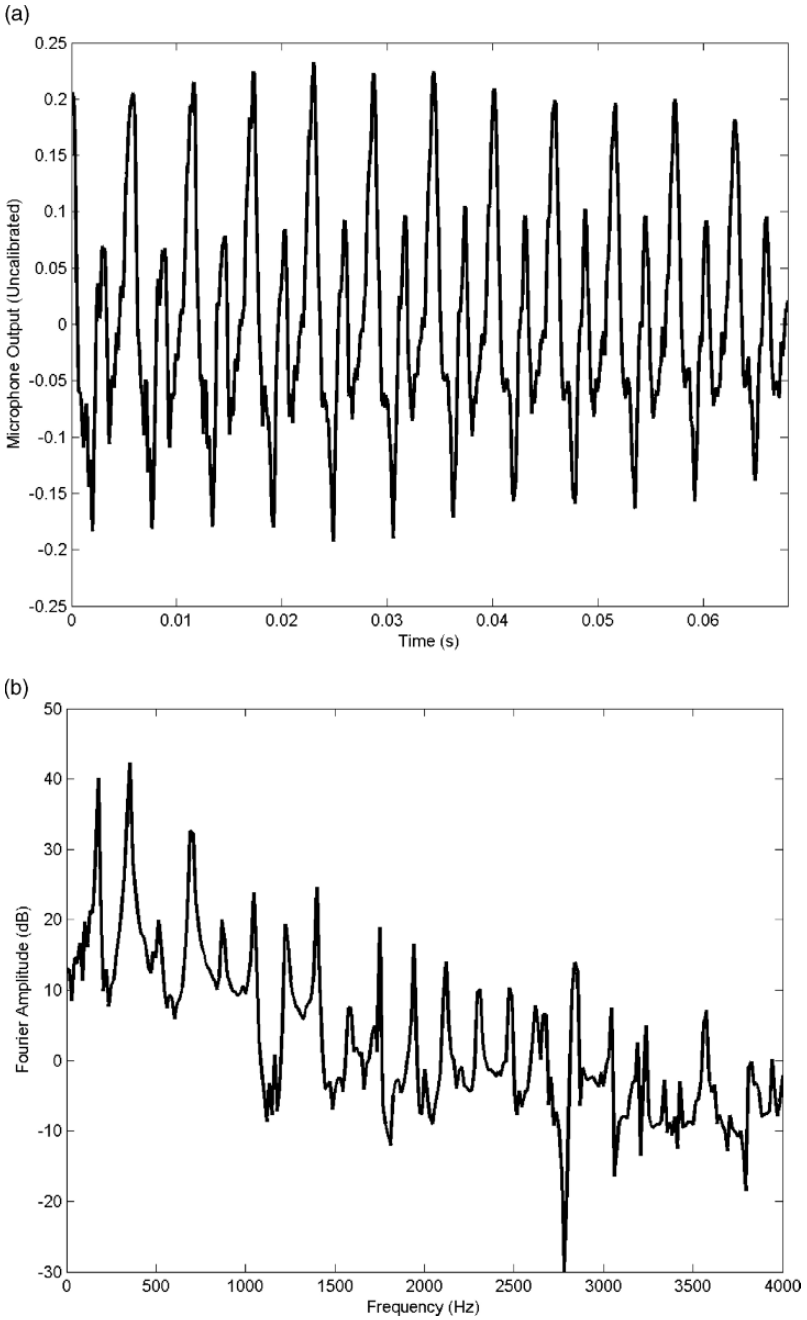


FIGURE 4 Example of a piano note: (**a**) time-series, (**b**) Fourier transform.

1.5 Stationary Random Processes

A stationary random process is one for which the statistical (probabilistic) structure is invariant under a shift in time. Strict stationarity requires that the joint probability density function

$$p_k(x_1, t_1; x_2, t_2; \ldots; x_k, t_k) = p_k(x_1, t_1 + T; x_2, t_2 + T; \ldots; x_k, t_k + T) \quad \text{for all } t_k. \tag{4}$$

This very strong condition is often relaxed to its application to $k = 1$ and 2 only, i.e. stationary to second order. This is referred to as weak (or wide-sense) stationarity.

Stationarity can only be approximate, and judging when the assumption is valid relates to this specific application. The assumption of weak stationarity readily allows the study of the process in the frequency domain using the power (auto) spectral density and the cross-spectral density. In many applications signals are processed by invoking an assumption of stationarity over short timescales.

The autocorrelation function of a stationary signal $x(t)$ is the average of the product of $x(t)$ with a time-shifted version of itself, i.e. the average of $x(t)x(t+\tau)$, and is written as $R_{xx}(\tau)$. It is related to the power spectral density $S_{xx}(f)$ of $x(t)$ by

$$S_{xx}(f) = \int R_{xx}(\tau)\, \mathrm{e}^{-2\pi i f \tau} \mathrm{d}\tau. \tag{5}$$

$S_{xx}(f)$ describes the decomposition of the power of the signal over frequency.

1.6 A Non-stationary Random Process

A non-stationary random signal is one for which p_k is not the same under a time shift.

It should be noted that the term "non-stationary" is now also used to denote any signal in which the "structure" changes as a function of time but which may not be immediately classified as random. An example is a deterministic chirp signal – for such signals the concept of a time–frequency analysis is often used.

For both classes of non-stationarity there are several methods that attempt to generalise the spectral representations [2]. Figure 5 shows an analysis of a section of the song of a nightingale (*Luscinia megarhynchos*).

1.7 Cyclo-stationary Processes

Cyclostationary processes are a class of non-stationary process for which there is an underlying periodic structure. The cyclo-stationary processes include periodic signals as a special case. An example of a more general cyclostationary process is a non-stationary process which is periodically correlated. A simple example of a cyclostationary process is modulated noise, e.g.

$$x(t) = \sin(2\pi f t)\, w(t), \tag{6}$$

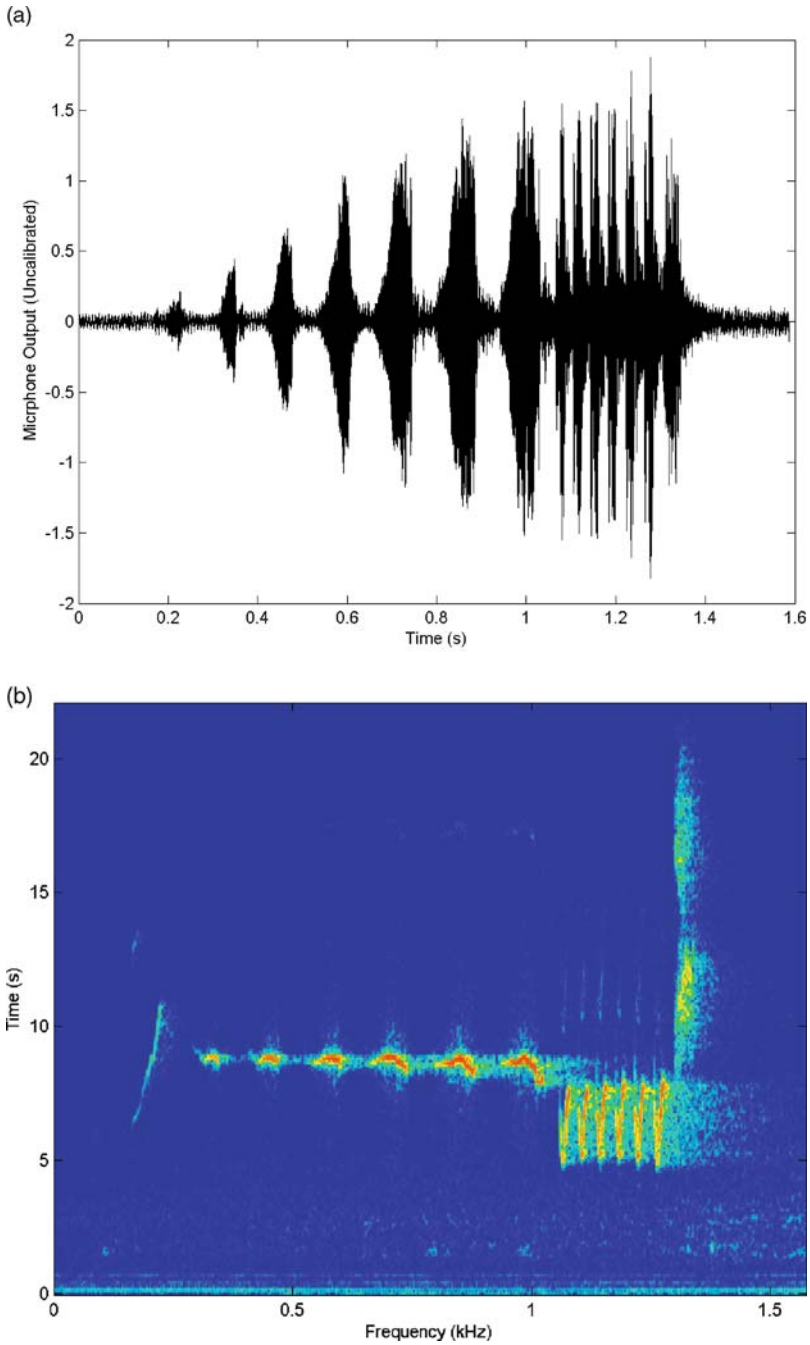


FIGURE 5 Example of a segment of the call of a nightingale: (**a**) time-series, (**b**) short-time Fourier transform.

where $w(t)$ is a broadband random process, with unit variance. The signal $x(t)$ has a structure (in this case a variance) which is periodic, whilst the signal itself is not periodic. Since it is the second-order statistics of the signal here which are the first to show periodicity the above signal is said to be second-order cyclostationary. Many acoustic signals associated with rotating systems, such as fans, contain components which are cyclostationary.

2 • System Definition and Classification

In signal processing the concept of a "system" is used alongside that of "signal". This may be because we wish to relate two (or more) signals that occur in the analysis of a system, or it is often convenient to regard a signal as having arisen as the output of a system (which may be "real" or simply a convenient model). This section describes some types of system.

2.1 Linear Systems

Figure 6 depicts a single input–single output system.

The link between input $x(t)$ and output $y(t)$ is linear if the concept of superposition applies – i.e. if $y_1(t)$ is the response to $x_1(t)$ and $y_2(t)$ is the response to $x_2(t)$ then the response to $ax_1(t)+bx_1(t)$ is $ay_1(t)+by_2(t)$.

Systems are often described by their response to a particular input – this provides both a characterisation of the system and the basis of how the response to more complex inputs may be built up. A fundamental description of a linear system is the response to a unit impulse (delta function). This is the impulse response function denoted $h(t,t_1)$ which is the response at time t due to the application of a unit impulse at time t_1.

Because the linearity of such a system is implied if an impulse of strength a is applied at t_1 and one of strength b is applied at t_2, the response to both is $ah(t,t_1)+bh(t,t_2)$.

The fact that the impulse response is a function of both t and t_1 indicates that the system, whilst being linear, may be *time variable*. The system is called *time invariant* if $h(t,t_1)=h(t-t_1)$, in which case what matters is the duration between t and t_1, not their explicit values.

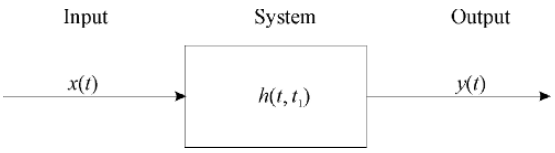


FIGURE 6 Schematic representation of a system.

A system is described as causal if $h(t, t_1) = 0$ for $t_1 > t$; in words the system only responds after the impact is applied. This clearly makes "obvious" sense for physical systems, but need not be so in mathematical models of data where the "past" and the "future" may be interpreted variously in, e.g., interpolation schemes for time histories or image processing.

A system is *stable* if $h(t, t_1)$ dies away "sufficiently quickly" as $t \to \infty$, i.e., the system "forgets" the impulse. There are several mathematical definitions of what constitutes a stable system. The most common definition is provided by the principle of bounded input bounded output (BIBO) stability, which states that a system is stable if for all bounded inputs, so $|x(t)| < M$ for all t, then one can find a value N such that $|y(t)| < N$. This is equivalent to requiring that

$$\int_{-\infty}^{\infty} |h(t, t_1)| \, \mathrm{d}t < \infty \quad \forall t_1. \tag{7}$$

These concepts can be put together for a linear, time-invariant, causal, stable system to yield an expression relating the system's input, $x(t)$, to its output, $y(t)$, as follows:

$$y(t) = \int_{-\infty}^{t} h(t-\tau) x(\tau) \, \mathrm{d}\tau = \int_{0}^{\infty} x(t-\tau) h(\tau) \, \mathrm{d}\tau = h(t) * x(t), \tag{8}$$

where the operator $*$ is referred to as convolution which is defined via the above integral relationships. In the frequency domain the input/output relationship can be represented as

$$Y(f) = H(f) X(f), \tag{9}$$

where the function $H(f)$ is referred to as the transfer function and can be obtained by Fourier transforming the impulse response, $h(t)$.

These concepts extend simply to multi-input multi-output systems. If the system is excited by the P inputs $x_p(t)$ which generated the Q outputs $y_q(t)$ then the system can be characterised by the set of impulse responses (or transfer functions) that relates each of the individual outputs. Specifically

$$y_q(t) = \sum_{p=1}^{P} \int_{0}^{\infty} h_{p,q}(\tau) x_p(t-\tau) \, \mathrm{d}\tau. \tag{10}$$

In the frequency domain this can be represented as a matrix product:

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) \qquad \mathbf{H} = \left\{ H_{p,q}(f) \right\}_{p=1,\ldots,P, q=1,\ldots,Q}, \tag{11}$$

where $\mathbf{x}(t) = [x_1(t) \ldots x_P(t)]^{\mathrm{t}}$ and similarly $\mathbf{y}(t) = [y_1(t) \ldots y_Q(t)]^{\mathrm{t}}$.

2.2 Distributed Systems

In reality the concepts of single input–single output arise simply because we choose to look at signals aiming at single points in some field, e.g., $y(t)$ may be the output of a microphone measuring pressure at some specific spatial position:

$$y(t) = p(t; \mathbf{r}). \tag{12}$$

This means that the bigger picture relates some general excitation field x(t, r) ($x \epsilon X$) to some response field y(t, r) ($y \epsilon Y$) through a generalisation of a convolution, i.e.,

$$y(t, \mathbf{r}) = \int\int h(t - t_1; \mathbf{r}, \mathbf{r}_1)\, x(t_1; \mathbf{r}_1)\, \mathrm{d}\mathbf{r}_1 \mathrm{d}t_1. \tag{13}$$

Every point of the input field X contributes to the response at the individual point $\mathbf{r}$ and this is contained in the spatial (three-dimensional) integral.

The impulse response function (often called the Green's function in this context) is $h(t,t_1;\ \mathbf{r},\mathbf{r}_1)$, and if the system is time variant and spatially homogeneous it can be expressed as $h(t - t_1; \mathbf{r} - \mathbf{r}_1)$.

A classical example of the use of impulse responses of acoustic systems is found in the calculation of the reverberation time for an enclosure. The reverberation time is defined as the time taken for the reverberant energy to decay by a specified number of decibels, in this case 60 dB. There are a variety of methods that can be used to compute the reverberation time. For illustration the decay curve shown in Figure 7c is based on Schroeder integration [3,4], which is defined as a time-reversed integration of the instantaneous power of the signal. The Schroeder integral of an enclosure's impulse response, an example of which is shown in Figure 7a, can be used to compute the reverberation time. Note this is a very large room with an unusually long reverberation time.

2.3 Time-Varying Systems

For systems that are time-varying, and causal, the relationship between the input and output is given by a generalisation of the convolution integral, so that

$$y(t) = \int_{-\infty}^{t} h(t, t_1)\, x(t_1)\, \mathrm{d}t_1. \tag{14}$$

2.4 Non-linear Systems

There is no simple relationship between the input and output of a general non-linear system. However, there are sub-classes of non-linear systems whose responses can be modelled through extensions of the linear convolution relationship. One of the most popular of which is the Volterra expansion:

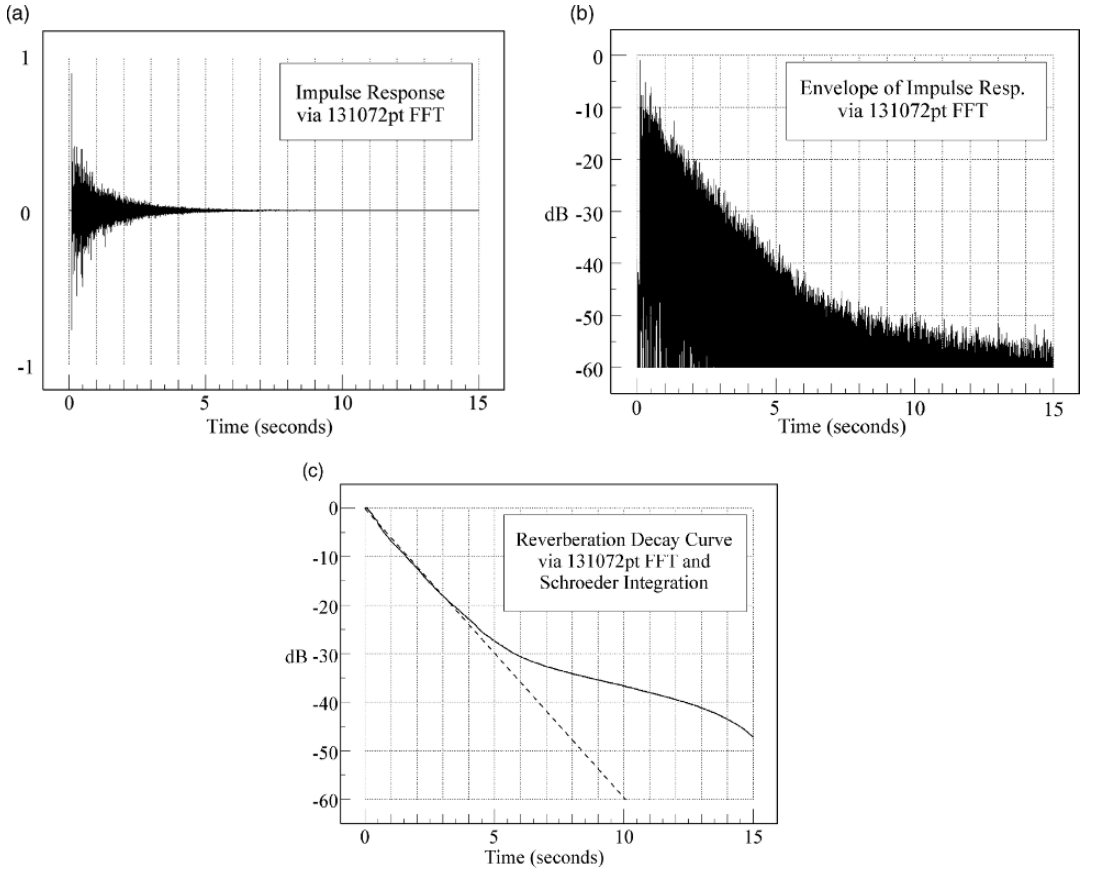


FIGURE 7 Computation of the reverberation time via the Schroeder integral (courtesy of K.R. Holland, ISVR, University of Southampton).

$$y(t) = \int_{-\infty}^{\infty} h_1(\tau)\, x(t-\tau)\, \mathrm{d}\tau + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, \mathrm{d}\tau_1 \mathrm{d}\tau_2 + \cdots + \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_k(\tau_1, \ldots, \tau_k)\, x(t-\tau_1) \cdots x(t-\tau_k)\, \mathrm{d}\tau_1 \cdots \mathrm{d}\tau_k + \cdots \tag{15}$$

This can be regarded as a generalised Taylor series expansion. To characterise the non-linear system one needs to define the set of kernel functions h_k. It is common practice to truncate this expansion at low order, e.g. the second or third terms. Such expansions have been used to describe the non-linear behaviour of loudspeakers [5].

References

1. F.J. Fahy and J.G. Walker (eds), *Advanced Applications in Acoustics in Noise and Vibration*, Spon Press, London (2004).
2. J.K. Hammond and P.R. White, "The Analysis of Non-stationary Signals Using Time Frequency Methods", *Journal of Sound and Vibration*, **190(3)** (1996), 419–447.
3. M.R. Schroeder, "New Method of Measuring Reverberation Time", *Journal of the Acoustical Society of America*, **37(6)** (1965), 1187–1188.
4. M.R. Schroeder "Integrated-Impulse Method Measuring Sound Decay Without Using Impulses", *Journal of the Acoustical Society of America*, **66(2)** (1979), 497–500.
5. A.J.M. Kaizer, "Modelling of the Nonlinear Response of an Electrodynamic Loudspeaker by a Volterra Series Expansion", *Journal of the Audio Engineering Society*, **35(6)** (1987), 421–433.

2

Acoustic Data Acquisition

David C. Swanson

The Applied Research Laboratory, The Pennsylvania State University, Philadelphia, PA, USA

By "acoustic" data we also mean vibrations in structures, internal stress waves, and structure–fluid interactions which involve acoustic radiation. There are a number of acoustic sensors that produce alternating voltage fluctuations that are then sampled at regular time intervals for analysis in the time, spatial, or frequency domains. The most essential part of data acquisition can be seen as objective signal calibration and characterization of signal-to-noise ratio (SNR). Calibration for multichannel systems involves not just signal amplitude, but relative phase or time delay between channels, as well as signal leakage, or cross-talk, between channels. Regardless of how expensive the data acquisition system is, these parameters of calibration, SNR, cross-talk, and relative phase are always an issue, albeit one that good engineering practice can manage.

1 • Electronic Noise

The electronics in the signal path from sensor to data acquisition all contain broadband noise in one form or another, and if the electrical impedance of the signal cables is high, noise interference from other electrical sources can be coupled into the sensor signals. Magnetic induction is very difficult to shield a cable from and can result in microamps of induced currents

from nearby power lines. If the impedance of the signal path to ground in the sensor cable is high, say 50,000 ohms or greater, the signal current corresponding to 1 V RMS is also small (20 μA for 1 V over 50k ohms). For a high impedance signal path, it is very easy to have these weak induced currents appear as hundreds of millivolts along with the signal, especially if the cables are long. However, if one shunts the signal with say a 150 ohm resistor at the data acquisition end (assumed to have a high input impedance), and the sensor has a unity gain amplifier, the milliamps of sensor signal current driving the 150 ohm shunt resistor will be huge compared to the magnetically induced currents (still in the microamp range), and the noise interference disappears. Sometimes these magnetically coupled signals are called "ground loops", but this is actually a slightly different problem, although one also sees harmonics of the power line frequency (60 Hz in the USA, 50 Hz in Europe, etc.) which can be the same symptom.

Ground loops are caused from the sensor being locally grounded at a place some distance from the data acquisition system's ground. This is not only bad practice, but it can be a dangerous electric shock hazard if the sensor is located on equipment with high electric power such as motors, generators, electric heaters, or high-voltage systems. If a ground loop exists and an equipment electrical short occurs, the large short circuit current can make its way into the data acquisition system ground, electrifying anyone touching the chassis grounds as well as damaging the electronics. Removing all ground loops and using a single ground point not only keeps the sensor signals free of interference but also provides an added measure of safety.

If possible, one should listen to the electronic background noise, if not also examining it with an oscilloscope at every junction of the signal path from the sensor to the data acquisition system to identify and isolate grounding problems.

The following guidelines should be followed to avoid magnetic induction and ground loops in the sensor signal path to the data acquisition system:

1. Use a shielded twisted pair rather then a coaxial "BNC" cable, grounding the shield at the data acquisition system and floating the shield at the sensor.
2. Amplify the sensor signal as close to the sensor as possible. Place a resistor (150 to 2k ohm) across the twisted pair *at the data acquisition end* of a long cable run so that the current from the sensor amplifier drives a strong sensor signal current (well beyond 1 mA for a 1 V signal) in the cable relative to magnetically coupled interference, which is often in the microamp or

nano-amp range. This will greatly reduce power line and other electromagnetic interference. Most data acquisition systems have high input impedance, sometimes even in the giga-ohm range. One nano-amp of electromagnetic noise creates significant voltage interference relative to a typical sensor signal voltage when input into such a high impedance. By lowering the cable impedance and amplifying the sensor signal current, the sensor signal dominates the noise interference in the acquired data as desired.

3. Always amplify the sensor signal *before* filtering. Filters are designed for unity gain in the pass band and always have some added residual noise that should not be amplified. Many data acquisition systems have built-in signal amplifiers that may also be amplifying signal conditioner noise along with the sensor signal.
4. Calibration should be designed to be conveniently done in the field on the complete system, from sensor to digital data. This should be repeated before and after each data acquisition use to validate that nothing changed during the recording. The calibration signal should closely match the expected sensor signal in frequency and amplitude and should include the sensor. For example, devices called speakerphones or pistonphones are available for standard microphones for calibration purposes. Accelerometer calibrators are also available to provide a repeatable 1 *g* RMS (one "*g*" is 9.8 m/s^2) acceleration at about 160 Hz. Calibration should be routine and records should be archived along with the sensor data.

2 • Analog-to-Digital Conversion

There are two main classes of analog-to-digital converters (ADC) called successive approximation (SA) and delta–sigma (DS) converters (also known as sigma–delta converters). Because the DS requires no external filters, it has become widespread particularly for audio-bandwidth applications. The type of built-in filtering in the DS converter can introduce signal delays as well as prevent DC coupling. However, the SA converter has an advantage of low latency in that the digital sample is available immediately and that you can select a specific external filter to suit the data acquisition requirements. This is very important for real-time control systems where acquired data must be processed to produce a control signal as rapidly as possible. Filters have a mathematical trade-off between transient response and roll-off in frequency response into the stop band. Sharper roll-off produces longer transient response "ringing". So for applications

requiring real-time control and/or specific filter responses, the SA type of ADC is preferred, even though some DS converters have filters with very short delays. The DS converter typically internally samples the signal using only 1 bit at a very high rate with a modest prefilter, and then digitally filters and down-samples the data, increasing the sample word length during the processing. This processing requires processing a number of samples, thus causing a small latency which is not a problem if all one is doing in recording. Not all DS converters filter the same way, so it is very important to check the manufacturer's specifications in detail.

Digital signals are by definition band limited to a frequency bandwidth less then half the sampling frequency to insure all digital frequencies have corresponding analog frequencies. If a frequency outside the bandwidth defined by the sample rate and type is digitized, it will appear as an *aliased* frequency in the digital signal. The anti-aliasing filters used to limit the analog signal bandwidth prior to sampling generally should begin their roll-off at around one third of the sample rate. They should have a sharp roll-off of 48 dB/octave or greater so that any aliased frequencies are well attenuated in the digital signal. However, aliased frequency residuals can often be seen if there is low background noise and the digital signal has a bit depth (word length) sufficient to detect a wide dynamic range of signals. For 12-bit systems, 48 dB/octave filtering should be sufficient, while for 16- to 24-bit systems, a roll-off of 96 dB/octave is preferred. A good system test is to input a signal at the maximum undistorted amplitude and sweep the frequency from just below 1/2 the sample rate (called the Nyquist frequency) up to the sampling frequency slowly and use a Fourier transform to measure the aliasing filter leakage.

2.1 Successive Approximation

An SA converter uses a counter, a digital-to-analog converter, and a comparator to approximate a digital number output to the analog voltage of the input. A digital-to-analog converter is simply a bank of switches and a summing circuit to add the corresponding voltages. First, the analog voltage input is sampled and held on a capacitor during the digitization process. The voltage comparator signals the counter to count downward when the analog voltage is below the current digital approximation and to count upward when the analog voltage is above the digital approximation. The SA process converges very fast to approximate the sampled analog input voltage latched on the sample-and-hold capacitor.

For 16-bit samples, the number range is 0–65,535 for single-ended and −32,768 to +32,767 (90.3 dB dynamic range) for

bi-polar input voltages. For 18-bit, it is −131,072 to +131,071 (102 dB), and for 24-bit, it is −8,388,608 to +8,388,607 (138 dB) because one binary number is reserved for zero. The maximum input voltage and the bit depth determine an effective "numeric gain" for the analog-to-digital converter of

$$G_{\mathrm{ADC}} = \frac{2^{Q-1}}{V_{\max}}, \tag{1}$$

where Q is the number of bits (for bi-polar voltages) and $V_{\max}$ is maximum input voltage (peak, not RMS). Equation (1) provides a very useful calibration metric for an analog-to-digital conversion system. Dividing the digital samples by this factor converts the integers to volts. Additional calibration factors for amplifier gains and transducer sensitivities can be applied to save the data in calibrated engineering units, which is always preferred unless only limited storage space is available.

2.2 Increasing Bit Depth Through Integration/ Decimation

One can compensate for limited bit depth by sampling at a much higher rate and averaging successive groups of samples into fewer samples (decimation) where each sample has more bits. For example, summing pairs of 8-bit samples together gives 9-bit samples at half the sample rate. The summing operation is also an effective low-pass filter with a zero at the original Nyquist frequency. Suppose we wanted to measure vibrations at 1 kHz and below that had a dynamic range of over 90 dB, but only had a 12-bit SA converter that could sample at rates up to 128 kHz. If we sampled at 12-bits and 128 kHz, we could integrate and decimate to 18-bits at 2 kHz sample rate and capture signals with up to 102 dB dynamic range. One gains 6 dB of dynamic range (1-bit) with each halving of the decimated sample rate. This is a very simple and affordable solution that is often overlooked when specifying a data acquisition system.

2.3 Delta–Sigma Conversion

If we are not interested in DC signals (although some DS converters can handle them), we can do even better than the 6 dB per halving of the sample rate by differentiating the signal samples prior to summing and decimating. This is also called "noise shaping" because we are effectively filtering with a zero at 0 Hz and a +6 dB/octave roll-up across the signal bandwidth [1,2]. Summing the differences suppresses the low-frequency side by another 3 dB, thus giving a net dynamic range gain of 9 dB for each halving of the sample rate. The differencing is the "delta" part and the summing is the "sigma" part of the DS converter algorithm. Manufacturers' literature detail many variations and enhancements on this basic approach. The differences

are seen mainly in the signal latency, phase responses near the Nyquist rate, and distortions that appear for low-level signals. For a DS converter to digitize an audio CD signal at 16-bits and 44.1 kHz, it begins with a 1-bit signal sampled at around 45 MHz and probably low-pass filtered at around 100 kHz and 12 dB/octave roll-off. About 10 "delta–sigma" operations would happen to enhance the bit depth to 16 bits at 44.1 kHz. Additional digital filtering may also be done to sharpen the roll-off near the Nyquist frequency, smooth the phase response to minimize audio distortion, and reduce ringing for transient signals. It is not uncommon for all this process to introduce a net latency of anywhere from 8 to 32 samples at the final decimated rate. This is only a problem for real-time control applications. Delta–sigma converters are quite good for digital recording of acoustic and vibration signals, although the DS sampling filter can cause problem if the analog signal is multiplexed. If a sample rate below around 8 kHz is desired, an external anti-aliasing filter is usually required. Again, since not all DS converters are alike, it is very important to check the specifications of the particular DS converter chip in use because the data acquisition system manufacturers often do not provide such detail.

3 • Anti-aliasing Filters

Figure 1 shows a typical collection of low-pass filter responses. While one might be tempted to always choose a filter type based on the sharpness of the roll-off, selection is even more critical considering stop-band leakage and phase response.

First, we briefly describe the filter types. The filters in Figure 1 are all pole-zero designs with a theoretically infinite impulse response [3]. In practice, filters with a sharper roll-off will tend to "ring" longer in response to a transient or impulsive signal. But as seen in Figure 1, the sharper roll-off comes at the expense of leakage in the stop band, which is unacceptable for anti-aliasing filters. The classic Butterworth filter frequency response is designed by placing the poles on a semicircle on the left-half complex s-plane ($s = \sigma + j\omega$), where the circle crosses the $j\omega$ axis at the positive and negative cut-off frequencies. The Chebyshev type I response places poles on a semiellipse in the left-hand s-plane. This gives a response with a slight ripple in the pass band and a faster roll-off in the stop band. The greater the eccentricity of the ellipse is, the larger the pass-band ripple and stop-band roll-off becomes. Generally, a ripple of about ± 0.5 dB or less is acceptable. The Chebyshev type II filter is derived from a type I by converting it from low pass to high pass and then inverting

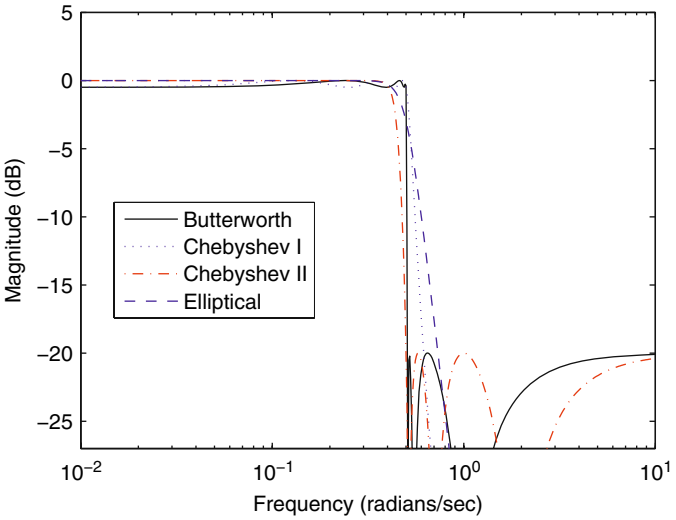


FIGURE 1 Several fifth-order (36 dB/octave) low-pass filters (cutt-off at 0.5 rad/s) showing sharper roll-off for Chebyshev II and Elliptical, but also greater high frequency leakage. Butterworth or Chebyshev type I (0.5 dB ripple specified here) is most desirable as anti-aliasing filters.

the polynomial. The poles of the type I polynomial that caused pass-band ripple are now zeros in the stop band for the type II filter. This leaves not only a flat response in the pass band, and a sharper roll-off, but also undesirable leakage bands in the stop band. The elliptical filter is similar to the Chebyshev type II except it is based on elliptical polynomials giving the sharpest roll-off, a flat pass band, but even greater leakage in the stop band. The sharp roll-off of the Chebyshev type II and elliptical filters also cause significant phase distortion near the cut-off frequency in the pass band. A fifth type of filter called a Bessel filter has a perfectly linear phase response (and the greatest group delay) but a very weak roll-off in the stop band [4]. While one should weigh all these factors in designing a data acquisition system, we recommend either Butterworth or Chebyshev type I anti-aliasing filters. Table 1 summarizes the filter characteristics as anti-aliasing filters showing why the Butterworth filter is desirable, especially if one samples over a wider bandwidth than needed.

Clearly, the octave from half the Nyquist frequency (one quarter of the sample rate) to the Nyquist frequency (half the sample rate) will have most of its response attenuated by the anti-aliasing filter roll-off. The sharper the roll-off, the more of this bandwidth can be used, which is why there is interest in sharp roll-off "brick

TABLE 1 Anti-aliasing filter characteristics

Filter type	Roll-off sharpness	Pass-band ripple	Stop-band leakage	Transient response
Butterworth	Weak	Excellent	Acceptable	Excellent
Chebyshev I	Acceptable	Poor	Good	Weak
Chebyshev II	Good	Weak	Poor	Acceptable
Elliptical	Excellent	Poor	Poor	Poor

wall" anti-aliasing filters. Using higher-order filters helps, but adds complexity, phase distortion, and time delay. So, in all cases, one has to deal with some level of leakage from the stop band into aliased frequencies in the pass band. It can be very difficult to recognize aliased interference, especially if it is broadband noise. Since it is always present to some degree, it is a good practice to measure the aliasing leakage by sweeping the frequency of a maximum amplitude sinusoid up past the Nyquist rate and measure the leakage directly using spectral analysis. Even delta–sigma converters with their sharp roll-off decimation filters can have leakage issues, especially if low sample rates are used.

4 • Basic Digital Signal Calibration

Calibration is an essential technique for validating the integrity of data. For example, in legal cases involving sound level measurements for noise mitigation, *one must calibrate both before and after a particular measurement* in order for the measurement to be admissible in the litigation. This is because there is no way to prove that the data acquisition system did not fail at some time during the measurement process if before and after calibrations are not available. One difficulty with electronics, battery, or connection failures when measuring noise is that the failure also produces noise. Even intermittent failures (say from a poor signal cable connection) can be identified using frequent calibration during a field test. However, to be practical, field calibration has to be fast and convenient. For microphones, products such as handheld speakerphone calibrators are an excellent choice for field work. They typically have selectable tones of 250 Hz, 1 kHz, and sound levels ranging from 94 dB (approximately 1 Pa RMS) to 124 dB. More expensive bench-type speakerphones cover many more frequencies and sound levels. The most conservative microphone calibrator is a mechanical pistonphone and barometer. The pistonphone has

adapters to compensate for the air volume displaced by different-sized microphones and a barometer to compensate for the air density effect on the generated sound level in the chamber. For accelerometers, handheld calibrators typically produce 1*g* RMS (9.8 m/s^2) at around 150 Hz as a reference and have a threaded hole for securely mounting the vibration sensor. Finally, a non-sensor calibration source is also quite convenient such as a 1 V RMS oscillator. These can be obtained with a convenient BNC-type connector for signal calibration of everything except the sensor.

4.1 Time Domain

Calibration of digital data is most often done using a simple "calibration constant" to convert the digital integers from the data acquisition system to floating point numbers in engineering units (Pascals, m/s^2, g's, etc.). One can estimate the "open-loop" calibration constant by collecting the advertised sensor voltage sensitivity (say S in mV/Pa for a microphone), the voltage gains of any amplifiers or attenuators in the signal path G_A, and the data acquisition "numerical gain" defined in Equation (1) as G_{ADC}. The calibration constant "C" is simply the inverse of the product of all the gains and the sensor sensitivity [5]:

$$C = \frac{1}{S \cdot G_A \cdot G_{ADC}}. \tag{2}$$

Multiplying the integers from the data acquisition system by C in Equation (2) should convert the data into the appropriate engineering units. However, as a check, one could put a known calibration signal through the system and back out the "field" calibration factor to the "open-loop" calibration factor as a reference. A good way to do this in practice is to have handy a little notepad with the estimated "open-loop" factor, the actual measured "field" factor, and a date and time. If one uses a field calibrator such as a speakerphone, a simple list with calibration data, time, and level will suffice. These entries are usually RMS signal averages. If a battery is near failure, a sensor is damaged, or a connection is intermittent, the RMS levels will show more than a few percent change, giving you an important clue to check the system before proceeding. Pick a quiet place for calibration. One should not try to calibrate a microphone outdoors on a windy day, near loud noise sources, or using a speakerphone sound level significantly louder or quieter than the signals one expects to record. Likewise, one should not try to calibrate an accelerometer in a moving vehicle, a boat in high sea state, or on a factory floor with machinery running. Isolation of the calibration signal from background noise is essential.

4.2 Frequency Domain Calibration

Using fast Fourier transforms (FFT) and spectral analysis offers an even more precise calibration technique than the basic time-domain calibration factor. This is because spectral analysis can separate sinusoids from broadband background noise. The larger the size N of the FFT, the greater the signal-to-noise ratio (SNR) gain ($\sqrt{N/2}$) will be in the frequency domain. The estimated spectrum is often averaged over a period of time which averages variations in the inherent background noise, although this does not enhance the SNR. Finally, one often uses a data window, such as a Hanning (cosine tapered) window, on the data prior to an FFT to reduce spectral leakage. All of these factors must be taken into account when preparing a frequency domain calibration [6].

We begin with a known sinusoidal signal of amplitude A in appropriate engineering units. Calculation of an N-point FFT of the real-valued signal (we assume the imaginary part of the FFT input is zero) will yield an amplitude of $AN/2$ if the frequency of the sinusoid lies *exactly* on one of the FFT frequency bins, that is, some multiple of Fs/N Hz. If the sinusoid is in between adjacent FFT bins, the amplitude will be less than $AN/2$ by as much as 3 dB, depending on which data window is used. Clearly, spectral leakage must be managed for any meaningful frequency domain calibration. Unfortunately, this important detail has rarely been addressed in the literature. To produce a calibrated spectrum, we first choose the appropriate frequency domain units, either RMS units or spectral density (units-squared per Hz). RMS units are convenient when analyzing sinusoidal signals or acoustic responses with a few narrowband resonances. For broadband acoustic signals, the units for spectral density are preferred.

Table 2 summarizes frequency domain calibration and the effects of the data window on narrowband sinusoids and broadband noise. We assume an N-point FFT on a sinusoidal signal of amplitude A in Gaussian noise with variance σ^2 and a sample rate of Fs Hz. We use two-sided frequency domain calibration meaning that a real time-domain signal is assumed and the calibration is applied to the positive frequencies only. The "RMS"-calibrated FFT scales the raw FFT output so that a bin-aligned sinusoid with a rectangular window produces the bin amplitude in the correct RMS units. The "narrowband" correction factor is just a scale factor for the type of data window used. If the sinusoid frequency is not bin aligned, or if a non-rectangular data window is used, one must sum the squared RMS bin values over a frequency range capturing the spectral leakage and take the square root to produce the correct RMS level of the sinusoid. The power spectrum, often called the periodogram, produces the correct noise variance level in

TABLE 2 Frequency domain calibration techniques

Data type	Formula	Signal level	Noise level	Units
Time domain	$[y_1 y_2 \cdots y_N]$	$\frac{A}{\sqrt{2}}$	σ	RMS units
Basic FFT	$\eta^{-1} \sum_{n=1}^{N} W_n y_n e^{-j\omega n}$	$\frac{NA}{2}$	$\sigma\sqrt{N}$	units/bin
"RMS" FFT (two sided)	$\frac{\sqrt{2}}{N}\eta^{-1}\left\lvert \sum_{n=1}^{N} W_n y_n e^{-j\omega n}\right\rvert$	$\frac{A}{\sqrt{2}}$	$\frac{\sigma}{\sqrt{N/2}}$	RMS units/bin
"Power spectrum" (two-sided)	$\beta^{-1}\frac{2}{N}\left\lvert \sum_{n=1}^{N} W_n y_n e^{-j\omega n}\right\rvert^2$	$\frac{NA^2}{2}$	σ^2	Units^2/bin
"Spectral density" (two sided)	$\beta^{-1}\frac{2}{Fs}\left\lvert \sum_{n=1}^{N} W_n y_n e^{-j\omega n}\right\rvert^2$	$\frac{N^2A^2}{2Fs}$	$\frac{\sigma^2}{Fs/N}$	Units^2/Hz
"Narrowband" scale factor η	$\eta = \frac{1}{N}\sum_{n=1}^{N} W_n$	Normalizes window amplitude to match rectangular window allowing narrowband signal calibration		
"Broadband" correction factor β	$\beta = \frac{1}{N}\sum_{n=1}^{N} \lvert W_n\rvert^2$	Normalizes window spectral leakage allowing broadband signal calibration		

every bin but only if the broadband correction factor is applied to account for spectral leakage. The spectral density, sometimes called the power spectral density (PSD), simply normalizes the power spectrum by the width of each bin in hertz. Many electronic component noise specifications are given as the square root of the PSD (i.e., units of $V/\sqrt{Hz}$). Calibration of broadband signals applies to the power spectrum and spectral density while narrowband signal calibration in the frequency domain should use the RMS FFT normalization and summing across the spectral leakage around the peak.

Averaging large numbers of non-overlapped spectral "snapshots" of sinusoids in Gaussian noise will yield a smoother looking spectrum but will not change the SNR. The noise statistics start out at the first average, with the mean-square and variance (over frequency bins) equal. After M spectral averages, the noise level mean stays the same while the standard deviation of the noise estimated over the frequency bins reduces by $\sqrt{M}$. Thus, the noise spectrum appears "smoother" over frequency, but its mean level is unchanged.

This implies that the relationship between the noise level and the noise standard deviation in the spectrum can tell you approximately how many averages are in the spectral estimate (if the underlying noise is a single Gaussian process). This is true only approximately but is a very useful relationship. It can be seen

that each frequency bin of the averaged spectrum has Gaussian probability density with mean equal to the spectral noise mean and standard deviation equal to $1/\sqrt{M}$ of the noise mean. For a bin with a sinusoid plus the noise, the noise mean is now shifted up by the RMS sinusoid amplitude, but with the same standard deviations as the surrounding FFT bins. This is true for M greater than around 10 or for very high SNR. For low SNR or few averages, the noise-only bins have chi-square distribution of order $2M$, while the signal plus noise bin has a Rician distribution, which are more skewed toward zero [7].

Nonetheless, the simplified model of a shifted Gaussian for each spectral bin distribution is quite reasonable and implies that one can precisely specify a confidence interval about the averaged spectral level in each bin based on the estimated mean and standard deviation of the noise-only bins around the sinusoid bin. For example, if the RMS spectrum peak is 1 V (RMS) and the standard deviation of the surrounding noise bins is 100 mV (RMS), then the 68 % confidence interval for the spectral peak is from 1.1 to 0.9 V (RMS). For precision calibration, statistical metrics such as these are important to report explicitly along with the estimated calibration data. For high SNR and reasonable numbers of averages, the estimated spectrum is quite accurate. However, this conversely implies a minimum SNR for which the calibrations meet or exceed the defined confidence interval.

5 • Multichannel System Calibration

Many data acquisition systems for acoustic and vibrations measurements use an array of sensors at known locations to spatially process the waveforms for propagation direction, mode measurement or isolation, intensity flux, or wavenumber interference measurements for holographic field reconstructions. These systems require very precise spatial phase calibration and are intolerant of uncorrected time delays between channels. However, most data acquisition systems share the actual analog-to-digital converter device between channels using a multiplexing scheme. Signals can leak or "cross-talk" between channels and there can be a common noise signal across all the channels. Therefore, it is very important to characterize these effects through direct measurement, remove any biases such as multiplex time delays, and provide estimates of the single and cross-channel SNR and coherence.

5.1 Simultaneous Sample and Hold

This is a feature of many data acquisition systems designed for multichannel use. The sample-and-hold circuit is like a snapshot frame of a movie. The moving signal voltage is captured by quickly charging a capacitor (or a more sophisticated sample-and-hold circuit) to meet the signal voltage and then isolating the capacitor to hold the voltage until the analog-to-digital conversion is completed. Simultaneous sample and hold implies an array of these switched capacitors to take a snapshot of a group of channels where the analog-to-digital converter converts each voltage to a number very quickly so that all channel conversions in the group are completed prior to the next sample cycle. There is no time delay for each channel simultaneously sampled in the group, which is obviously what is preferred. However, for the capacitors to charge quickly and the captured voltages to have minimal drift, the electrical impedance is quite high, making these circuits susceptible to electronic noise interference and cross-talk, sometimes caused by a weak design element in the analog converter switching channels. A good way to measure cross-talk and noise is to connect a circuit short to every other channel, with maximal level different frequency sinusoids in the non-shorted channels. In a good data acquisition system, the cross-talk and noise in the shorted channels is quite small, but is still measurable using large FFTs and spectral averaging.

The specifications for sophisticated data acquisition systems can be misleading. For example, a PCI-bus data acquisition board may have eight simultaneously sampled channels each with a multiplexer for up to eight channels and advertise up to 64 data channels with simultaneous sample and hold. However, acquiring all 64 channels, one only has groups of eight that are simultaneously sampled where each group is delayed by 1/8 the sample time interval. This practical limitation introduces a time delay bias that can easily be removed in the frequency domain and verified using broadband frequency domain calibration.

5.2 Multiplexed Channel Phase Correction

For systems where a fixed time delay exists between channels due to a multiplexed converter, one can connect a broadband signal to the two channels simultaneously and measure the transfer function and coherence between the two channels. The broadband input can be either a natural Gaussian noise from an analog electronic circuit, a pseudo-random maximal length noise sequence, or a frequency chirp. Chirp signals should cycle synchronized with the FFT data blocks so that the chirp in each block is identical. The signal need not be spectrally white, but it should cover the spectrum of interest with a reasonable SNR. A reasonable number of spectra are averaged to maintain

a small confidence interval for each spectral bin. With the same signal patched into both channels, an ideal intra-channel transfer function is unity amplitude with zero phase and unity coherence. The coherence $\gamma_{xy}^2(f)$ is a very useful quantity to measure because it provides a measure of the SNR and the transfer function errors [8]. The coherence is the magnitude-squared of the cross-spectrum average divided by the product of the input and output autospectra. One channel can be considered as the "input", while the other channel is considered the "output".

The SNR is calculated from coherence by

$$\mathrm{SNR}(f) = \sqrt{\frac{\gamma_{xy}^2(f)}{1-\gamma_{xy}^2(f)}}; \tag{3}$$

so, as the coherence approaches unity, the SNR is approaching infinity. The transfer function amplitude error variance is approximated using

$$\sigma_{xy}^2(f) = \frac{1-\gamma_{xy}^2(f)}{2M\gamma_{xy}^2(f)}\,|H(f)|^2, \tag{4}$$

where $|H(f)|^2$ is the magnitude-squared of the transfer function and M is the number of averages used. The phase error standard deviation is estimated using

$$\sigma_{\theta_{xy}}(f) = \tan^{-1}\left[\sqrt{\frac{1-\gamma_{xy}^2(f)}{2M\gamma_{xy}^2(f)}}\right]. \tag{5}$$

The statistical metrics for M spectral averages of the transfer function in Equations (4) and (5) provide a means to define a confidence interval around the measured cross-channel amplitude or phase bias.

This cross-channel bias is corrected by simple complex spectral normalization. If the measured transfer function H is based on channel 2 as output (y) and 1 as input (x), channel 2 can be made electrically identical to channel 1 by dividing it (divide the amplitude and subtract the phase) by this transfer function. As such, every channel in the array can be normalized to channel 1 in this manner, thus removing any amplitude or phase variations across the array in the frequency domain. If time-domain transfer function corrected signals are needed, one would invert the transfer functions and inverse Fourier transform to create a non-causal FIR filter for each channel for postprocessing, which would be very computationally intensive. Non-causality is needed because some channels may require a time advance relative to the reference channel.

6 • Data Validation and Diagnostic Techniques Summary

This section has briefly outlined the most important considerations for data acquisition and calibration and how to validate digital data. Below is a summary checklist outline.

1. Signal integrity

   (a) Keep line impedances low. Amplify at the sensor and drive a strong current across a low impedance resistor at the data acquisition end.
   (b) Use twisted shielded pair cables with the shield floated at the sensor end.

2. Closed-loop calibration

   (a) Adopt or develop a convenient field calibration technique using a NIST-traceable source and the sensor in the loop.
   (b) Repeat calibration recordings frequently and always before and after critical recordings.
   (c) Keep a calibration log with date, time, and calibration factors or levels (for a specific calibration source).

3. SNR assessment

   (a) Measure the system SNR using a maximal amplitude sinusoid where the background noise is estimated away from obvious spectral leakage.
   (b) Record the environmental background noise in separate file as needed to describe the background in the absence of any signal.
   (c) Short the data acquisition input and measure the residual electronic noise spectrum using a large number (40–100) of spectral averages.

4. Multichannel issues

   (a) Short every other input channel and input channel-specific frequencies at maximal amplitude to measure cross-talk. Repeat with the sinusoids and shorted channels switched.
   (b) Input a common broadband signal into all channels simultaneously and record the transfer functions and coherence from all channels (as outputs) relative to a reference channel input (say channel 1).
   (c) Repeat 4b with all channels shorted to measure the common noise levels and coherence.

References

1. P.M. Aziz et al., "An Overview of Sigma-Delta Converters", *IEEE Signal Processing Magazine* (Jan. 1996), 6183.
2. M.W. Hauser, "Principles of Oversampling A/D Conversion", *Journal of the Audio Engineering Society* **39**(**1/2**) (Jan. Feb. 1991).
3. A.V. Oppenheim, A.S. Willsky and Ian T. Young, *Signals and Systems*, Prentice – Hall International, London (1983).
4. T.W. Parks and C.S. Burrus, *Digital Filter Design*, Wiley, New York (1987).
5. D.C. Swanson, *Signal Processing for Intelligent Sensor Systems,* Chapter 1, Marcel Dekker, New York (2000).
6. D.C. Swanson, *Signal Processing for Intelligent Sensor Systems,* Chapter 6, Section 1, Marcel Dekker, New York (2000).
7. W.S. Burdic, *Underwater Acoustic System Analysis, 2nd ed.,* Chapters 9 and 13, Prentice-Hall, Englewood Cliffs (1984).
8. J.S. Bendat and A.G. Piersol, *Engineering Applications of Correlation and Spectral Analysis,* Wiley, New York (1993).

3

Spectral Analysis and Correlation

Robert B. Randall

University of New South Wales, Sydney, Australia

1 • Introduction

Spectral analysis is the process of breaking down a signal into its components at various frequencies, and in the context of acoustics there are two very different ways of doing this, depending on whether the result is desired on a linear frequency scale with constant resolution (in Hz) or on a logarithmic frequency scale with constant percentage resolution. The fundamental connection between the time domain and the frequency domain, the Fourier transform, is most easily interpreted in terms of linear time and frequency scales, at least in the practical version now used to calculate it, the FFT (fast Fourier transform). However, expressing a spectrum on a linear scale automatically restricts its frequency range, since the upper frequency decade occupies 90% of the scale and the upper two decades 99% of the scale. The human ear has a frequency range of three decades (20 Hz–20 kHz) which can only be expressed on a logarithmic frequency scale with constant percentage bandwidth, i.e., where the bandwidth of a filter at any frequency is a fixed percentage of its center frequency, as explained in Chapter 7. Moreover, the human ear tends to interpret equal intervals on a logarithmic frequency scale as equal steps (e.g., octaves with a frequency ratio of 2 between the highest and lowest frequencies covered by the band, or 1/3-octaves with a ratio of $2^{1/3}$). For estimating spectra

in terms of 1/n-octaves, it is most efficient to pass the signal through a series of filters with constant percentage bandwidth and measure the power transmitted by each filter. The filters can be either analogue or digital, but there is an increasing tendency to use the latter, since, as explained below, by efficient "timesharing" it is possible to use a very limited number of actual filter units to filter in parallel over any number of octaves, simply by processing data streams with different sampling frequencies varying in octave (2:1) steps.

The "power" of a signal is directly related to the square of its instantaneous value, but is normally interpreted as its "mean square" value, obtained by averaging the instantaneous value over a defined averaging time (or over an ensemble of realizations for a random process). For continuous deterministic signals made up of sums of sinusoids, the mean square value of each sinusoid is half the square of its amplitude or peak value, and the total power or mean square value is simply the sum of the powers of the components. This is one version of Parseval's theorem. For stationary random signals, the power is continuously distributed with frequency (rather than being concentrated at a single frequency as for a sinusoid), and spectra are expressed in terms of "power spectral density" or PSD, which has to be integrated over a finite frequency band to give a finite power. The version of Parseval's theorem applicable to them is that the total mean square value as calculated in the time domain is equal to the integral of the PSD over all frequency. Power is meaningful for stationary or slowly varying signals, but for transient events, such as vibratory shocks or acoustic pulses, the fundamental quantity is their "energy" obtained by integrating their instantaneous "power" over their total length. The square of the magnitude of their Fourier transform gives the distribution of their energy with frequency, the so-called energy spectral density or ESD, and the version of Parseval's theorem applicable to them is that the integral of ESD over all frequency equals the total energy obtained by integrating the instantaneous power over all time. Attention is paid to these differences in dimensions and units in the following sections where the practical spectral analysis of different types of signals is discussed in detail.

As will be shown later, correlation functions are the equivalent of spectral functions in the time domain, being related to them by a Fourier transform. They express the relationship between signals (possibly the same signal) with various amounts of delay, so that for example periodic functions continue to correlate perfectly every period, whereas white noise does not correlate

with itself for even the smallest delay, while narrow band noise does correlate over a number of periods of the basic oscillation, but the correlation deteriorates with increasing delay. The cross-spectrum (Fourier transform of the cross-correlation) between two signals and their two autospectra can be used to calculate their coherence, a measure of the degree of linear relationship between the two signals.

Theoretical frequency spectra represent averages over all time, but the ear is able to discern the way in which frequency patterns change with time, and so various time/frequency analyses have been developed to represent such changes in frequency patterns with time. A basic limitation here is the "Heisenberg uncertainty principle" which states that the product of the resolution in the time and frequency domains is a constant so that improvement in one comes at the cost of deterioration in the other. The various forms of time/frequency analyses discussed below represent different attempts to overcome this problem, for example to better represent what the human ear can achieve.

2 • The Fourier Transform and Variants

The basic connection between the time domain and the frequency domain is the Fourier transform, with the following equation [1]

$$G(f) = \int_{-\infty}^{\infty} g(t) \mathrm{e}^{-\mathrm{j}2\pi ft} \mathrm{d}t, \tag{1}$$

which transforms the time signal $g(t)$ into the frequency spectrum $G(f)$. For causal time signals (e.g., the impulse response function of all physical systems), the Fourier transform is equal to the Laplace transform evaluated along the imaginary axis of the Laplace plane (where the Laplace variable $s = \mathrm{j}\omega = \mathrm{j}2\pi f$) [2].

The time signal can be retrieved from its Fourier transform by the inverse transform

$$g(t) = \int_{-\infty}^{\infty} G(f) \mathrm{e}^{\mathrm{j}2\pi ft} \mathrm{d}f, \tag{2}$$

which can be seen to be almost identical to (1), so that many properties of the forward transform apply equally to the inverse transform, for example the convolution theorem [2]. The latter

states that a convolution in one domain corresponds to a multiplication in the other domain. For a convolution in the time domain represented by

$$g(t) = f(t) * h(t) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)\mathrm{d}\tau \tag{3}$$

where the asterisk is a shorthand way of representing the convolution operation, this corresponds to the following formula:

$$G(f) = \Im\{f(t) * h(t)\} = F(f) \cdot H(f), \tag{4}$$

where $\Im\{.\}$ represents the Fourier transform of the argument, and the uppercase frequency variables represent the Fourier transforms of the corresponding lowercase time variables. Since the convolution operation of (3) represents the output $g(t)$ of any physical system with impulse response function $h(t)$ subject to the forcing function $f(t)$, it is evident that the corresponding multiplicative relationship of (4) is much simpler.

2.1 The Fourier Series

Strictly speaking, the Fourier transform of (1) only applies to transient functions (whose integral over all time is finite) [1], but using the convolution theorem it can be extended to the case of periodic functions as illustrated in Figure 1 [2]. Any periodic function can be generated by convolving one period (of length T) with an infinite train of unit delta functions with spacing T. Since the Fourier transform (actually Fourier series) of such a train of delta functions is another train of delta functions with scaling $1/T$ and spacing $1/T$, the spectrum of the periodically repeated transients is the product of the spectrum of the single transient with the scaled train of delta functions, thus sampling it at intervals of $1/T$, and multiplying these values by $1/T$. Thus, the Fourier series spectrum for a transient $g(t)$ with Fourier transform $G(f)$, repeated with a period of T, can be calculated as

$$G(f_k) = \frac{1}{T}G\left(\frac{k}{T}\right) = \frac{1}{T}\int_{-T/2}^{T/2} g(t)\mathrm{e}^{-\mathrm{j}2\pi kt/T}\mathrm{d}t. \tag{5}$$

This is the complex version of the better known expression where a periodic signal $g(t)$ is expressed in terms of sine and cosine functions:

$$g(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(k\omega_0 t) + \sum_{k=1}^{\infty} b_k \sin(k\omega_0 t), \tag{6}$$

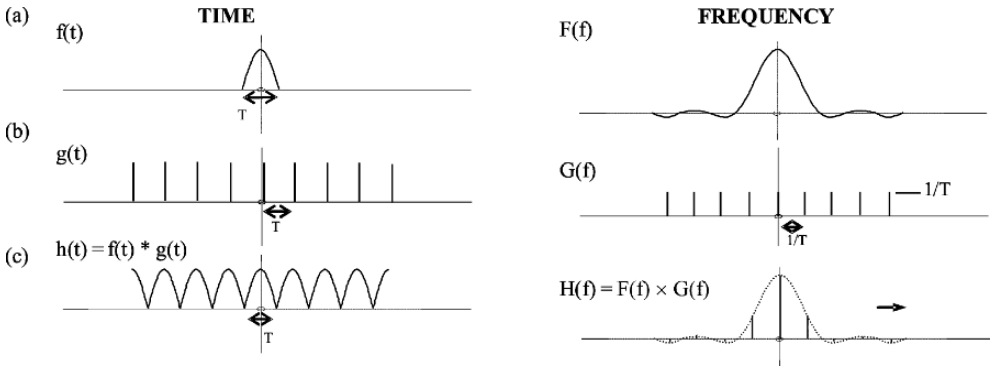


FIGURE 1 Illustration of how the Fourier series of a periodically repeated transient (a) can be obtained by sampling the Fourier transform of the transient.

where

$$a_k = \frac{2}{T}\int_{-T/2}^{T/2} g(t)\cos(k\omega_0 t)\mathrm{d}t \tag{7}$$

and

$$b_k = \frac{2}{T}\int_{-T/2}^{T/2} g(t)\sin(k\omega_0 t)\mathrm{d}t, \tag{8}$$

meaning that

$$G(f_k) = \frac{a_k}{2} - \mathrm{j}\frac{b_k}{2} \tag{9}$$

Note that (6) is for zero and positive frequencies only, whereas the frequency components of (5) extend from minus to plus infinity. Each sinusoidal component is made up of a sum of a positive and a negative frequency component, each with half the amplitude of the corresponding sinusoid (as indicated by (9)) since when they align, the total amplitude is double that of the individual components. This fact is important in scaling frequency spectra.

2.2 The Discrete Fourier Transform (DFT)

Equations (1) and (2) can be replaced by the following versions, when the signals are discretely sampled (in both time and frequency domains) and of finite length (implicitly one period of a periodic signal) [3,4]:

$$G(k) = \left(\frac{1}{N}\right)\sum_{n=0}^{N-1} g(n)\mathrm{e}^{-\mathrm{j}2\pi kn/N}, \tag{10}$$

$$g(n) = \sum_{k=0}^{N-1} G(k) e^{j2\pi kn/N}, \tag{11}$$

where (10) is the forward transform and (11) is the inverse transform. This version gives correctly scaled Fourier series spectra, and the scaling must be modified if other types of signals are processed, as described below in Section 3, "Practical FFT Analysis". The widely used signal processing package Matlab® makes the division by N in the inverse transform, giving a version more closely related to the Fourier transform pair of (1) and (2). However, the sum will then only represent the integral (over a finite length time signal) when the forward transform is multiplied by Δt, the time interval between samples.

Figure 2 shows the relationship between the Fourier transforms of (1) and (2) and the DFT of (10) and (11). It will be seen that the discrete sampling in each domain implies periodicity in the other domain so that the second half of the spectrum (from $f_s/2$ to f_s, where f_s is the sampling frequency) actually represents the negative frequency components from $-f_s/2$ to zero. This highlights the fact that a time signal sampled at f_s must not contain frequencies higher than half the sampling frequency before digitization, as discussed below.

2.3 The Fast Fourier Transform (FFT)

The FFT is simply a very efficient way of calculating the DFT using a number of operations of the order of $N \log_2 N$ instead of the N^2 required for a direct calculation [3–5]. For a typical transform size of 1,024 (2^{10}), this is more than 100 times faster. However, the FFT has all the properties of the DFT, and in particular the three "pitfalls" as described below in Section 3, "Practical FFT Analysis".

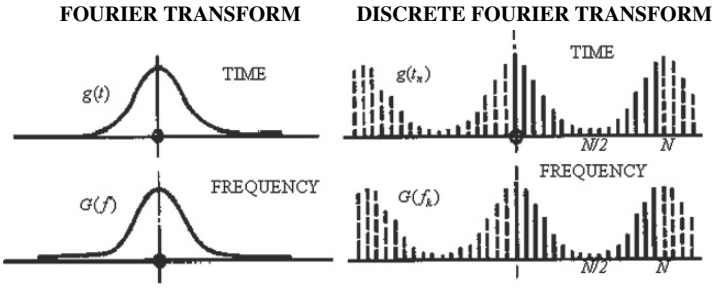


FIGURE 2 Relationship between the Fourier transform and the DFT.

2.4 Zoom FFT Analysis

As can be seen in Figure 2, the DFT always covers a frequency span from zero to half the sampling frequency. Sometimes it is desired to analyze a smaller range of frequency with high resolution, and this can be achieved by "zoom FFT". It is of course possible to achieve a high resolution by simply increasing transform size N, but a large transform size can be avoided if the signal can be preprocessed in real time by a zoom processor [4]. As illustrated in Figure 3, the real-time signal is first displaced in frequency by multiplying it by the (complex) unit vector $\exp(-j2\pi f_c t)$, which subtracts f_c from every frequency in the signal and shifts the frequency origin to f_c. The resulting signal is complex, but can be lowpass filtered by a digital filter to remove all frequency components outside the desired zoom band. The lowpass filtered signal can be resampled at a correspondingly lower sample rate without aliasing, meaning that a smaller number of samples N will correspond to a given record length T in seconds. Independent of the frequency shift, the frequency resolution is always given by $1/T$ so that a finer resolution is achieved in the zoom band. The FFT obtains the Fourier spectrum of the complex frequency-shifted time signal, which can simply be displayed as representing the original zoom band, rather than plus/minus half the new sampling frequency. The lowpass filtering and resampling can conveniently be done in octave steps, as explained below for digital filters, and so zoom is most commonly available in factors corresponding to a power of 2.

It is perhaps worth mentioning that the complex output of a zoom processor can also be used for demodulation, since its amplitude (modulus) represents the amplitude modulation of the original signal, and its phase the phase modulation signal,

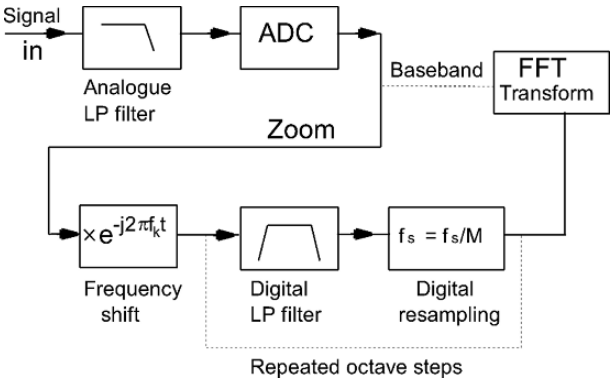


FIGURE 3 Schematic diagram of FFT zoom process.

for demodulation at the zoom center frequency [2,4]. The phase will have to be "unwrapped" to a continuous function of time (for example, using the Matlab® function UNWRAP) and can be differentiated to give the frequency modulation signal.

3 • Practical FFT Analysis

3.1 Aliasing, Windows, Leakage, and Picket Fence Effect

Aliasing, leakage, and picket fence effect are the three pitfalls of the FFT referred to above [4]. Figure 2 shows that before analogue to digital conversion (ADC) the signal must be lowpass filtered to exclude frequencies above half the sampling frequency because the periodic nature of the spectrum will otherwise cause "aliasing". To be able to use 80 % of the calculated spectrum, the antialiasing filter must be very steep (120 dB/octave).

Because the signal record transformed by the FFT is effectively joined into a loop and repeated periodically, for general signals this can give a discontinuity at the join which is not part of the actual signal, and which modifies the results. The effects can be reduced by multiplying the signal record by a "data window", since otherwise the signal is effectively multiplied by a rectangular window of length equal to the record length. By the convolution theorem, this corresponds in the frequency domain to a convolution with the Fourier transform of the window function (for a rectangular window, a sin x/x function). The Fourier transform of the window function thus acts as a filter characteristic, with which each frequency component is replaced. This phenomenon is known as "leakage", since power originally at one frequency is leaked into adjacent frequencies in proportion to the square of this filter characteristic. The actual leakage depends not only on this filter characteristic but also on how it is sampled in the frequency domain. The sampling in the frequency domain results in another window property known as the "picket fence effect". If a sinusoidal frequency corresponds to an integer number of periods along the record length, it will coincide with a frequency line, and the sin x/x window characteristic will be centered on this line, and thus sampled at the zero crossings for all other frequency lines. In this special case, neither the leakage nor the picket fence effects will be visible, but in practice this can rarely be achieved for more than one periodicity at a time, and the frequency characteristic will be sampled at other than the zero crossings. The various window functions are usually compared for the worst case situation where the frequency is halfway between two lines (and, for example, the sidelobes of the sin x/x function sampled at their peaks).

Leakage is primarily indicated by three parameters, the highest sidelobe (−13.4 dB for the rectangular window), the rate of rolloff of the remaining sidelobes (−20 dB/decade for rectangular), and the noise bandwidth compared with the line spacing (1.0 for rectangular). The picket fence effect gives the maximum error of observed peak values compared with the actual peak (−3.9 dB for rectangular). The noise bandwidth is the width of ideal filter (i.e., rectangular in the frequency domain) that would transmit the same amount of power from a white noise source, i.e., with the same integral under the squared amplitude characteristic. Table 1 gives the properties of a number of window functions that are applied in practice to stationary signals.

The Hanning window is found in most FFT analyzers and is a useful general-purpose window with steep filter characteristic, small noise bandwidth, and reasonable picket fence effect. The flat top window is designed to eliminate the picket fence effect and is most useful when the spectrum is dominated by discrete frequency components, since their RMS values can be read directly without correction. This is particularly valuable when using a calibration signal to calibrate a frequency spectrum. However, it has a very large noise bandwidth, and so the discrete frequencies do not protrude so much from background noise in the spectrum (4 dB less than with Hanning). The Kaiser-Bessel window has the steepest filter characteristic in the sense of giving the best separation of closely spaced components of very different level, but such separation can also be achieved by using zoom.

Other windows are used for analyzing transient signals. For example, a decaying exponential window is used to force lightly damped responses to near zero at the end of the data record to avoid the discontinuity that would otherwise result from joining the signal into a loop. This corresponds to an added extra damping that is known exactly and can thus be subtracted from measured results.

TABLE 1 Properties of various windows

Window	Noise bandwidth	Highest sidelobe (dB)	Sidelobe rolloff (dB/decade)	Picket fence effect (dB)
Rectangular	1.0	−13	20	3.9
Hanning	1.5	−33	60	1.4
Kaiser-Bessel	1.8	−60	20	0.8
Flat top	3.8	−70	20	<0.1

3.2 Spectrum Averaging

After correction for picket fence effect, each estimate of a discrete frequency component will be the same, so little is achieved by averaging, except to confirm that a particular frequency component is deterministic.

For random signals, each FFT transform gives an estimate of the power or autospectrum, but a number of them must be averaged to obtain a valid estimate. The general formula for the standard deviation of the error in the RMS amplitude of an autospectrum estimate is [4,6]

$$\varepsilon = \frac{1}{2\sqrt{BT}}, \tag{12}$$

where B is the bandwidth of each narrow band filter and T is the effective averaging time. In FFT analysis, the BT product given by each spectrum is 1, and thus for averaged FFT spectra, (12) can be used with BT replaced by n, the number of independent spectra averaged. For rectangular weighting, "independent" means from non-overlapping time records, but for the Hanning window most commonly used with random signals, up to 50% overlap of successive records can be used without significant loss of statistical reliability. Two-thirds and more overlap gives a temporally uniform time weighting, but n should still be estimated as for 50% overlap.

3.3 Spectrum Scaling

When the FFT according to (10) is used, and correction is made for picket fence effects, discrete frequency components will be scaled as half the amplitude of the equivalent sinusoidal components, as indicated in (9). Thus, the value must be multiplied by $\sqrt{2}$ to obtain the equivalent RMS value or by 2 to obtain the sinusoidal amplitude [2,4].

For random signals, the required averaging does not change the scaling, and thus the measured power obtained by squaring the value in each line must be divided by the equivalent filter bandwidth to give power spectral density (PSD). To include the power of the negative frequency components, the value obtained by squaring each positive frequency component should be multiplied by 2. The filter bandwidth by which this is to be divided is the line spacing $1/T$ multiplied by the relative noise bandwidth of the filter (as given in Table 1 for the most common windows) [2,4].

For transient signals which fit into the record length, no window need be applied, since there will be no discontinuity when the ends are joined, and so the noise bandwidth is 1. To convert to energy spectral density (ESD), the measured power in each line should be multiplied by T^2 [2,4].

4 • Spectral Analysis Using Filters

This will only be considered for stationary signals, since even though it is possible to analyze transients using filters it is quite complicated and beyond the scope of this section. An exception is the analysis of acoustic transients as expressed in terms of SEL (sound exposure level), since the result is still expressed in terms of power, as though the transient had constant power over 1 s. Thus the measurement is done in terms of power (e.g., L_{eq}) over whatever (linear) averaging time is necessary to capture the whole event, and then the L_{eq} can be converted to SEL by adding $10 \log_{10}(T_A)$ where T_A is the averaging time in seconds used for the L_{eq}. Note that the length of the event is not determined by the original signal alone. It may be dominated by the length of filter response (at low frequencies) which can be taken as $3/B$, where B is the filter bandwidth in hertz.

With $1/n$th-octave (constant percentage bandwidth) filters, the spectrum is usually expressed as power or RMS value in each band, though often expressed in dB re the appropriate reference level. The power is simply measured as the mean square value of the signal transmitted by each filter over the specified averaging time, or for stationary or slowly varying signals the equivalent obtained by exponential averaging with the corresponding time constant [4]. For exponential averaging, the equivalent linear averaging time T_A is given by 2RC, where RC is the time constant of the electronic smoothing circuit used for the averaging, or the equivalent time constant for digital exponential averaging [4]. Thus for sound level meters set to "slow" (time constant 1 s), the equivalent averaging time is 2s. For stationary random signals, the averaging time must be chosen in accordance with (12), keeping in mind that it will always be the lowest frequency filter that has the minimum bandwidth B, and thus the lowest BT product for a given averaging time [4].

$1/n$th-octave filters are efficiently generated digitally, since the characteristics of a digital filter are directly proportional to the sampling frequency, and so for 1/3-octave analysis, for example, only three sets of filter coefficients are required for an octave. If the sampling frequency is halved (by discarding every second sample), the same coefficients generate the 1/3-octave filters one octave below, and so on. Before reducing the sampling rate, the signals must be lowpass filtered to remove the information in the highest octave, and thus avoid aliasing, but this can also be done with just one set of filter coefficients. If M is the number of samples per second in the highest octave, because the sample rate is halved for each lower octave, the total number of samples to be processed per second for any number of octaves equals

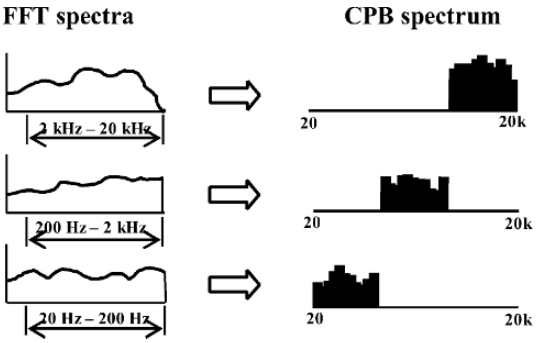


FIGURE 4 Conversion from FFT spectra to a 1/nth-octave spectrum.

$M(1+1/2+1/4+1/8+\cdots)=2M$. Thus if the processor is able to calculate twice as fast as the input sample rate, any number of octaves can be calculated in real time. The same procedure can be used for exponential averaging, as this is nothing other than a lowpass filtration by a first-order filter [4].

1/nth-octave spectra are sometimes calculated from FFT spectra by the process illustrated in Figure 4, where the upper decade of three separate FFT spectra is converted into constant percentage bandwidth and assembled into a single spectrum over three decades. The original FFT spectra are assumed to have constant PSD in each line, so the spectrum can be considered as a bar graph. The lower and upper cutoff frequencies of each 1/nth-octave filter are calculated, and the PSD in the corresponding lines (and parts of lines) of the FFT spectrum is integrated to give the total power in the 1/nth-octave filter. The problem with this approach is that the filter characteristic of the synthesized filter changes abruptly at the junctions between decades, since the slope changes by a factor of 10. The resulting filter characteristics may not meet the requirements of international and national standards [2].

To reduce this problem, some analyzers do the conversions over frequency bands of one octave rather than a decade.

5 • Correlation Functions

Correlation functions give a measure of how well one signal correlates with another as a function of the time displacement between them. The cross-correlation function between a signal $x(t)$ and a signal $y(t)$ is given in general by the expression [2]

$$R_{xy}(t,\tau)=E\left[x\left(t-\tau/2\right)y\left(t+\tau/2\right)\right], \tag{13}$$

which gives the value centered on time t for a total displacement τ. The symbol $E[.]$ represents the expected value or statistical average over an ensemble of different realizations of the statistical process that produced the signals x and y. For stationary processes, there is no statistical variation with time, and the average can be taken over time, as represented by

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x\,(t-\tau/2)\, y\,(t+\tau/2)\,\mathrm{d}t. \tag{14}$$

which is independent of time t, and a function of displacement τ only. Likewise because of the independence with respect to time t, the displacement does not have to be symmetric, and so the cross-correlation function is often represented as [7]

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x\,(t)\, y\,(t+\tau)\,\mathrm{d}t. \tag{15}$$

When $x(t) = y(t)$, the so-called autocorrelation function results, giving the equivalent of (13), for non-stationary functions, as

$$R_{xx}(t,\tau) = E\left[x\,(t-\tau/2)\, x\,(t+\tau/2)\right] \tag{16}$$

and of (15), for stationary functions, as

$$R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x\,(t)\, x\,(t+\tau)\,\mathrm{d}t. \tag{17}$$

The autocorrelation function gives a measure of how well a signal correlates with itself as a function of displacement. Any sinusoid correlates perfectly at zero displacement and for displacements corresponding to any exact number of periods. Thus, the autocorrelation function for any sinusoid (independent of initial phase) is a cosine function. Similarly, any periodic function has a periodic autocorrelation function consisting entirely of cosines so that it tends to be more impulsive than the original function (all harmonics are aligned at time zero).

It can be shown that the Fourier transform of the autocorrelation function is equal to the power spectrum, and before the advent of the FFT algorithm this was the normal means of calculating the power spectrum [6]. This gives a means of determining the autocorrelation function of noise signals as illustrated in Figure 5.

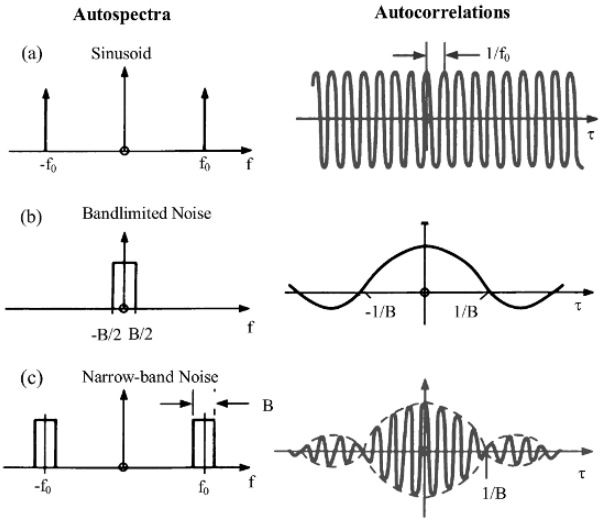


FIGURE 5 Autocorrelation vs autospectrum for three signals. Note that the spectrum of (**c**) is the convolution of (**a**) and (**b**).

Because the narrow band spectrum of Figure 5c is the convolution of a and b, the autocorrelation function of c is the product of the other two, and thus it can be seen that the "correlation length" of noise of bandwidth B is of the order of $1/B$ regardless of whether the noise is bandlimited or bandpass. This gives a limitation in the use of the autocorrelation function to detect echoes, since even if the echo is a perfect scaled copy of the original signal, there will only be a well-defined peak at the echo delay time if the bandwidth of the original signal is fairly large. The same applies when using the cross-correlation function to determine whether one signal is simply a scaled delayed version of another.

5.1 Cross-Spectrum

In the same way that the Fourier transform of the autocorrelation function is the autospectrum, or power spectrum, the Fourier transform of the cross-correlation function is the "cross-spectrum", which can also be calculated directly as [2,7]:

$$G_{xy}(f) = E\left[G_x^*(f) \cdot G_y(f)\right]. \tag{18}$$

This has an amplitude given by the product of the amplitudes of the two spectra (at each frequency) and a phase given by the difference of the two phases (i.e., the phase change from x to y). The effect of additive random noise tends to zero as a result of the averaging operation over a number of realizations.

5.2 Coherence

The coherence between two signals $x(t)$ and $x(t)$ is given by the formula

$$\gamma^2(f) = \frac{|G_{xy}(f)|^2}{G_{xx}G_{yy}} \tag{19}$$

which is effectively the (square of the) correlation coefficient between the frequency components as a function of frequency [4]. It has a value between zero and one and gives a measure of the degree of linear relationship between the two signals as a function of frequency. It is reduced by the addition of noise in either of the two signals, or if the relationship is non-linear. In practical terms, it will also be reduced by any delay between the equivalent parts of the two signals, if this is of the same order as the lengths of record used to generate the frequency spectra of the realizations. Note that the first estimate of coherence (for one realization only) will always be one, and it is only after averaging a number of realizations that the vector sum of the complex individual estimates of the cross-spectrum will be less than if they were aligned (as is the case when the relationship is fully linear).

As an example of the application of the coherence, the product of it with one of the autospectra used to calculate it gives the "coherent power", the amount of that autospectrum resulting from the linear part of the relationship between $x(t)$ and $y(t)$. This can be useful for identifying the sources of a measured spectrum [8].

6 • Time–Frequency Analysis

In principle, frequency analysis represents an average over all time, but it can be advantageous to investigate how the frequency content of a signal varies over time, as the human ear is certainly able to discern this.

6.1 Uniform Frequency Resolution

The most direct way of producing a time–frequency diagram is via the short time Fourier transform (STFT), which is an array of spectra obtained by moving a window (such as Hanning) along the time record and generating a spectrum for each position. This is illustrated in Figure 6, for a series of transients representing impulse responses. For some applications the individual spectra can be complex, but when they are autospectra, and particularly for sound signals, e.g., speech, the diagram is sometimes known as a sonogram.

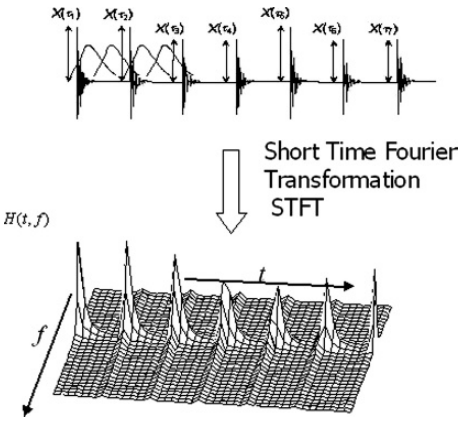


FIGURE 6 Generation of STFT.

The limitation with the STFT is that the product $\Delta t \times \Delta f =$ constant, where Δt is the resolution in time and Δf is the resolution in frequency (by the Heisenberg uncertainty principle). Thus, better resolution in one can only be obtained at the expense of poorer resolution in the other.

The Wigner–Ville distribution (WVD) was proposed as a means of obtaining better simultaneous resolution. It is defined as [9]

$$\mathrm{WVD}(t,f) = \int_{-\infty}^{\infty} x\left(t-\frac{\tau}{2}\right)x^*\left(t+\frac{\tau}{2}\right)\mathrm{e}^{-\mathrm{j}2\pi f\tau}\mathrm{d}\tau. \tag{20}$$

The WVD appears to give better resolution but suffers from interference (cross) components that are difficult to distinguish from the real physical components in the signal. To overcome this problem, various smoothed versions have been proposed [3] that largely eliminate the interference at the expense of somewhat reduced resolution (but still better than the STFT).

Figure 7 gives an example comparing the STFT with the WVD, and a smoothed version, the so-called smoothed pseudo-Wigner–Ville distribution. It represents the time–frequency variations within a firing cycle of the vibration signal measured on the head of a diesel engine.

A distribution similar to the WVD is obtained by performing a Fourier transform with respect to time lag τ of the autocorrelation of (16). It is known as the Wigner–Ville spectrum (WVS). The difference is the averaging given by taking the expected value, instead of performing the Fourier transform on a single realization. Antoni [10] has shown that for second-order cyclostationary signals, the interference components average to zero,

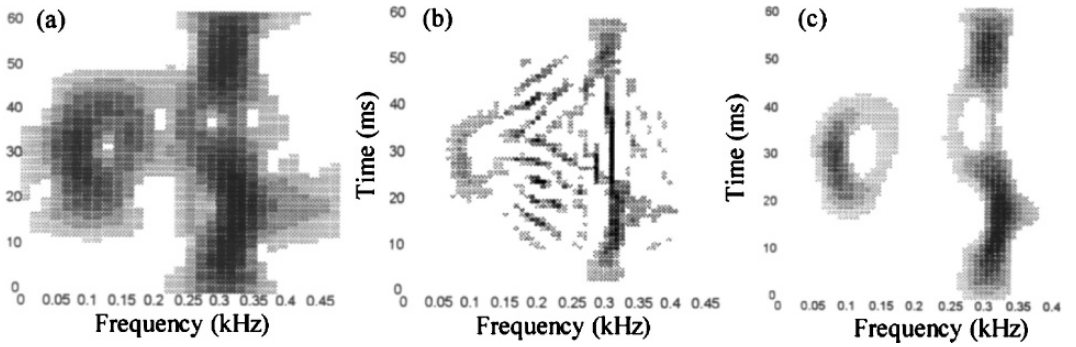


FIGURE 7 Comparison of time–frequency distributions for a diesel engine vibration signal: (**a**) STFT, (**b**) Wigner–Ville distribution, (**c**) smoothed pseudo Wigner–Ville distribution.

while the resolution remains unchanged. Figure 8 compares the WVS with the WVD for the impulse response of a nonlinear single degree-of-freedom (SDOF) system. In the WVS, the decreasing natural frequency with amplitude is shown very clearly (with the same resolution as the WVD), whereas the interference terms in the latter cloud the issue. The system was excited by a burst random signal which is typically second-order cyclostationary because the mean value (first-order statistic) is zero, while the variance (second-order statistic) is a square wave (periodic).

6.2 Logarithmic Frequency Resolution

The STFT and other so-called Cohen class distributions have uniform resolution in both time and frequency, but another time–frequency representation with inherent logarithmic frequency resolution is given by wavelet analysis [11]. This is far too big a topic to be treated in detail here, but wavelet analysis can be said to be the result of filtering a signal with a number of matched filters obtained from a basic "mother wavelet" by dilation in geometric steps (thus giving a logarithmic frequency resolution), each dilation corresponding to a particular "scale", and displaced in time by amounts proportional to the wavelet length at each scale. Wavelets can be either continuous or discrete and either orthogonal or non-orthogonal.

Discrete orthogonal wavelets give the most efficient representation of a signal, with the possibility of analysis/synthesis and possible filtering in the wavelet domain. If the wavelets are chosen to be similar to features in the original signal which it is desired to enhance, then considerable denoising can be achieved by eliminating components below a defined value.

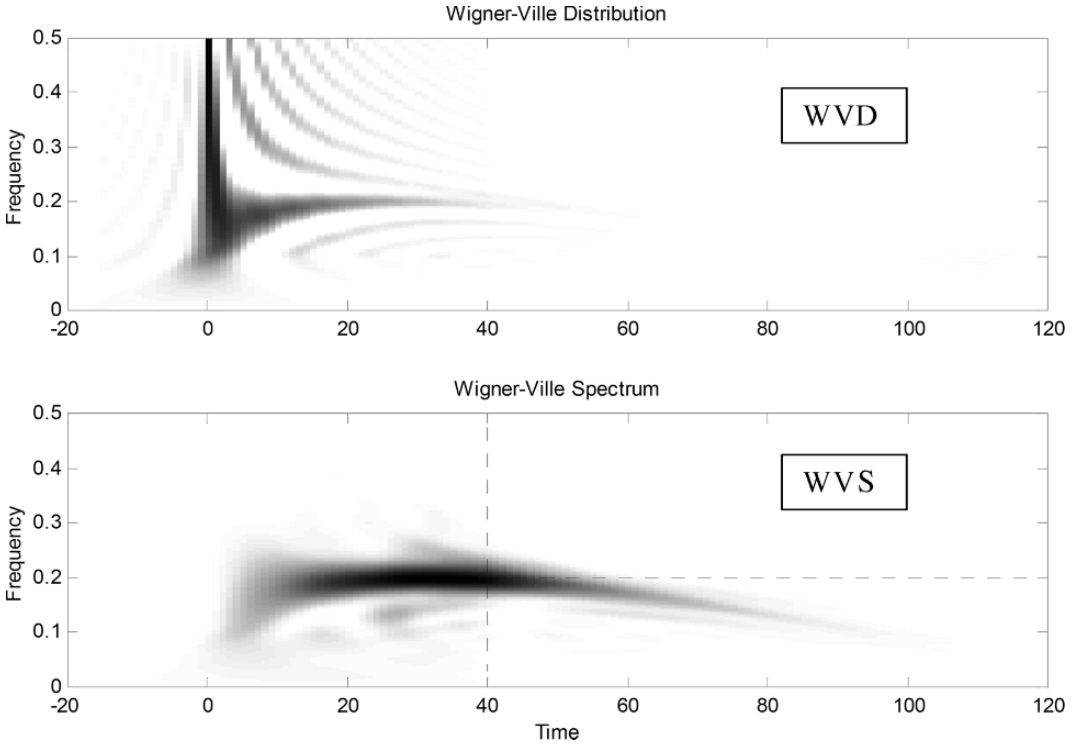


FIGURE 8 Impulse response of a non-linear SDOF system expressed as a Wigner–Ville spectrum (excitation by a burst random signal) and compared with the Wigner–Ville distribution with interference components [10].

The formula for performing a wavelet analysis is

$$W(a;b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{(t-b)}{a} \right) \mathrm{d}t, \tag{21}$$

where $\psi(t)$ is the mother wavelet, translated by b and dilated by factor a, where a is the scale.

Continuous non-orthogonal wavelets, such as Morlet wavelets, which are Gaussian-windowed sinusoids, have properties very similar to 1/nth-octave filters, except for being symmetrical and non-causal and thus with zero phase shift. The same can be achieved with digital 1/nth-octave filtering. They have the property, which is advantageous in some cases, that even though they are subject to the uncertainty principle, the time resolution is better at high frequencies where the bandwidth is greater, and thus impulsive events can be well localized in time. Figure 9 (from [12]) illustrates this for a gear vibration signal where a local fault is better resolved at higher frequencies.

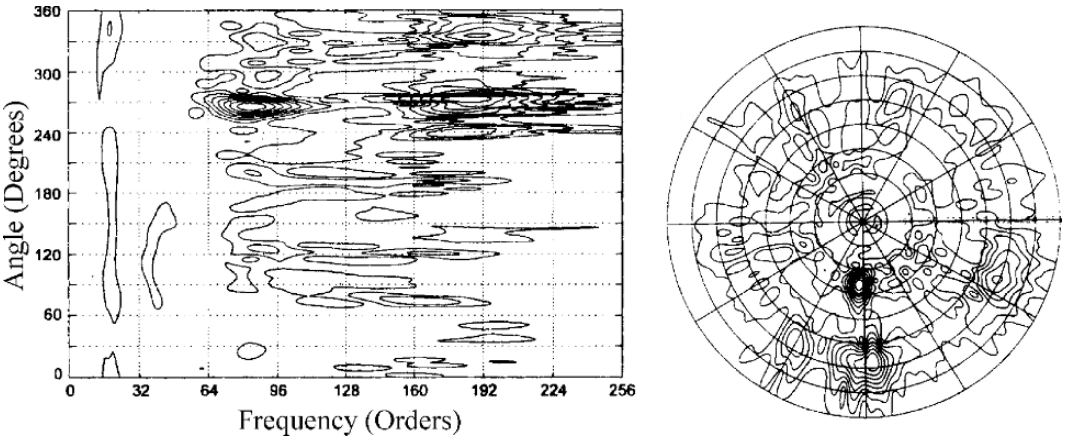


FIGURE 9 Use of wavelet transform to detect a local gear fault: (**a**) rectangular scales; (**b**) polar plot (frequency radial). From [6].

Lin and McFadden [12] proposed a radial display for this type of application (where the time axis can be related to rotation angle of a shaft) since it automatically gives better time resolution at higher frequencies when the frequency scale is radial.

References

1. Papoulis, P. (1962) *The Fourier Integral and Its Applications*. McGraw-Hill, New York.
2. Randall, R.B. (2001) "Signal Processing in Vibration Analysis". In *Structural Dynamics @ 2000: Current Status and Future Directions* (Eds. D.J. Ewins and D.J. Inman). Research Studies Press, Baldock, Hertfordshire, England.
3. Brigham, E.O. (1974) *The Fast Fourier Transform*. Prentice-Hall, New Jersey.
4. Randall, R.B. (revised ed. 1987) *Frequency Analysis*. Bruel & Kjaer, Copenhagen.
5. Cooley, J.W. and Tukey, J.W. (1965) "An Algorithm for the Machine Calculation of Complex Fourier Series". *Math. Comp.*, Vol. 19 No. 90, 297–301.
6. Bendat, J.S. and Piersol, A.G. (1971) *Random Data: Analysis and Measurement Procedures*. Wiley–Interscience, New York.
7. Bendat, J.S. and Piersol, A.G. (1980) *Engineering Applications of Correlation and Spectral Analysis*. Wiley, New York.
8. Halvorsen, W.G. and Bendat, J.S. (1975) "Noise Source Identification Using Coherent Output Power Spectra". *Sound Vibration*, Vol. 9, 15.
9. Cohen, L. (1995) *Time–Frequency Analysis*. Prentice-Hall, New Jersey.

10. Antoni, J. (2003) "On the Benefits of the Wigner–Ville Spectrum for Analysing Certain Types of Vibration Signals". Wespac8 Conference, Melbourne.
11. Newland, D.E. (1993) *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*, 3rd ed. Pearson Education (formerly Longmans). Reprinted Dover, 2005.
12. Lin, S.T. and McFadden, P.D. (1995) "Vibration Analysis of Gearboxes by the Linear Wavelet Transform". *Second International Conference on Gearbox Noise, Vibration and Diagnostics*, IMechE, London.

4

The FFT and Tone Identification

John C. Burgess[1]

Department of Mechanical Engineering, University of Hawaii, 2540 Dole Street, Honolulu, HI 96822, USA, e-mail: jcb@hawaii.edu

1 • Introduction

Practical signal processing is based on the fast Fourier transform (FFT), an efficient algorithm for computing a discrete Fourier transform (DFT). While the FFT has a speed advantage, it is limited to data record lengths that are powers of 2. The slower DFT can operate on data records of any length. Except for the restriction on an FFT, *FFT* and *DFT* are interchangeable.

An FFT provides a spectral representation of a data record, a finite amount of sampled data taken from a stream of data. Because the data record has a finite length, the spectrum provided by an FFT is almost always corrupted by leakage. Data windows can reduce leakage.

When used with optimum data windows, an FFT can provide accurate estimates of sine wave parameters (tones). It can provide good estimates in the presence of moderate noise and moderate non-stationarity. This has practical application with

Copyright © 2004–2006 John C. Burgess. All rights reserved. No part of this manuscript may be published in any form without the prior written permission of the copyright owner.

[1]Emeritus. Home address: 1269 Mokulua Dr., Kailua, HI 96734, USA

quasi-stationary sounds, such as music, speech, bio-acoustics, and acoustic probes transmitted through the air, ocean, and biological specimens.

2 • The FFT (and DFT)

There are two distinct sets of equations. The set used here,

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1}\mathbf{X}(k)\mathrm{e}^{\mathrm{j}2\pi kn/N},\ 0 \le n \le N-1, \tag{1}$$

$$\mathbf{X}(k) = \sum_{n=0}^{N-1}x(n)\mathrm{e}^{-\mathrm{j}2\pi kn/N},\ 0 \le k \le N-1, \tag{2}$$

is based on a *z-transform* derivation (called "the discrete Fourier transform" by Oppenheim and Schafer [1]), where n is an integer time index, k an integer frequency index, $x(n)$ a discrete data record, $\mathbf{X}(k)$ a (complex) discrete spectrum, and N the length of a data record as well as the FFT (block) size. FFT designates the direct transform, Eq. (2), and IFFT designates the inverse transform, Eq. (1).

Another set of equations sometimes encountered is based on a *Fourier series* derivation [2] (Randall, Chapter 3, Eqs. (10), (11)), in which the scaling factor $1/N$ is applied to the FFT instead of the IFFT. *This slight difference can create difficulty in rationalizing published results.*

A Fourier series is defined as a one-sided, frequency-discrete function [2]. This deceptively implies that the discrete values are all that one has, and leads to terms such as the "picket-fence" effect.

2.1 An FFT Spectrum Is Not an Analog Spectrum

While an FFT spectrum has a superficial resemblance to an analog spectrum, the differences require careful interpretation. Leakage, data record, window, Nyquist frequency, and aliasing, important with digital spectra, have no meaning with analog spectra. Filter passband characteristics, important with analog spectra, have only limited meaning with digital spectra. *Reading a digital spectrum as if it were an analog spectrum can lead to significant interpretation error.*

2.2 The z-Transform and the FFT

The z-transform of a time-discrete data record is frequency-continuous, but periodic with period f_s, the sampling rate (or sampling frequency) [1]. This changes the signal analyzed into a new, periodic sequence. The FFT is defined as a sampled

z-transform. The signal "seen" by an FFT is a repeated data record, with period $T_r = N/f_s$, where T_r is the time length of a data record.

Both the z-transform and the FFT provide two-sided spectra, a feature essential to understanding leakage and its control. The limits 0 to $N-1$ on Eq. (2) are equivalent to $-N/2+1$ to $N/2$, meaning that half of the FFT represents positive frequencies and the other half negative frequencies.

These characteristics have practical effects:

- An FFT can be interpolated to whatever degree of spectral detail is desired (of the data record; seldom the true signal: Section 2.7 in [3]).
- If a data record includes an entire pulse (e.g., a finite impulse response), there is no leakage, and an FFT can provide a true spectrum of the pulse.

2.3 The Nyquist Frequency and Aliasing

The Nyquist frequency is defined as half the sampling rate, $f_{Ny} = f_s/2$. Data from an analog signal at frequencies higher than f_{Ny} are mirrored down in frequency and corrupt data at lower frequencies. For example, if the sampling rate is 10 kHz, the spectrum of a tone at 7 kHz in an analog signal will appear in the FFT spectrum centered at 3 kHz. *Once an analog signal is sampled, there is no way to separate aliased data from true data.*

There are only two ways to prevent aliasing: set the sampling rate high enough, $f_s > 2f_{max}$, or use low-pass (anti-aliasing) filters. Anti-aliasing filters have a disadvantage: the FFT spectrum will be corrupted by the filter's amplitude and phase distortion.

2.4 Frequency Relationships

Spectral plots can be presented as functions of frequency f, frequency index k, or normalized frequency, f_{norm}. Frequency and frequency index are related by $f(k) = kf_s/N$. Normalized frequency is defined by $f_{norm} = f(k)/f_s = k/N$. Corresponding frequency ranges are $-f_{Ny} \le f \le f_{Ny}$ for the z-transform, $-N/2 \le k \le N/2$ for an FFT, and $-0.5 \le f_{norm} \le 0.5$ for normalized frequency.

3 • Leakage

Leakage is a significant problem. It does not depend on whether the true signal is periodic or random. It can easily lead to misinterpretation.

For example, when a signal generator creates a pure tone, humans "hear" the tone and are trained to "see" a line spectrum,

which has a single value at the tone frequency. But the z-transform of a data record shows tone energy spread over the entire spectrum. This is leakage. If a data record of a pure tone is not truncated exactly at an integer multiple of the tone period, an FFT will have N non-zero spectral values, none of which accurately represent tone amplitude, frequency, and phase (Section 5). Leakage can corrupt band-pass filtering (e.g., octave-band analysis) of signals containing tones.

3.1 What Causes Leakage?

An unmodified data record is "seen" through an open window (often called a rectangular window). But what about the unknown data outside the window? The misleading concept of a rectangular window implies that data outside the window are zero. This is usually incorrect in fact. It is also usually incorrect in a different way when an FFT is used. An FFT treats a data record as if it were periodically extended, not zero beyond its "ends" [3]. Almost always, this results in sharp discontinuities at the "ends" of the periodically extended data record. This results in leakage.

3.2 Can Leakage Be "Cured"?

Leakage cannot be "cured" with data records taken from extended signals. But it can be controlled by using a data window that depresses spectral window side lobes.

An attractive idea to "cure" leakage is to "taper the window gradually to zero at its ends". But this is not a necessary and sufficient requirement. While the Hann window (often called Hanning) does taper to zero at its ends, the Hamming and other optimum windows do not (Table 1).

TABLE 1 One-sided Fourier Series Coefficients (FSC) that define several data windows [1,12]

Name	$S(k)$	SAR (dB)	B_0	B_1	B_2	B_3
open	1.625	−13.3	1.0			
Hann(ing)	3.743	−31.5	0.5	−0.5		
Hamming	3.82	−42.7	0.54	−0.46		
opt43	3.946	−43.2	0.53836	−0.46164	0	
opt50	4.277	−50.0	0.50577	−0.47539	0.01565	−0.00319
Blackman	5.65	−58.1	0.42	−0.05	0.08	
opt59	5.075	−59.7	0.461445	−0.491728	0.046827	0
opt71	5.907	−71.5	0.424379	−0.497340	0.078281	0
opt80	6.644	−80.0	0.399770	−0.496413	0.101353	−0.002464
opt90	7.315	−90.0	0.379212	−0.493342	0.121205	−0.006241
opt98 [13]	7.929	−98.2	0.363579	−0.489179	0.136602	−0.010640

FSC are normalized to make data window center magnitudes 1.0. S is given in units of k as a real number. Values of B_i stated as 0 are zero in fact; those not stated are arbitrarily set to zero. The opt43 window is a more accurate Hamming window

4 • Windows

Practical windows are time-discrete functions of length N. Two kinds are used for spectrum analysis: data windows and lag windows. Data windows are used in the ordinary time domain, t, to modify a data record. Their two primary functions are to control leakage and provide insight into the parameters of tones (sinusoidal signal components). Data windows can also be used to design chirps for system identification [4] and for FIR filter design [5].

Lag windows [6] are used to modify correlation functions in the lag time domain, τ. Their primary function is to smooth fluctuations in the spectra of broadband signals. While data windows provide some smoothing, it is not significant.

The remainder of this section concentrates on the two primary functions of data windows.

4.1 Data Windows

The basic window is the open window, which is no window at all. It passes everything in a data record without change, altered only by leakage. All other windows, called data windows, modify an open window. There are a plethora of them [7].

Of all the possible applications for data windows, the two primary functions (Section 4) are the most useful. Attractive windows perform both functions. These windows are defined with one-sided Fourier series coefficients (FSC),

$$b(n) = \sum_{i=0}^{M-1} B_i \cos(2\pi i n/N), \ 0 \le n \le N-1, \tag{3}$$

where i is an integer, B_i the FSC, M their number, and N is even. Data windows are implemented by block multiplying a data record by the window,

$$\tilde{x}(n) = b(n)x(n), \tag{4}$$

where $\tilde{x}(n)$ represents a windowed data record.

4.2 Spectral Windows

The spectral result of Eq. (4) is mathematically equivalent to convolving the spectral window (spectrum of the data window) with the spectrum of the data record,

$$\tilde{\mathbf{X}}(k) = \mathbf{B}(k) * \mathbf{X}(k), \tag{5}$$

where $\tilde{\mathbf{X}}(k)$ represents the windowed spectrum, $\mathbf{B}(k)$ the spectral window, and $*$ represents convolution.

The practical meaning of Eq. (5) is this: the spectral features of a data window depend only on the FSC that define it, not on

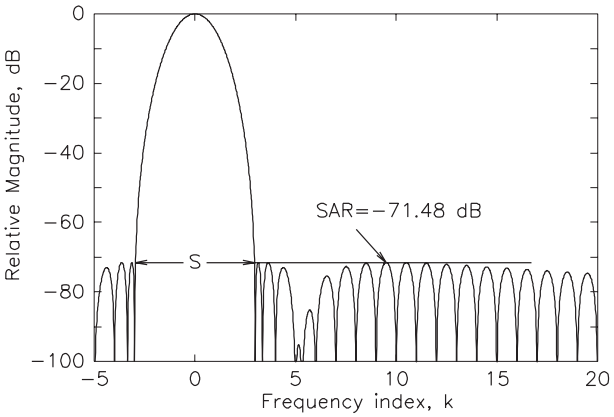


FIGURE 1 Spectrum ($20\log_{10}|\mathbf{B}|$) of opt71 optimum data window (from [12]).

its use. The most important features of a spectral window are its selectivity, S (width of its main lobe), and its selectivity amplitude ratio, SAR (minimum depression of its side lobes) (Figure 1). Selecting a window is a trade-off between S and SAR, since it sets both.

There are two ways to define windows: with a few FSC or with equations. Table 1 lists S, SAR, and the FSC for several windows defined by a few FSC. Three windows defined by equations are the Dolph–Chebyshev [8–10], Kaiser [1,10,11], and Gaussian [10] windows.

4.3 Optimum Data Windows

All data window spectra are characterized by a main lobe and many side lobes (Figure 1). They are interdependent. Reducing side lobe amplitudes automatically increases main lobe width. An optimum data window must not only depress side lobes to control leakage, it must minimize main lobe width. Among the plethora of window shapes, only a few are optimum.

The Dolph–Chebyshev window is the best. It has the narrowest possible main lobe, and depresses all side lobes to the same amplitude. It has a computational disadvantage: $M = N$ different FSC are required for each side lobe amplitude depression and for each block size N.

The Kaiser and Gaussian windows not only have the same computational disadvantage as the Dolph–Chebyshev window, their main lobe widths S, like the Hann(ing) and Blackman windows (Figure 2), are greater than those for corresponding Dolph–Chebyshev windows at the same SAR.

Optimum data windows [12] depress side lobe amplitudes to a specified maximum value, have a main lobe width, S, nearly as narrow as a Dolph–Chebyshev window with the same SAR,

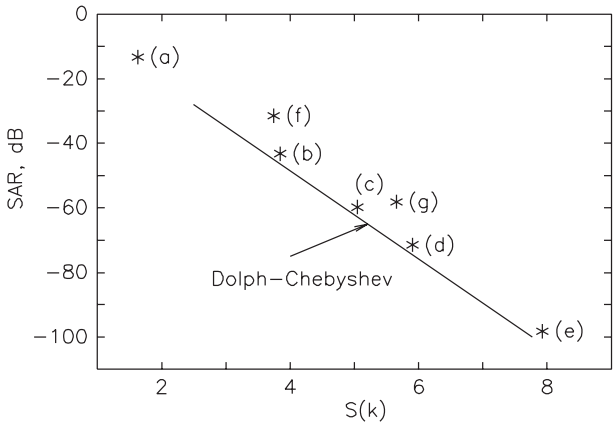


FIGURE 2 Selectivity S versus selectivity amplitude ratio SAR for various windows compared with Dolph–Chebyshev windows: (**a**) open; (**b**) Opt43 and Hamming; (**c**) opt59; (**d**) opt71; (**e**) opt98; (**f**) Hann(ing); (**g**) Blackman [12, Figure 2].

and the FSC are the same for any block size. Only four FSC are needed to depress side lobe amplitudes by 98 dB. While the FSC are intended for large N, they can provide useful results for $N \geq 16$. Figure 2 shows the relationship of various windows to Dolph–Chebyshev windows.

5 • Estimation of Amplitude, Frequency, and Phase

The amplitude, frequency, and phase angle of a pure tone are associated with the peak value of the main lobe in its spectrum (Figure 3). While discrete FFT values $X[k]$ (Fourier coefficients, often misleadingly called "bins") almost never provide this information, optimum data windows make it possible to extract accurate estimates.

5.1 Data Record

The data record for the mth single tone of frequency f_m can be expressed as

$$x_m(n) = A_m \cos(2\pi mn/N - \alpha_m), \ 0 \leq n \leq N-1, \tag{6}$$

where $m/N = f_m/f_s$, A_m is the signal amplitude, and α_m is the signal phase angle. The variable m is the (real) number of tone periods in the data record. It can be expressed by

$$m = \frac{T_r}{T_m} = p + q, \tag{7}$$

where (Figure 3) p is the integer part of m, q is the fractional part of m, $T_r = N\Delta T$ is the data record period ($\Delta T = 1/f_s$ is the

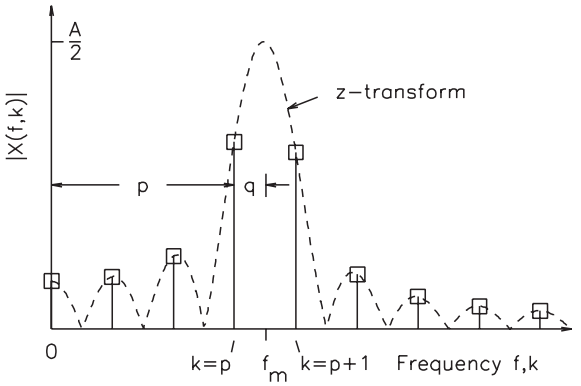


FIGURE 3 Pure tone spectrum for $m = 3.486$, showing p, q, and the principal values at $k = p$ and $k = p + 1$. The *solid vertical lines* are FFT values. These and the z-transform illustrate leakage [3, Figure 12].

sampling interval), and $T_m = 1/f_m$ is the tone period. *Estimating q is the key to the procedure* (Section 5.3) [14].

5.2 Spectrum

The two-sided spectrum of a pure tone (z-transform or FFT) is mathematically the complex vector sum of two contributions, a direct contribution with its main lobe centered on the tone frequency f_m in the positive range and an image contribution centered on $-f_m$ in the negative range [14]. There appears to be no known way to separate the two after computation of an FFT.

The side lobes of each contribution corrupt the main lobes of the other. Optimum data windows provide a way around this effect by depressing the side lobes of one contribution within the main lobe of the other. When all side lobes are sufficiently depressed, the relatively simple direct contribution can provide a good approximation to the FFT [14, Eq. (3)]

$$\tilde{X}(k) \cong \tilde{X}(k)_{\text{direct}}. \tag{8}$$

Only two complex values of $\tilde{X}(k)$, called principal values, are required to obtain accurate estimates for tone amplitude, frequency, and phase angle. The principal values are the two values of $\tilde{X}(k)$ on either side of a tone frequency, at $k = p$ and $k = p + 1$ (Figure 3),

$$\tilde{\mathbf{X}}(p) = \tilde{X}(p)\mathrm{e}^{\mathrm{j}\tilde{\phi}(p)}, \tag{9}$$

$$\tilde{\mathbf{X}}(p+1) = \tilde{X}(p+1)\mathrm{e}^{\mathrm{j}\tilde{\phi}(p+1)}. \tag{10}$$

A single tone has two main lobes, one in the direct contribution, one in the image contribution. When the tone frequency is near

either *0* or f_{Ny}, the two main lobes may overlap. The following procedure applies as long as the main lobes are sufficiently separated (Figure 10 in [14]). The minimum required separation of the main lobe centers is

$$\Delta k_{\min} \geq \frac{S}{2} + 1 + \Delta k_{\mathrm{shift}}, \tag{11}$$

where S is the value of window *selectivity* (Table 1), and $\Delta k_{\mathrm{shift}}$ allows for quasi-stationarity and noise [14]. While $\Delta k_{\min}$ controls the minimum value of N required to properly analyze a single tone, values of N used in practice normally are much larger.

5.3 Procedure

The measured ratio of the two principal values is defined as

$$R = \left|\tilde{X}_p\right| \Big/ \left|\tilde{X}_{p+1}\right|. \tag{12}$$

While small values of N can be used, increasing N decreases error by increasing the separation of direct and image main lobes. Error for a single tone signal is a minimum when the frequency is $f_m = f_s/4$, which positions it midway between zero and Nyquist frequencies. Error can be decreased significantly by using optimum windows (Table IV in [14]).

For a multi-coefficient window, R can also be specified as a function of q and the window coefficients B_i [14],

$$R = \left| \frac{q}{1-q} \sum_{i=0}^{M-1} \frac{B_i}{q^2 - i^2} \Bigg/ \sum_{i=0}^{M-1} \frac{B_i}{(q-1)^2 - i^2} \right|, \tag{13}$$

which can be written as a polynomial in q,

$$\sum_{j=0}^{J-1} \sum_{i=0}^{M-1} a_{ji} B_i q^j = 0. \tag{14}$$

Only one root of q will be in the range $0.0 \leq q < 1.0$. Optimum window coefficients B_i can be found in Table 1, and coefficients a_{ji} can be found in [14].

Accurate approximations for tone amplitude, phase angle, and frequency are given by

$$A \simeq \frac{2\pi \tilde{X}_p}{Nq \sin \pi q} \frac{1}{\sum_{i=0}^{M-1} B_i \frac{1}{(q^2 - i^2)}}, \tag{15}$$

$$\alpha \simeq -\tilde{\phi}_p + \pi q \frac{N-1}{N} \simeq -\tilde{\phi}_p + \pi q, \tag{16}$$

$$f \simeq \frac{p+q}{N} f_s. \tag{17}$$

5.4 Multiple Tones

Multiple tones create multiple main lobes in the direct and image contributions. The single-tone procedure applies to each tone provided that all direct and image contributions meet the minimum separation requirement, $\Delta k_{\min}$, Eq. (11). Error may increase as a result of multiple side lobe interactions with principal values.

For each tone, determine $|m| = p$ by inspection of the spectrum, get the principal values, calculate the ratio R_m, determine q, and then the tone parameters.

5.5 Quasi-stationary Signals and Noise

When used with non-stationary signals or signals corrupted by noise, an FFT can provide only approximate values for the spectrum of the true, underlying signal. With quasi-stationary signals, there will be time intervals when the errors are small enough that results are useful. A reasonable estimate for Δk_{shift} (Eq. (11)) is necessary. Reference [14] includes discussion of expected errors resulting from moderate non-stationarity and noise.

6 • Applications

Two applications will illustrate how optimum windows can be used to provide useful data.

6.1 Probing Physical Systems

While one or more pure tones are useful as probes through physical media, the signal measured will almost always be quasi-stationary and corrupted by noise. An example concerning sound transmission through the atmosphere is given in [14].

Probe signals composed of one or more pure tones must have a minimum data record size N, a minimum sampling rate, $(f_s)_{\min}$, and are most efficiently analyzed with optimum data windows. Data record size N can be restricted by experimental factors, such as an upper limit on signal time length (e.g., the presence of echos), whether the signal is stationary or non-stationary, and the presence of unwanted experimental artifacts in the measured signal (e.g., power glitches [14]).

When the probe signal consists of a single tone, the minimum data record size is $N_{\min} = 2k_{\min}$. $N_{\min}$ varies from 4 for the open window to 10 for the opt98 window (N will normally be greater).

As an example, consider a single tone probe signal with the conditions $f_m = 1\,\text{kHz}$, and $T_r \leq 0.1\text{s}$. To minimize the effects of image side lobes, f_m should be placed at $f_{\text{Ny}}/2 = f_s/4$. This sets $f_s = 4\,\text{kHz}$ as the minimum sampling rate. The sampling interval $\Delta T = 1/f_s = 0.25\,\text{ms}$. The condition $T_r \leq 0.1\text{s}$ sets $N_{\max} = 0.1/\Delta T = 400$. Thus $N = 256$ is the largest FFT block size meeting the test

conditions. The requirement for $\Delta k_{\min}$ is more than met for any optimum window. A general method for designing multi-tone probe signals, with examples, is given in [14].

6.2 System Identification

Efficient waveforms for system identification must have a low peak factor and a broad spectrum. Maximum length sequences (MLSs) are well known. Linear chirps provide another. A linear chirp has the advantage over an MLS sequence that its power can be concentrated in a specified frequency range. But, when used with an FFT, it suffers from leakage. By using optimum windows, leakage in a repeated chirp, called a *choip,* can be minimized, and a choip becomes a useful tool for system identification. Details are given in [4] [and Chapter 4 by Ning Xiang].

References

1. A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1989.
2. R. Bracewell, *The Fast Fourier Transform and Its Applications,* 2nd Ed., Rev. McGraw-Hill Book Company, New York, NY, 1986.
3. J. C. Burgess, "Practical considerations in signal processing", in *Encyclopedia of Acoustics* (M. Crocker, ed.), vol. 3, ch. 101, pp. 1261–1279. Wiley, New York, 1997.
4. J. C. Burgess, "Chirp design for acoustical system identification", *J. Acoust. Soc. Am.*, vol. 91, pp. 1525–1530, 1992.
5. J. Proakis and D. Manolakis, *Introduction to Digital Signal Processing*. Macmillan, New York, 1988.
6. G. Jenkins and D. Watts, *Spectral Analysis and Its Applications*. Holden-Day, San Francisco, 1969.
7. F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform", *Proc. IEEE*, vol. 66, pp. 51–83, 1978.
8. C. Dolph, "A current distribution for broadside arrays which optimizes the relationship between beam width and side-lobe level", *Proc. I.R.E*, vol. 34, pp. 335–348, June 1946.
9. H. Helms, "Nonrecursive digital filters: design methods for achieving specifications on frequency response", *IEEE Trans. Audio Electroacoust.*, vol. AU-16, pp. 336–342, 1968.
10. R. A. Roberts and C. T. Mullis, *Digital Signal Processing*. Addison-Wesley, Reading, MA, 1987.
11. L. Rabiner, B. Gold, and C. McGonegal, "An approach to the approximation problem for nonrecursive digital filters", *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 83–106, 1970.
12. J. C. Burgess, "Optimum approximations to Dolph–Chebyshev data windows", *IEEE Trans. Signal Process*, vol. 40, no. 10, pp. 2592–2594, 1992.

13. A. Nuttall, "Some windows with very good sidelobe behavior", *IEEE Trans. Acoust., Speech Signal Process*, vol. ASSP-29, pp. 84–91, 1981.
14. J. C. Burgess, "Accurate analysis of multitone signals using a DFT", *J. Acoust. Soc. Am.*, vol. 116, pp. 369–395, 2004.

5

Measuring Transfer-Functions and Impulse Responses

Swen Müller

National Institute of Metrology, Xerém, Brazil

1 • Introduction

The impulse response (IR) and its associated Fourier transform, the complex transfer function (TF), describe the linear transmission properties of any system able to transport or transform energy in a certain frequency range. As the name suggests, the IR is the response in time at the output of a system under test when an infinitely narrow impulse is fed into its input (Figure 1). Any other signal present at the input will be *convolved* with the system's IR $h(t)$:

$$o(t) = \int_{-\infty}^{\infty} h(\tau) \cdot i(t-\tau)\,\mathrm{d}\tau = h(t) * i(t), \tag{1}$$

with $i(t)$ = input signal, $h(t)$ = system IR, $o(t)$ = output signal, * = convolution symbol.

This corresponds to *multiplying* the signal's spectrum with the system's complex *transfer function* $H(f)$:

$$O(f) = H(f) \cdot I(f). \tag{2}$$

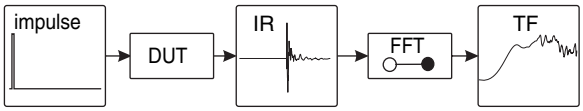


FIGURE 1 Basic setup to measure the IR of a device under test and derive its TF.

As the IR can be transformed into the TF via Fourier transform and back again into the IR via the inverse Fourier transform, both are equivalent and carry the same information, which can be extracted and visualized in different ways.

In the frequency domain, the best known is the *magnitude response*, which is the modulus of the complex TF and is most commonly plotted on a logarithmic (dB) *y*-scale. In acoustics, the frequencies on the *x*-axis are also normally plotted in a logarithmic fashion, which gives the well-known double-logarithmic display. The phase, which is the arc tangent of imaginary part divided by the real part of the complex TF, gives additional valuable information, for example to find the exact resonance frequency of a transducer. Sometimes even more important is the *group delay*, which is proportional to the negative derivative of the phase vs. frequency:

$$\tau_G = -\frac{d\varphi}{d\omega} = -\frac{d\varphi}{df \cdot 2\pi}. \tag{3}$$

The group delay is an estimate of the delay which each frequency suffers when traveling through the investigated system.

The transfer function often involves transformation from one physical entity to another, as is the case for transducers (voltage or current to sound pressure and vice versa) or in acoustics for solid materials (force to acceleration or velocity).

In room acoustics, the *room impulse response* between a sender (loudspeaker or impulsive noise source) and a receiver (microphone) plays a crucial role. It is used to evaluate a comprehensive set of parameters (most notably the frequency-dependent reverberation time) to characterize the acoustic properties of a room.

2 • Measurement Methods

Generally spoken, every kind of excitation signal can be used to determine the IR and the respective TF of a linear and time-invariant system, provided that it contains enough energy at every frequency of interest to overcome the noise floor. As can easily

be seen from equation (2), the TF can be obtained by *dividing* the output spectrum of the system under test by the input spectrum:

$$H(f) = \frac{O(f)}{I(f)}. \tag{4}$$

This usually implies Fourier-transforming input and output signals in order to perform the complex division in the spectral domain.

Alternatively to the processing in the spectral domain to obtain the TF, the IR can be retrieved directly in the time domain by convolving the DUT's response with the *inverse filter*. In case of a white excitation signal, this inverse filter simply is the *time-reversed* excitation signal (sometimes called *"matched filter"*; Figure 2). Convolution with a time-reversed signal is equal to *cross-correlation*. This plays an important role when using maximum-length sequences (MLS) as excitation signal, a special class of binary sequences with white spectrum presented in Chap. 6, because a very efficient cross-correlation algorithm operating in the time domain is available for them.

When using a non-white excitation signal, the general form to construct the inverse filter (sometimes misguidingly called *"mismatched filter"*) is to Fourier-transform it, invert its spectrum (put the modulus upside down and negate the phases) and back transform it to the time domain (Figure 3). The convolution itself is also most efficiently performed by Fourier-transforming the DUT's response to the spectral domain where the convolution is replaced by the complex multiplication with the inverse filter spectrum and then performing an IFFT, which yields the IR.

All efficient modern methods for obtaining IR and TFs rely on the use of the fast Fourier transform (FFT) to switch between the

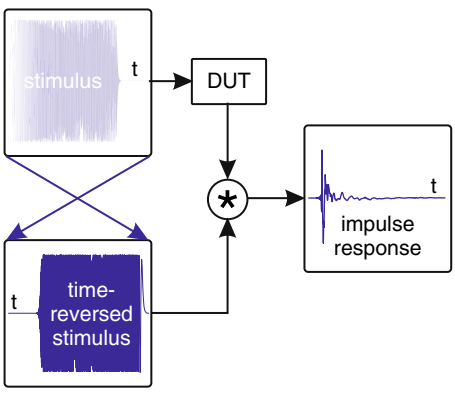


FIGURE 2 In the case of white excitation signals (here, a linear sweep has been used as example), the DUT's impulse response can be calculated by performing a cross-correlation, which is the convolution (denoted by the * *symbol*) with the time-reversed excitation signal.

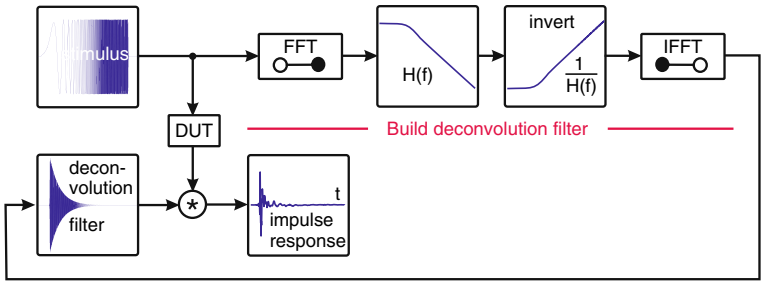


FIGURE 3 For any non-white excitation signal, the evaluation of the impulse response in the time domain can only be done by convolution with the inverse filter, which itself should be constructed in the frequency domain to yield exact results.

digitized time signal and its spectrum in the frequency domain. This has a few implications:

1. The analyzed signal must of course be band-limited to $f_S/2$ (f_S = sampling frequency) before digitalization. Care must be taken near the Nyquist frequency $f_S/2$ because the typical halfband digital filters used in modern oversampling AD converters only exhibit a fair attenuation (typically 6 dB) at $f_S/2$. Thus, a narrow frequency strip around $f_S/2$ usually contains invalid aliasing products.
2. Only signal parts with a length of 2^N samples (N = integer number) can be processed (at least for the commonly used decimation-in-2 algorithms).
3. The FFT only yields *exact* results if the signal part analyzed repeats periodically. This automatically means that the signal's spectrum becomes discrete and indeed only contains the frequencies given by $n \cdot f_S/2^N$, with N = FFT order and $n = 0, 1, 2, 3 \ldots, 2^{N/2}$.

As a consequence, the white noise constructed in the frequency domain shown later in Figure 10 will only exhibit the desired flat horizontal line in the magnitude spectrum if repeated periodically. It then consists of discrete frequencies with indeed all having exactly the same amplitude.

Conditions 2 and 3 frustrate the direct application of an FFT to maximum-length sequences (focused in Chap. 6), as these have a period length of $2^N - 1$. There is thus one value missing to complete the FFT block length and setting it to whatever value will lead to gross aberrations from the flat magnitude spectrum of a periodically repeated MLS.

While the periodic repetition of a deterministic noise signal is crucial to establish the desired spectral distribution, single pulses

or sweeps do not rely on periodic repetition to establish a smooth spectrum. Thus, the latter two have to be ejected only once in a TF measurement, while deterministic periodic noise should be switched on at least one period before starting the actual measurement in order to establish a steady-state condition.

2.1 Dual Channel FFT Analysis

Instead of working with deterministic excitation signals, it is also possible to use an asynchronous signal source. As its output is not predictable, it is necessary to simultaneously capture the input and output signals of the DUT in each measurement cycle and process them in the frequency domain after FFT to obtain the TF. This leads to the classical two-channel FFT analyzer arrangement (see [1] for a comprehensive description) depicted in Figure 4.

A classical dual channel analyzer captures both the input (channel A in Figure 4) and the output signals of the DUT (channel B in Figure 4). The signal is cut in contiguous or overlapping segments which are windowed and then transformed to the spectral domain via FFT. The two resulting spectra are involved in three averaging processes: two for the autospectra

$$G_{AA}(f) = \frac{1}{n}\sum |A(f)|^2 \quad \text{and} \quad G_{BB}(f) = \frac{1}{n}\sum |B(f)|^2$$
$$(n = \text{number of averages}), \tag{5}$$

which are the squared modulus of the spectra for channels A and B, and one for the cross-spectrum

$$G_{AB}(f) = \frac{1}{n}\sum A^*(f)\cdot B(f), \tag{6}$$

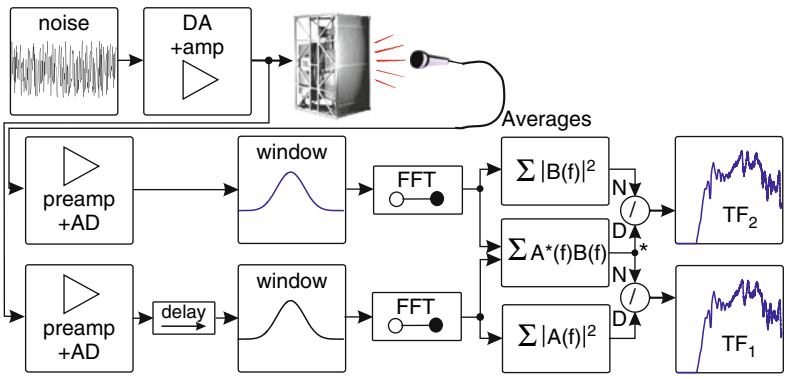


FIGURE 4 Simplified block diagram of classical two-channel FFT analysis.

which is the complex conjugate of the spectrum for channel A (spectrum A with its phases negated) multiplied with the spectrum for channel B.

While the two autospectra are real-valued and add up both signal and noise power in every measurement run, the cross-spectrum is complex, and uncorrelated noise with its random phase spectrum tends to average out with increasing number of averages.

The complex transfer function defined by

$$H(f) = \frac{B(f)}{A(f)} \tag{7}$$

can now be computed in two ways. Multiplying both the numerator and the denominator of the right side of the above equation with $A^*(f)$ yields

$$H_1(f) = \frac{A^*(f)B(f)}{A^*(f)A(f)} = \frac{G_{\mathrm{AB}}(f)}{G_{\mathrm{AA}}(f)}, \tag{8}$$

so $H(f)$ can be estimated by the division of the cross-spectrum with the autospectrum of the DUT's input. On the other hand, multiplying both the numerator and the denominator of the right side with $B^*(f)$ yields

$$H_2(f) = \frac{B^*(f)B(f)}{B^*(f)A(f)} = \frac{G_{\mathrm{BB}}(f)}{G^*_{\mathrm{AB}}(f)}, \tag{9}$$

so H(f) can also be computed by the division of the autospectrum of the DUT's output with the complex conjugated cross-spectrum. When noise is present neither in channel A nor in channel B, both methods obviously lead to identical results. With noise, however, the two results differ. $H_1(f)$ is a better estimate of the true TF when output noise prevails (the more typical case), while $H_2(f)$ comes closer to reality when input noise dominates.

To determine the amount of uncorrelated noise in the measurement, the two estimates of the transfer function can be divided:

$$\nu^2(f) = \frac{H_1(f)}{H_2(f)} = \frac{|G_{\mathrm{AB}}|^2}{G_{\mathrm{AA}} \cdot G_{\mathrm{BB}}}. \tag{10}$$

This *coherence function* assumes values between 1 and 0. When no uncorrelated noise is present, the coherence function will become 1, as $H_1(f)$ and $H_2(f)$ are identical. When there is only uncorrelated noise without signal, the coherence function drops to near

0 because the noise tends to average out in the cross-spectrum, while it is being added up energetically in the autospectra.

With the help of the coherence function, the power of the autospectrum G_{BB} (derived from the DUT's output) can be split into the coherent power $\nu^2(f) \cdot G_{BB}(f)$, which originates from the excitation signal passed through the DUT, and the non-coherent power $(1-\nu^2(f)) \cdot G_{BB}(f)$ originating from uncorrelated noise.

The signal-to-noise ratio becomes thus

$$\mathrm{SNR} = \frac{\nu^2(f)}{1-\nu^2(f)} \tag{11}$$

and it is an interesting feature of the dual channel analysis technique that this value is available without turning off the excitation signal.

Traditionally, FFT analyzers are operated with asynchronous noise sources. However, the use of an asynchronous excitation signal has some implications:

As already mentioned, the FFT yields correct results only for signals repeated with a period equal to the FFT block length. For non-periodic signals, start point and end point of the analyzed signal section generally do not match. This discontinuity introduces a considerable error, the famous *leakage*, which has to be lowered by windowing the analyzed sections prior to the FFT. However, windowing introduces a DC bias error and has a smoothing effect which reduces the linear spectral resolution, causing the *relative* spectral resolution (in fractions of an octave) to diminish toward low frequencies.

In measurement setups involving an acoustic transmission path, the propagation delay between sending and receiving transducer has to be carefully compensated with a digital delay of the same duration inserted in the analysis path for the direct signal. This is to guarantee that the same signal sections are submitted to the FFTs in every measurement cycle. The window attenuates all components outside the center, which can lead to errors when the DUT features a strong frequency-dependent group delay, as is the case for reflections or at frequencies near resonances.

The asynchronous noise generator commonly used as signal source has a white (or pink or other specified) spectrum when averaged over a long time, but a single snapshot of the noise signal has a very irregular spectral structure suffering from deep magnitude dips. Thus, a classical dual channel analyzer has to average over many individual measurements before being able to present a reliable result.

There is an interesting application which exploits the fact that any arbitrary signal source can be used, provided that at least in the long term, it covers all frequencies of interest: the unobtrusive measurement of a sound system during a performance, using the program material itself as the excitation signal. However, speech or music with their erratic spectral distribution require even longer averaging periods to achieve a meaningful result compared to asynchronous noise.

When the unobtrusiveness is not paramount, it is advantageous to replace the asynchronous noise (or other) signal source with a synchronous predetermined excitation signal that features a stable smooth spectrum over the whole frequency range of interest for the given FFT length. This can be a periodic noise or a zero-padded aperiodic sine sweep. Using these designed signals allows simplifying and considerably speeding up the entire analysis. First of all, due to the periodicity of the periodic noise (or the 0 start/end points for the sweep), there is no leakage and thus no need to window the DUT's input and output signals prior to the FFT. Furthermore, the inconvenient tunable delay to adjust the arrival time can be omitted (provided that the FFT block size accommodates all delayed components and, in case of the sweep, a sufficiently long gap of silence is introduced at the end). Any delay will be represented by the start position of the IR obtained by IFFT. Finally, as the excitation signal is smooth and has no dips at any frequency, a single measurement is sufficient to obtain a reliable result. Thus, the averaging process is normally not needed, provided that the SNR is not too poor. Without averages, however, the coherence function becomes meaningless.

Furthermore, using a synchronous deterministic excitation signal means that distortion components will always appear at the same locations in the estimated IR and hence do not diminish with increasing number of averages (Figure 5). In contrast, asynchronous noise gives the best linear fit of the transfer function in the long run (e.g., when performing many averages). Noise averages out, and so do all distortion products.

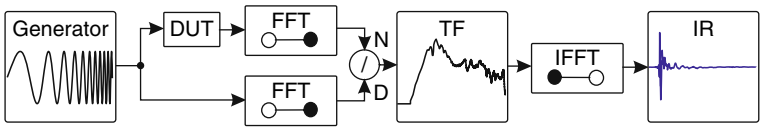


FIGURE 5 Simplified two-channel analysis to capture IR with deterministic excitation signal.

2.2 Single Channel FFT Analysis with Deterministic Excitation Signal

Employing a predetermined synchronous excitation signal that repeats exactly the same way in each measurement cycle means that the spectrum of the signal fed to the DUT will not change from measurement to measurement. In this case, it obviously has to be estimated only once at the beginning of a measurement session. The best way to do this *reference measurement* is to replace the DUT with a wire and pass the excitation signal over the same channel, as depicted in Figure 6. The captured signal is subjected to an FFT and the resulting spectrum inverted, yielding the deconvolution or *reference* spectrum. In the actual measurement of the DUT, this reference spectrum is applied as complex multiplication with the FFT of the DUT's output signal. The product will be the transfer function.

Using the same channel for both the reference and the actual measurement of the DUT has two advantages: It obviously eliminates the need for a second channel, and it leads to very high precision, as any gain error, any frequency response deviation and any delay of the components in the measurement chain (including amplifiers and converters, but not the microphone in the setup in Figure 6) are compensated. The resulting transfer function is actually a comparison to the frequency response of the flat wire used during the reference measurement.

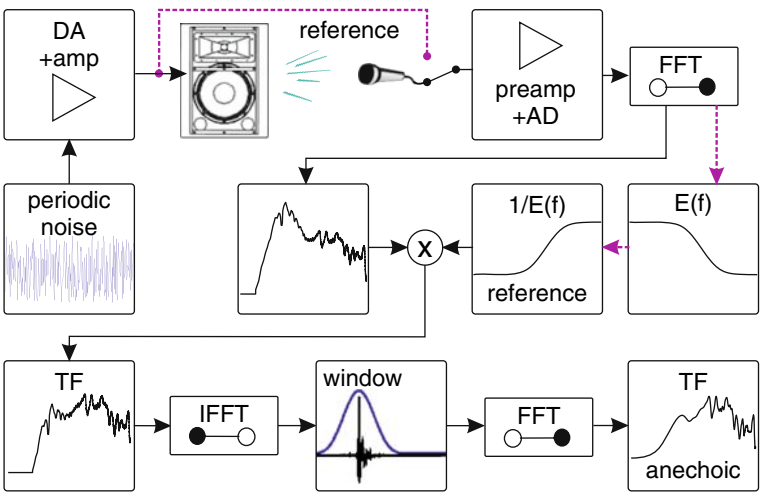


FIGURE 6 Single channel measurement, here with synchronous periodic excitation signal. The *dashed lines* depict the signal flow for the reference measurement. The processing also includes windowing in the time domain to reject unwanted reflection and noise, which allows quasi-anechoic measurements.

When working with a periodic excitation signal, the acquired impulse response will also be periodic. Care must be taken to guarantee that the chosen period is sufficiently long to accommodate all delayed components of the IR. In this case, one period of the periodic IR is virtually equal to the normal aperiodic IR. If the chosen period is too short, the IR's tail will smear into the beginning of the next period, an effect known as *time aliasing* [2].

2.3 Non-periodic Sweeps

Sweeps do not have to be periodically repeated, as their spectrum is similar for a single shot compared to periodical repetition. Thus, they can be used in a non-periodic fashion to perform TF measurements with the setup in Figure 6. In this case, a *linear deconvolution* is most appropriate to obtain the non-periodic IR. The linear deconvolution can be performed most efficiently by padding both the input signal and the output signal of the DUT with zeros in order to double the FFT block length to $2N$. Then both are Fourier-transformed and the output spectrum divided by the input spectrum. An IFFT will reveal the complete deconvolution result with positive and negative times. A noticeable amount of energy can show up in front of the rise of the actual impulse response. These are the distortion products, and it is a special feature of measurements with sweeps that they are all pulled to negative arrival times (relative to the direct sound), where they can be safely discarded, as no causal components of the IR can reside there (Figure 7). The IR itself stays totally free of harmonic distortion products, in contrast to a measurement with pseudo-random noise, where they are distributed as small IR repetitions over the whole period (*distortion peaks*).

For this reason, the dynamic range of an IR acquired with sweeps is virtually only limited by the background noise, whereas an IR obtained with a periodic pseudo-random noise stimulus

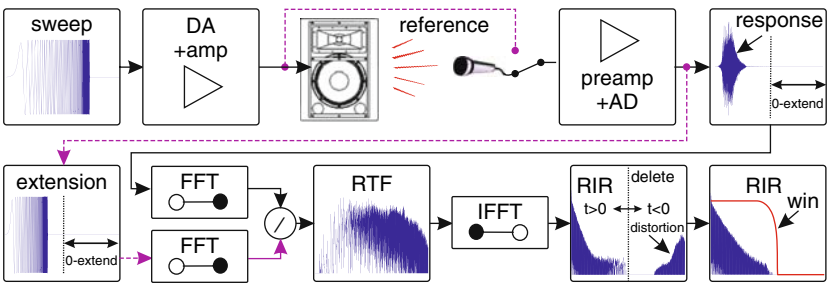


FIGURE 7 Linear deconvolution and separation of distortion products applied for room transfer function (RTF) and room impulse response (RIR) measurements with sweep.

(MLS including) is mostly restricted by the non-linearities of the transducers (most notably the sender). Reducing the level will reduce distortion, but when reducing it too much, background noise may dominate.

Performing synchronous averages helps to defeat the influence of uncorrelated background noise. Every doubling of the number of averages improves the SNR by 3 dB, provided that the undesired components are really of random nature. However, in an averaging process with synchronous periodic excitation, the distortion products are completely correlated with the excitation signal and thus averaging does not help to diminish their presence. This means that a compromise level (where distortion products and background noise have about equal contribution) has to been found by trial and error when working with synchronous noise excitation. In contrast, due to the ability to isolate the distortion products in the estimated IR, sweeps can be emitted with the maximum possible level, easily elevating the achievable SNR by 15–20 dB compared to a measurement with pseudo-random noise.

However, driving the transducer considerably into its non-linear range is only acceptable in a comparative measurement, in which the transducer itself is not the object of investigation. If the transducer's frequency response itself is to be estimated, operating in its non-linear range will of course lead to an apparent loss of sensitivity.

When a logarithmic (pink coloration) sweep is used to measure a transducer in anechoic conditions, a train of discrete spikes antecedes the actual IR. These are "harmonic impulse responses" (HIRs) in rising order from right to left. Due to the constant relative frequency increase (constant fraction of an octave per time unit) of the logarithmic sweep, the deconvolution process focuses the energy of each individual harmonic to a fixed arrival time over the whole frequency span, the same way as it happens for the fundamental (Figure 8). This makes it possible to determine the transfer function and the frequency-dependent level of all harmonics with one single log sweep and a bit of post-processing [3,4]. To do so, the fundamental and each individual HIR are isolated with appropriate windows and separately Fourier-transformed. The position of the HIRs can be found by

$$t_{\mathrm{HIR}(k)} = t_{\mathrm{FUND}} - \frac{\log_2(k)}{\mathrm{Sweep_rate}}, \tag{12}$$

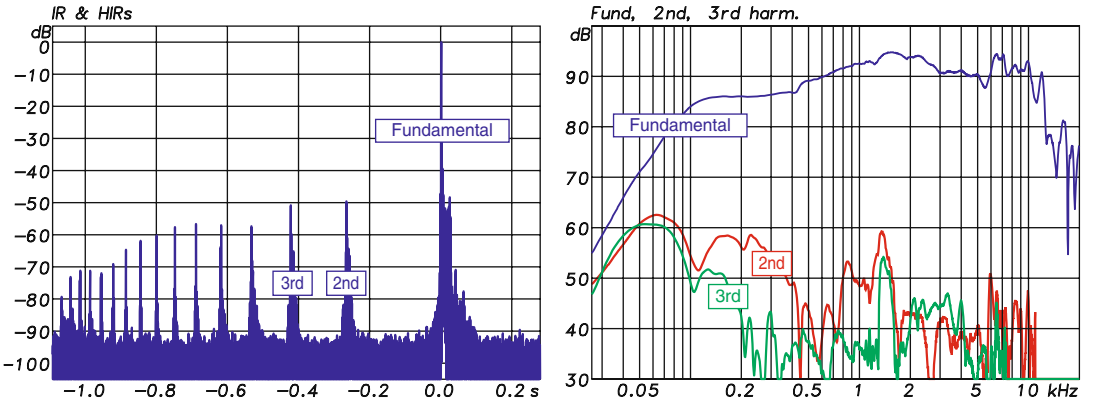


FIGURE 8 Extraction of just the second and third harmonics from a measurement with log sweep.

with $t_{\mathrm{HIR}(k)}$ = arrival time of HIR of order k, t_{FUND} = arrival time of fundamental and

$$\mathrm{Sweep_rate} = \frac{\partial f}{f \cdot \partial t}\ (octaves/s).$$

The spectra of the HIRs are now compressed toward the left side of the frequency axis by a factor corresponding to their order. For the third-order harmonic, for example, the former 3 kHz component will be relocated to 1 kHz, as this is the frequency of the fundamental which caused the 3 kHz component of the third harmonic to appear. To calculate the frequency-dependent percentage of distortion for each harmonic, its relocated spectrum has only to be divided through the spectrum of the fundamental.

Loudspeakers, when pushed hard, are sometimes haunted by subharmonics, most notably half the fundamental. This component can also be detected by the method. Its HIR appears on the right side of the fundamental IR, at the same distance as the second harmonic on the left side. To isolate this component, good anechoic conditions and a slow sweep rate are vital in order to avoid smearing of reverberation into the subharmonic's IR.

Apart from the complete rejection of harmonic distortion, TF measurements with sweeps are also considerably less vulnerable to time variance compared to measurements with (pseudo-random) noise. Time variance occurs when the sound medium is in movement (wind gusts, convection, water currents) or when the DUT itself is fairly time-variant (analogue recording gear). Breaking the exact periodicity of a periodic noise signal immediately disfigures its smooth spectrum, while a sweep suffering the same kind of time variance shows to be fairly immune.

2.4 Time Delay Spectrometry (TDS)

TDS is a vintage method to derive TF with the help of linear (white) sweeps (Figure 9). Devised by Richard Heyser specifically for measuring loudspeakers, it can be realized with purely analog means, which is the reason why it has been rather popular before the advent of powerful personal computers with FFT capabilities.

The mathematic framework of TDS has been extensively covered in [5]. To keep the description of the TDS principle and some practical aspects reasonably short, equations have been omitted here.

A TDS analyzer has a generator that produces both a swept sine and, simultaneously, a phase-locked swept cosine. The sine is fed to the DUT, and its captured response is multiplied separately by both the original sine (to get the transfer function's real part) and the 90° phase-shifted cosine (to get the imaginary part). The multiplier outputs are filtered by a low pass with fixed cutoff frequency. These multipliers work similar to the mixer circuits used in the intermediate-frequency stages of radio frequency receivers based on the ubiquitous super-heterodyne principle, producing the sums and differences of the input frequencies. The sum terms of both multiplier outputs must be rejected by the low-pass filters, whereas the difference terms may pass up to a certain frequency. If both the generated and the captured frequencies are almost equal, the output difference frequency will be very low and thus will not be attenuated by the low-pass filters.

As the sound that propagates from the DUT arrives with a delay at the microphone, its instantaneous frequency will be lower than the one coming directly from the generator. This causes a higher mixer output difference frequency that, depending on the cutoff frequency, will be attenuated by the low-pass filters. For this reason, the generated signal must be *time-delayed* by the propagation time of the sound between loudspeaker and microphone before being multiplied with the DUT's response. In this way, the difference frequency will be near DC. In contrast to the direct sound, reflections will always take a longer way and thus arrive

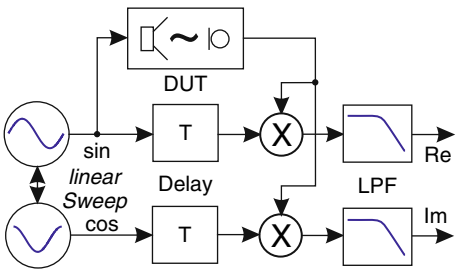


FIGURE 9 Basic TDS signal processing.

with a lower instantaneous frequency, causing higher frequency components in the multiplier outputs, which will be more attenuated by the low-pass filters, the longer their detour is. This is equivalent to windowing the respective IR. To obtain the latter, an IFFT of the complex TF has to be performed.

So, by proper selection of the sweep rate and the low-pass filter cutoff frequency, simulated quasi-free-field measurements are possible with TDS. In addition to the attenuation of unwanted reflections, distortion products, which arrive with a higher instantaneous frequency and thus cause high mixer output frequencies, are also suppressed very well. Extraneous noise in the wide band above the filter cutoff frequency will also be rejected.

The controlled suppression of reflections is the reason why TDS analyzers utilize a linear sweep ($\mathrm{d}f/\mathrm{d}t = \text{constant}$) as the excitation signal. The frequency difference between incoming direct sound and reflection stays constant over the whole sweep range, keeping the attenuation of each reflection frequency-independent.

For certain purposes, however, it could be interesting to use logarithmic sweeps with TDS – at least over some parts of the frequency span. This would lead to an effective IR window width that is inversely proportional to the frequency. This behavior corresponds to constant *relative*-bandwidth smoothing, a technique widely used for processing and displaying loudspeaker frequency responses. In contrast, standard TDS employing a linear sweep has a constant *absolute* bandwidth smoothing effect, just as windowing a broadband IR. Another important benefit of using a (semi-) log sweep would be to enhance the SNR at low frequencies where the linear (white) sweep is often not capable of injecting enough energy to sufficiently overcome background noise.

A common problem with TDS is ripple which occurs at low frequencies. As mentioned previously, the multipliers produce sum and difference terms of the "time-delayed" excitation signal and the incoming response. At higher instantaneous frequencies, the sum is sufficiently high to be attenuated by the output low-pass filter. But at the low end of the sweep range, when the sum is close to or lower than the low-pass cutoff frequency, "beating" will appear in the recovered magnitude response. To remedy this, the sweep can be made very long and the low-pass cutoff frequency reduced by the same factor.

Even then, some spectral ripple often appears at the beginning and at the end of the sweep frequency range because of the sudden onset and termination of the linear sweep. The abrupt switching corresponds to *multiplying* a continuous time signal

with a rectangular window, which corresponds to *convoluting* the spectrum of the sweep by the rectangular window's spectrum (that is, the *sin(x)/x* function). A common way to circumvent this problem is to let the excitation sweep start well below the lowest frequency of interest. Synthesizing the excitation sweep in the spectral domain is a better solution, as will be shown in Sect. 3.2.

2.5 Noise Rejection

The frequency-dependent SNR of a non-windowed IR measurement depends on the spectral distribution and energy of both the background noise and the excitation signal during the data acquisition period, but not on the type (impulse, noise, sweep or other) of the excitation signal. However, differences exist in how the different types of interfering noise are distributed over the entire period of the estimated IR.

Monofrequent noise (such as hum) and uncorrelated broadband noise (for example, from air conditioning) will still appear as such in the IR, independently of the measurement method. This is due to the fact that their general character is not altered by manipulating the phase spectrum in the deconvolution process.

Short, impulsive noise events, such as clicks and pops, will be transformed into noise when using any noise stimulus. In contrast, they will be transformed into small time-inverted sweeps in a sweep measurement. Depending on the strength and time of occurrence, these time-reversed repetitions of the sweep can be masked by the main peak of the IR or become audible in its reverberant tail, giving it a bizarre melodic touch.

3 • Excitation Signals

Achieving a good signal-to-noise ratio (SNR) in impulse responses can be quite a challenge. In purely electrical environments, the measurement of the IR can be accomplished in direct form by injecting a pulse as shown in Figure 1. If the noise floor of the equipment under test is low, the SNR of the direct impulse response can be satisfactory. The situation is often different for acoustic ambiences, where the prevailing background noise, transducer non-linearities (loudspeaker distortion) and occasionally time variance (air movement) make the task difficult.

To improve the signal to background noise ratio, other excitation signals than the simple impulse are commonly used. The basic idea is to spread out the energy contained in a single impulse over a broader time interval.

Two fundamental types of excitation signals can be used: independent asynchronous signals (coming from analogue noise generators, for example) and custom-tailored deterministic signals which usually have the same length as the analysis time interval and are synchronously clocked with the acquisition of the DUT's output signal.

Synchronous deterministic signals can be conveniently created by defining their properties in the spectral domain and then getting the desired signal by IFFT. This is shown in Figure 10, where three basic types of excitation signals are created from exactly the same linear (white) magnitude spectrum. The difference of the three types lies in their phase spectrum.

For the impulse, all phases have to be set to 0°, which corresponds to an equal arrival time for all frequencies. For a pseudo-random noise signal, the phases must assume random values. More sophisticated algorithms exist to keep the crest factor of the noise signal low [6]. To create a white sweep (sometimes called *time stretched pulse*, TSP, see [7,8]), the group delay (which is proportional to the negative derivative of the phase) has to increase proportionally with the frequency.

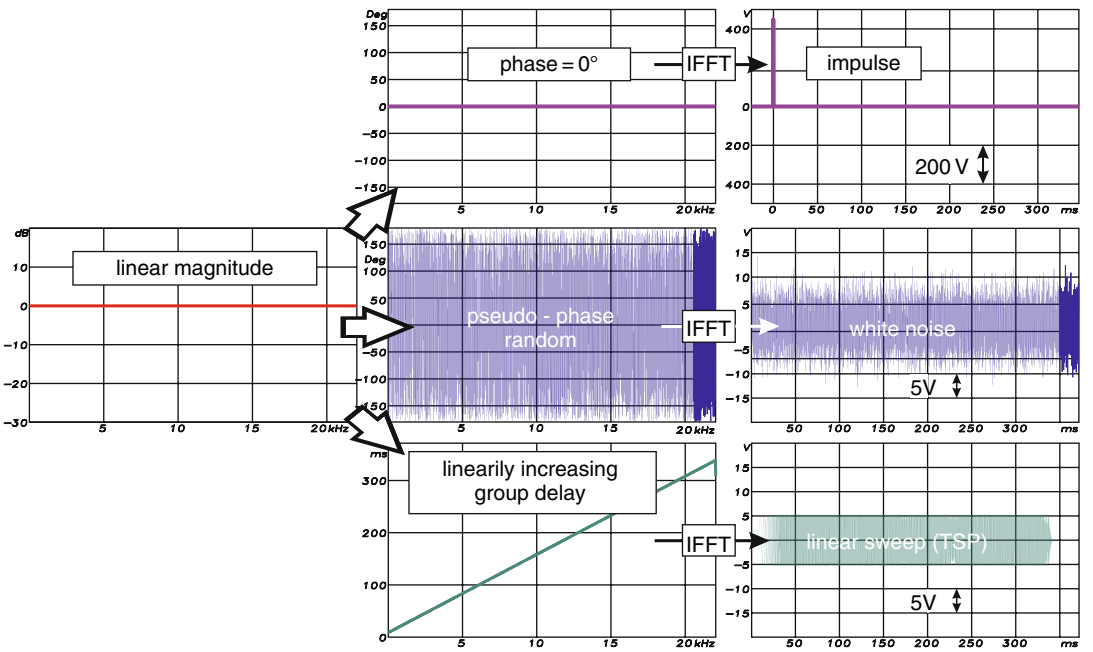


FIGURE 10 Three types of white deterministic excitation signals constructed in the frequency domain: impulse, pseudo-random noise and sweep.

In this example, the three excitation signals obtained after IFFT have identical energy. The sweep has the lowest peak value of all, about 6 dB lower than white noise. The impulse needs an amplitude of several hundred volts to concentrate the same energy as the sweep and noise signal, which shows that its use is unpractical.

3.1 Coloration of the Excitation Signal

In airborne and many other acoustical measurements, it is rarely useful to work with white excitation signals. The ambient noise floor almost always tends to increase considerably toward low frequencies. A white excitation signal often leads to insufficient SNR at the lower frequency bands [9,10]. In room acoustics, this may thwart the evaluation of parameters such as the reverberation time. A minimum SNR of at least 30 dB is desirable to obtain this value. For this reason, it is advisable to employ an excitation signal which strongly emphasizes the lower frequencies, for example by using pink (−3 dB/oct) or red (−6 dB/oct) coloration or a low-shelf filter, operating like the "bass" control in sound systems.

When a multi-way loudspeaker is used as sound source, the total energy fed to it can be increased considerably when working with pre-emphasis. The woofer often bears 10 times or more the power the tweeter can handle.

In many cases, a pink coloration is a reasonable choice to achieve sufficient SNR over the whole frequency range of interest. A low-shelf filter (see Figure 11) can be a better option in some cases. Above the upper corner frequency, its response gradually stops to decline and becomes linear. This can help overcome air absorption in long-distance open air measurements while maintaining substantial boost at low frequencies. This type of equalization can also be adequate if the background noise in a room is mainly caused by outside sources (such as traffic noise). In this case, the increased efficiency of the acoustical insulation of the walls can cause the acoustical noise level to drop below the microphone's or preamplifier's electrical noise level.

Additionally to the compensation of the background noise, the applied pre-emphasis can also include the loudspeaker's inverted frequency response. This is particularly useful when obtaining room IRs for auralization purposes. In this case, it is not acceptable that the loudspeaker's own response colors the results.

3.2 Generalized Sweep Synthesis

Sweeps used as excitation signal have a number of advantages presented earlier in this chapter. They can be created either directly in the time domain or indirectly in the frequency domain. In the latter case, their magnitude and group delay are synthesized and the sweep is obtained via IFFT of this artificial spectrum. This procedure avoids spectral ripple at the start and

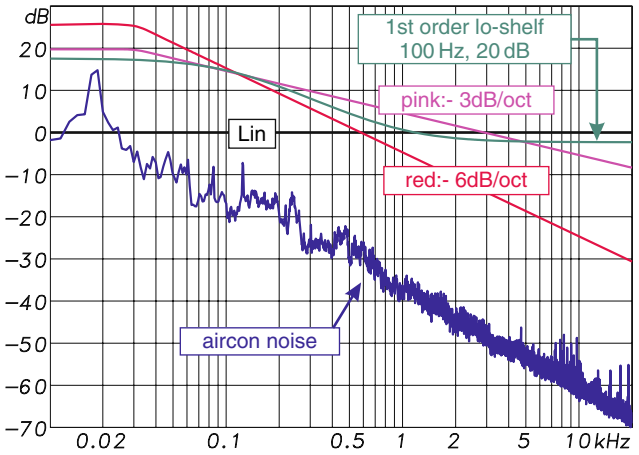


FIGURE 11 Typical background noise from an air condition unit (curve obtained from 30 energetic averages) and various pre-emphasis curves to overcome lacking SNR at low frequencies, compared to a white excitation signal of same total energy. To limit subsonic energy, the pink and red pre-emphasis curves have been chosen to not rise further below 30 Hz.

end frequencies which is mostly seen for sweeps built in the time domain. However, it can slightly worsen the crest factor, meaning that some cusps of the waveform can overshoot. Nevertheless, the generation in the frequency domain offers the crucial freedom to endow the sweep with a freely defined spectral distribution while almost perfectly maintaining its constant envelope (Figure 12).

This can be done by controlling the course of the group delay, which tells us the time at which each frequency is swept through. A steeply increasing group delay means that the corresponding frequency region is stretched out substantially in time, with the instantaneous frequency rising only slowly. This way, much energy is packed into this frequency region. On the other hand, an almost horizontal group delay course means that the corresponding frequency range is swept through in very short time, leaving little energy for it.

Thus, to construct a constant-envelope sweep with arbitrary spectral distribution, the group delay's inclination, or in other words, its derivative $\partial\tau_G(f)/\partial f$, has to be made proportional to the energy $|H(f)|^2$ at each frequency. For a discrete FFT spectrum, this means that the group delay τ_G has to increase according to

$$\tau_G(f) = \tau_G(f - \mathrm{d}f) + C \cdot |H(f)|^2, \tag{13}$$

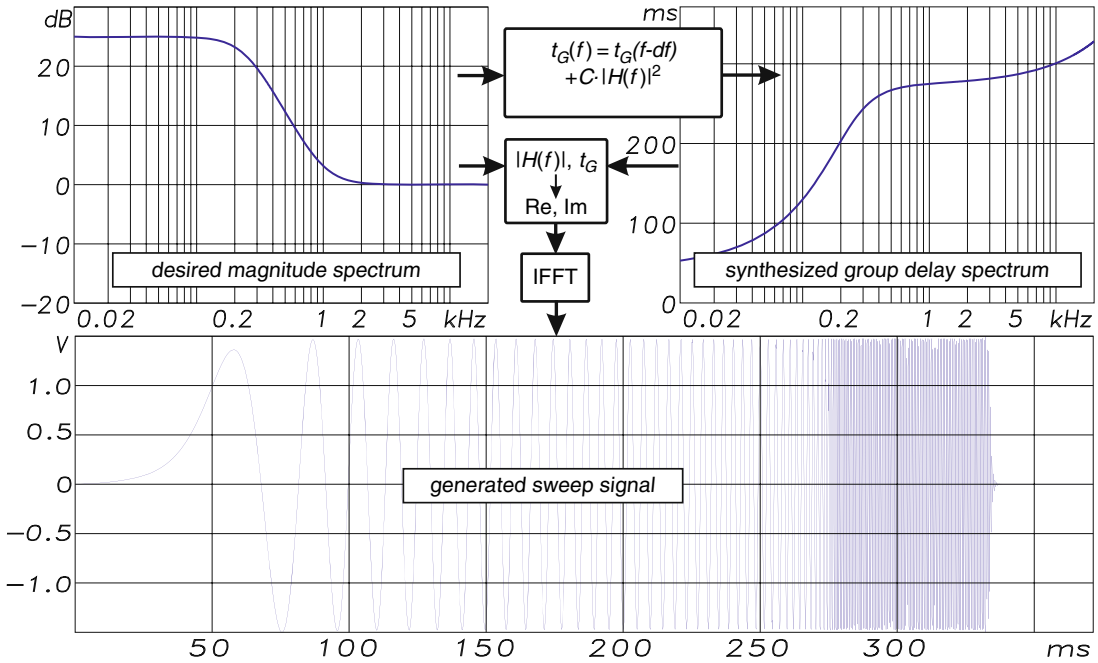


FIGURE 12 Generation of a sweep with constant envelope from an arbitrary magnitude spectrum.

with $\tau_G(f - df)$ being the predecessor (bin) of group delay $\tau_G(f)$ in the FFT spectrum, $|H(f)|^2$ the energy of the bin at frequency f, and C a normalization constant composed of the sweep's length divided by its total energy:

$$C = \frac{\tau_G(f_{END}) - \tau_G(f_{START})}{\sum_{f=0}^{f_S/2} |H(f)|^2}. \tag{14}$$

The sweep start time $\tau_G(f_{START})$ should be slightly above 0 s to allow the first half wave to evolve, the end time $\tau_G(f_{END})$ should be considerably smaller than the time interval $2^N/f_S$ (N = FFT order) to avoid smearing of the high-frequency tail into the sweep start and vice versa.

The phase is now easily calculated by integrating the group delay:

$$\varphi(f) = \varphi(f - df) - 2\pi \cdot df \cdot \tau_G(f) \quad \text{with} \quad df = f_S/2^N. \tag{15}$$

As a small detail, it is important that the phase resulting from the integration reaches exactly 0° or 180° at $f_S/2$. This condition

generally must be fulfilled for spectra of real-time signals. It can be achieved easily by subtracting values from the phase spectrum that decreases linearly with frequency until exactly offsetting the former phase at $f_S/2$:

$$\varphi_{NEW}(f) = \varphi_{OLD}(f) - \frac{f}{f_S/2} \cdot \varphi_{END_OLD}. \tag{16}$$

This is equivalent to adding a minor constant group delay in the range of ±0.5 samples over the whole frequency range.

An IFFT now reveals the synthesized sweep. Head and tail usually stretch out a little beyond the defined start and end time. Half-windows at the very beginning and end help to confine the sweep while unavoidably introducing a little deviation from the desired magnitude spectrum, although normally much smaller than the ripple of sweeps constructed by direct synthesis in the time domain.

The method can be easily extended to control the envelope in a frequency-dependent fashion instead of maintaining it constant. To do so, the desired magnitude spectrum is first *divided* by the *desired envelope spectrum* (which stipulates the frequency-dependent amplitude of the time signal). After synthesizing the group delay as previously, the desired magnitude spectrum is now *multiplied* with the desired envelope spectrum to re-establish the original spectral distribution. The final IFFT reveals a sweep which features both the desired spectral contents and the frequency-dependent amplitude. This generalized sweep generation method offers two crucial degrees of freedom: Any desired spectral distribution along with any freely definable frequency-dependent envelope will be transformed into a swept sine wave suitably warped in amplitude and time.

As a further sophistication, the sweep can also be split into various bands stored in separate channels to include the crossover functionality and equalization of a multi-way loudspeaker, this way replacing its passive or active crossover circuit. More details can be found in [4].

References

1. H. Herlufsen, “Dual Channel FFT Analysis (Part I, II)”, Brüel & Kjær Technical Review No. 1, 1984, www.bksv.com/pdf/Bv0013.pdf
2. D. D. Rife, J. Vanderkooy, “Transfer Function Measurement with Maximum-Length Sequences”, *J. Audio Eng. Soc.*, vol. 37, June 1989, pp. 419–444
3. A. Farina, “Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique”, *J. Audio Eng. Soc.*, vol. 48, 2000, p. 350, 108th AES Convention, Paris, Preprint 5093

4. S. Müller, P. Massarani, "Transfer-Function Measurements with Sweeps", *J. Audio Eng. Soc.*, vol. 49, June 2001, pp. 443–471
5. AES, R. C. Heyser, *Time Delay Spectrometry – An Anthology of the Works of Richard C. Heyser*, AES, New York, 1988
6. M. R. Schroeder, "Synthesis of Low-Peak-Factor Signals and Binary Sequences with Low Autocorrelation", *IEEE Trans. Inform. Theory,* 1970, pp. 85–89
7. N. Aoshima, "Computer-Generated Pulse Signal Applied for Sound Measurement", *J. Acoust. Soc. Am.*, vol. 69, May 1981, pp. 1484–1488
8. Y. Suzuki, F. Asano, H.-Y. Kim, T. Sone, "An Optimum Computer-Generated Pulse Signal Suitable for the Measurement of Very Long Impulse Responses", *J. Acoust. Soc. Am.*, vol. 97, February 1995, pp. 1119–1123
9. D. Griesinger, "Beyond MLS – Occupied Hall Measurement with FFT Techniques", *J. Audio Eng. Soc.*, vol. 44, p. 1174 (abstract), 101st AES Convention, Preprint 4403
10. E. Mommertz, S. Müller, "Measuring Impulse Responses with Preemphasized Pseudo Random Noise derived from Maximum Length Sequences", *Appl. Acoust.*, vol. 44, 1995, pp. 195–214

Further Reading

11. G.-B. Stan, J. J. Embrechts, D. Archambeau, "Comparison of Different Impulse Response Measurement Techniques", *J. Audio Eng. Soc.*, April 2002, pp. 249–262
12. J. Schoukens, R. Pintelon, E. van der Ouderna, J. Renneboog, "Survey of Excitation Signals for FFT Based Signal Analyzers", *IEEE Trans. Instrum. Meas.*, vol. 37, September 1988, pp. 342–352
13. J. Shoukens, R. Pintelon, "Measurement of Frequency Response Functions in Noise Environments", *IEEE Trans. Instrum. Meas.*, vol. 39, December 1990
14. A. J. Berkhout, D. de Vries, M. M. Boone, "A New Method to Acquire Impulse Responses in Concert Halls", *J. Acoust. Soc. Am.*, vol. 68, 1980, pp. 179–183
15. J. C. Burgess, "Chirp Design for Acoustical System Identification", *J. Acoust. Soc. Am.*, vol. 91, 1992, pp. 1525–1530
16. M. Poletti, "Linearly Swept Frequency Measurements, Time-Delay Spectrometry, and the Wigner Distribution", *J. Audio Eng. Soc.*, vol. 36, June 1988, pp. 457–468
17. A. Lundeby, T. E. Vigran, H. Bietz, M. Vorländer, "Uncertainties of Measurements in Room Acoustics", *Acustica*, vol. 81, 1995, pp. 344–353

6

Digital Sequences

Ning Xiang

School of Architecture and Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA, e-mail: xiangn@rpi.edu

1 • Introduction

This section discusses the applications of digital sequences in acoustical system identification and characterization and describes Golay codes and binary maximum-length sequences (MLSs) in some detail. Legendre sequences and other coded signals are briefly described. Golay codes and MLS have been used for acoustic applications for years. Applications of Legendre sequences have also been reported. Digital sequences of other classes such as, e.g., binary Gold sequences and Kasami sequences have only recently found applications in acoustical system identification and characterization.

2 • Golay Codes

Golay codes [1] have been used as acoustic excitation signals for determination of head-related transfer functions (HRTFs) of human subjects [2–4]. Golay codes $\{a_n, b_n\}$, a pair of complementary binary sequences of length $L = 2^n$, with integer $n > 1$, can be recursively constructed from a pair of complementary codes

$$\begin{cases} a_1 = \{+1 + 1\} \\ b_1 = \{+1 - 1\} \end{cases} \tag{1}$$

using the following recursion:

$$\begin{aligned}&\text{Append } b_{n-1} \text{ to } a_{n-1} \text{ to obtain } a_n\\&\text{Append } -b_{n-1} \text{ to } a_{n-1} \text{ to obtain } b_n.\end{aligned}$$

For example, appending b_1 to a_1 and appending $-b_1$ to a_1 yield $a_2 = \{+1+1+1-1\}$ and $b_2 = \{+1+1-1+1\}$. The most important property of this pair of codes is that the sum of the periodic autocorrelations of both codes is a two-valued function [3]:

$$\phi(k) = \sum_j \{a(j)a(k+j) + b(j)b(k+j)\} = \begin{cases} 2L \text{ for } k = 0 \\ 0 \quad \text{for } k \neq 0 \end{cases}, \qquad (2)$$

where $a(i)$ or $b(i)$ is the ith element of a_n or b_n, respectively. For periodic (circular) correlation all the indices are calculated modulo L. Figure 1 illustrates normalized periodic autocorrelation functions (PACF) of a pair of Golay codes of length 16. In this example, all functions in Figure 1 have to be normalized by 2×16 which leads to a peak value 0.5 of individual PACFs in Figure 1a and b, while the sum of them provides a perfect PACF of $\delta(i)$ as shown in Figure 1c. Since the complementary pair of Golay codes are of radix two in length, the fast Fourier transform (FFT) operations can be directly applied for cross-correlations in the frequency domain. So the sum of the PACFs of a pair of Golay codes in the frequency domain can be expressed as [3]

$$\Phi(f) = \text{FFT}(a_n)\text{FFT}^*(a_n) + \text{FFT}(b_n)\text{FFT}^*(b_n) = 2L, \qquad (3)$$

where FFT() and FFT*() stand for the fast Fourier Transform and its complex conjugates, respectively. In the context of acoustic transfer function measurements of a linear time-invariant (LTI) system under test as shown in Figure 2, $H(f)$ can be determined by exploiting the autocorrelation property expressed in Eq. (3) [3]:

$$H(f) = \frac{1}{2L}\{[\text{FFT}(a_n)H(f)]\text{FFT}^*(a_n) + [\text{FFT}(b_n)H(f)]\text{FFT}^*(b_n)\}. \qquad (4)$$

A measurement procedure for a system under test (SUT) can be implemented by generating a pair of Golay codes of sufficient length, longer than the expected system impulse response, to avoid time aliasing. The measurement procedure illustrated in Figure 2 is summarized as follows [4]:

- Excitation of the SUT periodically using a_n code of $k+1$ periods. After the SUT reaches its steady state, record the system response to a_n, conceivably with synchronous averaging over k periods.

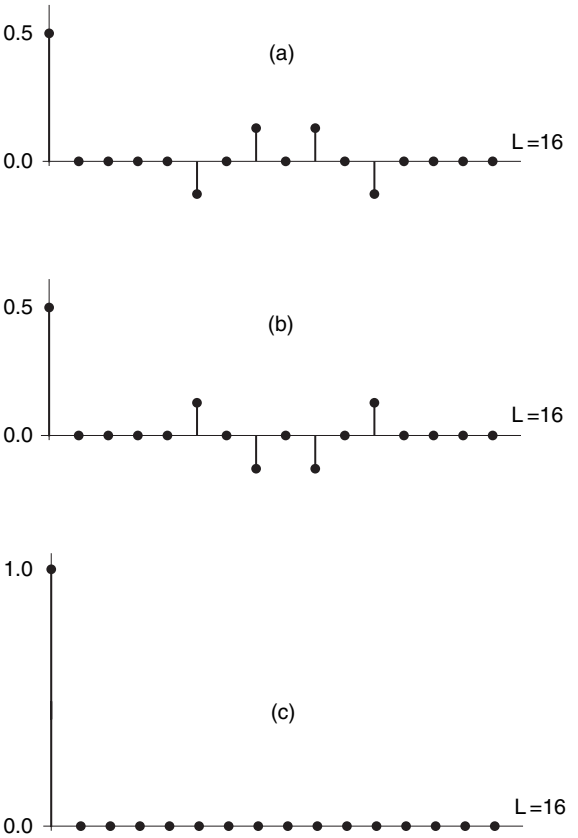


FIGURE 1 Normalized periodic autocorrelation functions of Golay codes of length $L = 16$ for the individual complementary a_4 and b_4 codes, shown in (**a**) and (**b**), respectively, along with the sum of both autocorrelation functions, shown in (**c**). The normalization factor of $2L$ leads to a peak value 0.5 of individual autocorrelation functions in (**a**) and (**b**), while the sum of them provides a perfect PACF of $\delta(i)$ as shown in (**c**) (from [4]).

- Wait until the system settles down, normally after one more period.
- Excitation of the SUT periodically using b_n code with $k+1$ periods. After the SUT reaches its steady state, record the system response to b_n, conceivably with synchronous averaging over k periods.
- Cross-correlation of one resulting period of response to a_n with the original code a_n in the frequency domain. In practice, its complex conjugate $\text{FFT}^*(a_n)$ is prepared in memory in advance for a direct multiplication as shown in Figure 2.

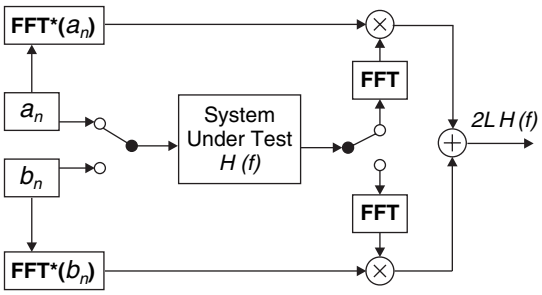


FIGURE 2 Flow diagram for determining the transfer function of a linear time-invariant system using a pair of Golay codes. FFT() and FFT*() stand for the fast Fourier Transform and its complex conjugates, respectively.

- Cross-correlation of one resulting period of response to b_n with the original code b_n in a similar fashion.
- Sum of the two correlation results and normalize it by $2L$, the transfer function of the SUT is determined.

Ease of generation using the recursive algorithm and the fact that FFT can be directly used for determining the transfer function of a SUT have been considered positive in the works by Zahorik [4]. Its shortcoming is that two sequences have to be used sequentially for acoustic excitation. The steady state of a LTI system has to be reached twice that makes any unavoidable time-variance effect more troublesome than those signals such as chirps [5], sweeps (see Chapter 5 by S. Müller), MLSs and other pseudorandom noise signals. The system characterization tasks, such as experimental determination of human HRTFs, are actually critical examples for the mentioned time variances [4].

3 • Binary Maximum-Length Sequences

Since the 1970s, binary maximum-length sequences (MLSs) have been introduced in acoustic applications [6] in shaping wall surfaces for diffuse sound reflections and in acoustic system identification, such as room impulse response measurements. This section briefs the acoustic system identifications using binary MLSs.

A n-stage linear feedback shift-register device (see Figure 3) can generate a binary periodic sequence $\{a_i\}$ with $a_i \in \{0, 1\}$, or equivalently $\{a_i\}$ can be derived by the linear recurrence [7]

$$a_i = \sum_{k=1}^{n} c_k a_{i-k}, \quad 0 \le i < L, \tag{5}$$

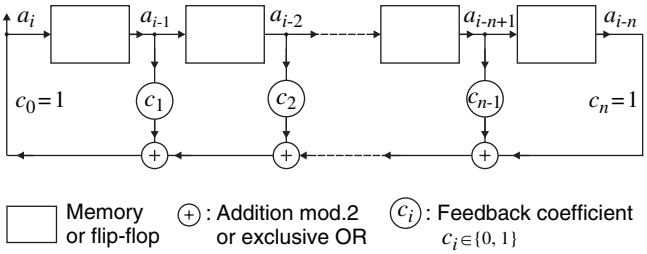


FIGURE 3 An n-stage linear shift register for generating binary periodic sequences.

with $c_k \in \{0, 1\}$ but $c_0 = c_n = 1$. The summation and indices are calculated modulo 2. Exclusive one all-zero state, there are $2^n - 1$ possible states for the shift register. When its feedback taps are appropriately connected, or equivalently an appropriate set of c_k is chosen, the periodic sequences arrive at their maximum period length of $L = 2^n - 1$. In this case, the sequences are referred to as maximum-length sequences (MLSs). A characteristic polynomial $f(x) = \sum_{k=0}^{n} c_k x^k$ expressing its feedback connection is then referred to as primitive. The positive integer n is said to be the degree of the MLS as well as its primitive polynomial (PP). A table of PPs for $n \leq 40$ can be found in [8]. In mathematical treatment MLSs are convenient in their binary form while a bipolar form $\{m_i\}$ is often used in practice to generate waveforms with $m_i = 1 - 2a_i \in \{-1, +1\}$.

MLSs enjoy a number of attractive properties that make them widely useful in broad scientific and engineering fields. They possess properties similar to those of random noise, but they are periodically deterministic and have a strict time structure within their periods. They are also termed pseudorandom sequence/noise. This section briefly reviews some basic properties pertaining to acoustic applications. Detailed description and definitions can be found in Refs. [6–11].

3.1 Auto-correlation Function

The most important property of binary MLS is that the normalized PACF of a bipolar MLS within one period is a two-valued function [12] with

$$\phi(i) = \frac{L+1}{L}\delta(i) - \frac{1}{L}. \tag{6}$$

Figure 4 illustrates a normalized PACF of MLS of $L = 15$. When the period length $L = 2^n - 1$ is large enough the PACF approximates the unit sample sequence

$$\phi(i) \approx \delta(i). \tag{7}$$

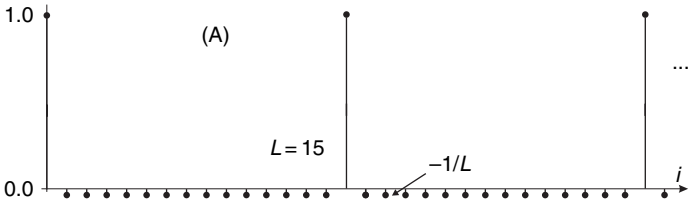


FIGURE 4 Normalized periodic autocorrelation function of an MLS of degree 4 ($L = 15$).

From its PACF, it is then readily deduced that the power spectrum of an MLS is completely flat, expect for a "dip" at the DC. One can also replace all -1s of a bipolar MLS $\{m_i\}$ by a value $-q$ with

$$q = \frac{1}{1 + 2/\sqrt{L+1}}, \tag{8}$$

so that the slightly modified MLS possesses a perfect PACF of $\delta(i)$ [11].

3.2 Shift-and-Add

A (modulo 2) sum of an MLS $\{a_i\}$ with its phase-shifted version $\{a_{i+\tau}\}$ results in the same MLS with another phase shift

$$\{a_i\} \oplus \{a_{i+\tau}\} = \{a_{i+g(\tau)}\}. \tag{9}$$

$g(\tau)$ is termed offset function, being a function of τ which can be explicitly determined in terms of Zech's logarithm [6]. The addition of modulo 2 becomes multiplication for bipolar MLS. The shift-and-add property leads directly to the proof of the two-valued autocorrelation function of bipolar MLSs (see Section 3.1). Along with Zech's logarithm it is also of practical significance for nonlinear system identification [13].

Generally a nonlinear system can be described by Volterra series

$$\begin{aligned} y(n) = & \sum_{i=0}^{\infty} h(i)m(n-i) \\ & + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} h(i,j)m(n-i)m(n-j) \\ & + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} h(i,j,k)m(n-i)m(n-j)m(n-k) \\ & + \cdots, \end{aligned} \tag{10}$$

where $h(i)$, $h(i,j)$ and $h(i,j,k)$ are first, second and third orders of Volterra kernels, respectively. For a linear system, in particular, second and all higher orders of Volterra kernels become zero. If the input signal of the nonlinear system $m(n)$ is an MLS, the periodic cross-correlation between $y(n)$ and $m(n)$ becomes

$$\begin{aligned}\phi(\tau) = &\sum_{i=0}^{L-1} h(i) \sum_{n=0}^{L-1} m(n-i)m(n+\tau) \\ &+ \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} h(i,j) \sum_{n=0}^{L-1} [m(n-i)m(n-j)]m(n+\tau) \\ &+ \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\sum_{k=0}^{L-1} h(i,j,k) \sum_{n=0}^{L-1} [m(n-i)m(n-j)m(n-k)]m(n+\tau) \\ &+ \cdots . \end{aligned} \tag{11}$$

According to the shift-and-add property, items in [.] of second and third line of Eq. (10) can be considered as phase-shifted MLS, while quantities following $\sum_{n=0}^{L-1}$ in each line represent shifted autocorrelation of the MLS as elaborated in Section 3.1. Equation (10) well explains "erroneous" behaviors in resulting impulse responses when the SUT contains weak nonlinearities, as reported via experimental observations [14,15]. For purpose of a linear system identification in practice, distinct spikes distributed along the impulse response are strong evidence and indication that the MLS excitation has already stretched beyond the linear range of the SUT. An amplitude reduction of the MLS excitation will often remove the erroneous nonlinear behaviors. Equation (10) also indicates an approximation scheme of nonlinear system identification relying on the determination of the offset function given the MLS. With some constrains, the higher order of Volterra kernels can be identified and separated from the linear response $h(n)$ of the SUT [13].

With respect to nonlinear immunity, inverse repeat sequences by appending an amplitude-inverse MLS to the original MLS can be used for improving even-order nonlinearity immunity [14, 16]. An overall high nonlinear immunity can be obtained using logarithmic sweep signals (see Chapter 5).

3.3 Characteristic MLS

An MLS $\{b_i\}$ can always be derived from a given one $\{a_i\}$ in terms of a circular phase shift such that $b_i = a_{i+\tau}$ so that an invariant decimation $b_i = b_{2i}$ can be satisfied. The MLS $\{b_i\}$ is designated as characteristic MLS and self-similar MLS [17], also termed idempotent MLS [6]. This invariant decimation holds only for

characteristic MLSs with decimation factors of 2^k with k being a positive integer [7]. There exists a unique characteristic MLS for a given PP and a unique initial state of the shift register to generate it. An algorithm of deriving the required initial states given the PP has recently been documented in acoustics journals [17]. This special form of MLS is of practical significance in the interleaved sampling technique for ultrasonic applications where current analog-to-digital converters undesirably limit the upper limit of the sampling frequency in specific applications [17,18]. Ultrasonic non-destructive testing and evaluation generally work in frequency range of tens or hundreds of MHz. Figure 5 illustrates the interleaved sampling scheme. A self-similar MLS of 2^k periods or more, generated at an appropriate clock rate, is applied to an ultrasonic SUT given a LTI system. After the system arrives at its steady state a clearly lower sampling rate derived from the master clock can drive the digitizer to receive one period of the system response to the MLS. In this way a high-resolution, low-cost device can conveniently be employed [18]. Since at least 2^k periods have to be used to derive one period of the system response, a highly stationary (time-invariant) measurement environment is crucial to the success of this application.

3.4 Preferred MLS Pairs

With some proper decimation factors d, decimation of $\{a_i\}$ by d may yield distinctly different MLSs $\{b_i\} = \{a_{di}\}$ associated with a distinctly different PP of the same degree. When the decimation factor is properly chosen, the decimation will lead to a pair of MLSs, whose periodic cross-correlation function (PCCF) is of clearly lower value than the peak of the PACF. This cross-correlation property is well-known in the spread spectrum communication systems, in particular, in CDMA technology [19]. Recent research in outdoor sound propagation for acoustic atmospherical tomography [20] calls for a critical measurement

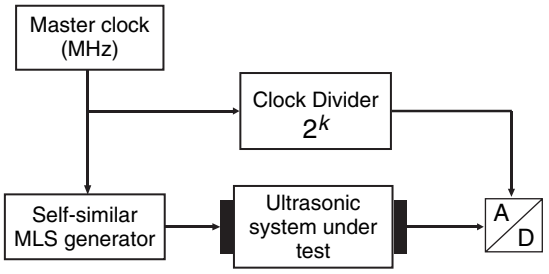


FIGURE 5 Characteristic MLS for ultrasonic applications. Interleaved sampling can take advantage of low sampling rate, high-resolution analog-to-digital converters [18].

technique for simultaneous sound speed determination using multiple sound sources through the atmosphere over the ground layer. The proper decimation of MLSs can make a significant contribution to the acoustic atmospherical tomography [21] (see Section 5).

If the degree of an MLS is a multiple of 4, then a decimation factor $d(n) = 2^{(n+2)/2} - 1$ will lead to an MLS pair whose PCCF takes on only four values. For such a decimation, the upper bound on the four-valued PCCF is

$$l_4(n) = \frac{2^{(n+2)/2} - 1}{2^n - 1}. \tag{12}$$

If the degree of an MLS is not a multiple of 4, then an appropriate decimation of that MLS will yield pairs of MLSs having relatively small three-valued PCCF [22]. For MLSs whose degrees are not a multiple of 4, some of decimation factors are of the form $d = 2^k + 1$ or $d = 2^{2k} - 2^k + 1$, where k is chosen such that $n/\gcd(n, k)$ is odd [22], with gcd(,) being the greatest common divider, the cross-correlation bounds take on preferred small values expressed by

$$l_3 = \frac{2^{\lfloor (n+2)/2 \rfloor} + 1}{2^n - 1}, \tag{13}$$

where $\lfloor x \rfloor$ denotes the integer part of the real number x. This section will refer to MLS pairs with both three-valued and four-valued PCCFs as preferred pairs. Figure 6 illustrates the PCCF of a preferred pair of degree 13 in comparison to the PACF of one of the MLSs. As one can see from the figure, the three-valued PCCF is bounded at a small value about 36 dB lower than the peak value of the PACF.

In addition to the preferred MLS pairs, the time-reversed sequence of every MLS and the MLS itself form a reciprocal pair. The reciprocal pair of every binary MLS of degree n also possesses small-valued PCCF [23] bounded by Eq. (12), yet not limited to three or four values.

Table 1 lists all bound values of preferred and reciprocal MLS pairs between degree 8 and 24 for an easy comparison. A careful comparison between columns in Table 1 reveals that the small-valued PCCF of the preferred MLS pairs are close to that of reciprocal ones for even-numbered degrees while for the odd-numbered degrees the preferred MLSs are approximately 1.4 times (3 dB) smaller than that of the reciprocal MLSs. Some acoustics applications require the bound values as small as possible, such as the acoustic atmospherical tomography [20]. When applying the cross-correlation technique in

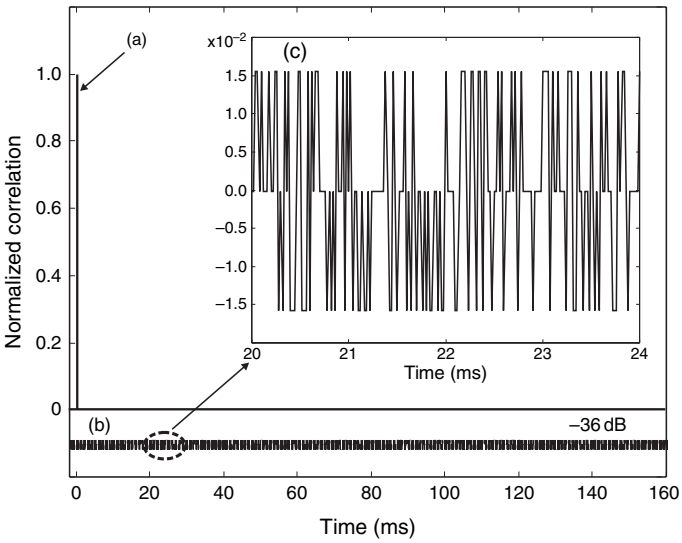


FIGURE 6 Normalized periodic correlation functions of a preferred MLS pair of degree 13 generated at sampling frequency of 50 kHz. (**a**) Autocorrelation function of an individual MLS. (**b**) Cross-correlation function between a preferred three-valued MLS pair (shifted downward beneath the autocorrelation function for a convenient comparison while keeping the same amplitude scale). The bound value of the cross-correlation amounts to 1.57×10^{-2}, 36 dB lower than the peak value of the autocorrelation. (**c**) Zoomed presentation of a segment from (**b**).

acoustic measurements using simultaneous sources driven by signals with above-mentioned correlation properties, individual channel impulse responses associated with each of individual sources can be straightforwardly resolved [21]. The smaller the bound values of PCCFs the less interference from other unwanted source channels.

While a pair of MLSs selected from two different preferred pairs of the same degree may not possess the preferred small-valued PCCFs as listed in Table 1, some pairs selected from different pairs do. A set of the MLSs for which each pair of the set has these preferred small PCCF values is referred to as connect set. The largest possible connected set is termed a maximum connect set (MCS) and the size of such a MCS is denoted by M_n. One can find the values of M_n for $n \leq 16$ with $M_n \leq 4$ [22]. The length of MLSs of these degrees and their reasonably low PCCF-bound values are of practical interest in a wide variety of acoustics measurements, such as in acoustic channel estimation using multiple simultaneous sources [21].

TABLE 1 Cross-correlation bound values of preferred and reciprocal MLS pairs, of Gold and Kasami sequences

Degree	Period length	Reciprocal MLS pairs	Preferred MLS pairs and Gold sequences	Kasami sequences
8	255	0.1216	0.1216	0.067
9	511	0.0881	0.0646	
10	1, 023	0.0616	0.0635	0.0322
11	2, 047	0.0437	0.0317	
12	4, 095	0.0310	0.0310	0.0159
13	8, 191	2.20E-2	1.57E-2	
14	16, 383	1.56E-2	1.56E-2	7.87E-3
15	32, 767	1.10E-2	7.84E-3	
16	65, 535	7.82E-3	7.82E-3	3.92E-3
17	131, 071	5.52E-3	3.91E-3	
18	262, 143	3.90E-3	3.90E-3	1.96E-3
19	524, 287	2.76E-3	1.95E-3	
20	1, 048, 575	1.95E-3	1.95*E*-3	9.77E-4
21	2, 097, 151	1.38E-3	9.77E-4	
22	4, 194, 303	9.76E-4	9.76E-4	4.89E-4
23	8, 388, 607	6.90E-4	4.88E-4	
24	16, 777, 215	4.88E-4	4.88E-4	2.44E-4

However, the number of MLSs in the MCS is strongly limited. One needs to use properly combined MLSs to form Gold or Kasami sequences (see Section 5) if specific applications require a large number of simultaneous sources.

4 • Fast MLS Transform (FMT)

For the purpose of experimental determination of impulse response of a linear time-invariant (LTI) system, we apply a periodic bipolar MLS $\{m_i\}$ to the LTI system under test. At the output of the system we receive one period or several periods of the system response to the MLS after the system arrives at its steady state. The receipt of several periods is for a synchronous averaging in order to further improve the signal-to-noise ratio of the measurement. Owing to the PACF of MLS in Eq. (7) the impulse response $h(i)$ can be determined in terms of cross-correlation between the excitation MLS and the system response to the MLS $\mathbf{Y}$ by

$$\mathbf{h} = \frac{1}{L+1}\mathbf{M}\mathbf{Y}, \tag{14}$$

where $\mathbf{h}$, $\mathbf{Y}$ are vectors of L elements. $\mathbf{M}$ represents the MLS matrix of dimension $L \times L$, its rows contain sequentially right-

cyclically shifted MLS. Equation (14) is historically termed MLS transform [24].

4.1 Hadamard Transform-Based FMT

The MLS matrix $\mathbf{M}$ in Eq. (14) is permutationally similar to Hadamard matrix [25]

$$\mathbf{M} = \mathbf{P_2}\mathbf{H_N}\mathbf{P_1}, \tag{15}$$

where $\mathbf{P_1}$, $\mathbf{P_2}$ denote the permutation and the repermutation matrix, respectively. $\mathbf{H_N}$ is the Sylvester-type Hadamard matrix, a square matrix of dimension N with $N = 2^n, n$ being a positive integer, its row vectors contain mutually orthogonal elements being either -1 or $+1$. $\mathbf{H_N}$ can be recursively constructed by

$$\mathbf{H_{2N}} = \begin{pmatrix} +\mathbf{H_N} & +\mathbf{H_N} \\ +\mathbf{H_N} & -\mathbf{H_N} \end{pmatrix} \quad \text{with} \quad \mathbf{H_1} = [1]. \tag{16}$$

For example,

$$\mathbf{H_2} = \begin{pmatrix} +1 & +1 \\ +1 & -1 \end{pmatrix}, \quad \mathbf{H_4} = \begin{pmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{pmatrix}. \tag{17}$$

This example indicates that the Hadamard transform (HT) consists of recursive butterfly structures, similar to the FFT, but requiring only addition operations. $\mathbf{H_2}$ makes its basic butterfly structure directly comprehensible. Figure 7 illustrates the fast Hadamard transform (FHT) of degree 3. The computation efficiency is $O(N \log_2 N)$ of add/subtraction.

Equation (15) implies a fast algorithm referred to as fast MLS transform (FMT) [24], since the FHT for Hadamard matrix $\mathbf{H_N}$ is adopted in the calculation. Figure 7 illustrates a flow diagram for the HT-based FMT of degree 3. It consists of three major steps:

- Permutation of the SUT response to the MLS being used as excitation ($\mathbf{P_1 Y}$).
- Fast Hadamard transform of the permuted vector ($\mathbf{H_N}(\mathbf{P_1 Y})$).
- Re-permutation of the transformed vector $\mathbf{P_2}[\mathbf{H_N}(\mathbf{P_1 Y})]$.

As illustrated in Figure 7, an impulse response in the time domain is directly obtained as a result right after the re-permutation except for a scale factor.

The Hadamard transform-based fast MLS transform generally requires two permutation matrices as indicated in Eq. (15). They contain only one non-zero element of value "1" in each row or column of the matrices. Instead of the matrix, the indices

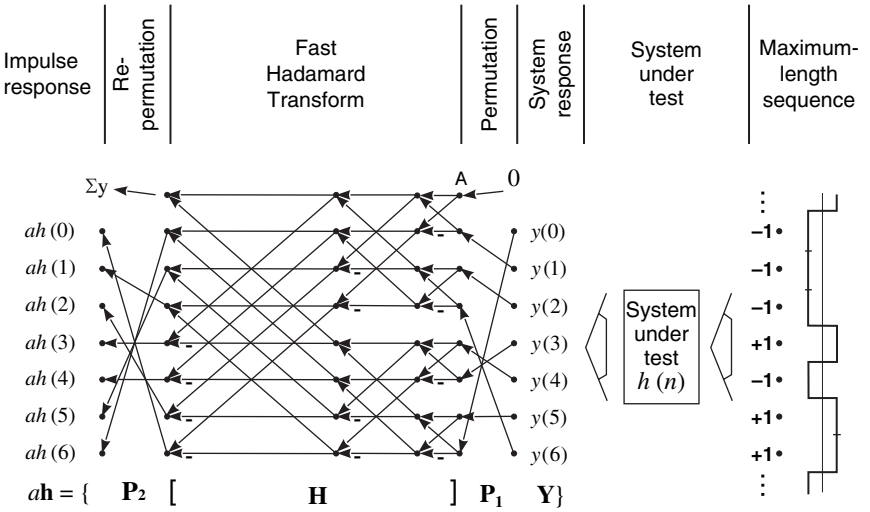


FIGURE 7 Flow diagram of the Hadamard transform-based fast MLS transform with $n = 3$, $L = 2^3 - 1$ and a primitive polynomial $f(x) = 1 + x + x^3$. Apart from a scale factor $a = (L+1)$, the impulse response $h(i)$ of LTI system can be efficiently derived in time domain from the system response $y(i)$.

of non-zero elements from each row or column are usually stored [23]. Since the fast MLS transform performs inherently the cross-correlation between the MLS itself and the system response to the MLS, the two permutation matrices together can be considered as the original binary MLS in index form. There exist several different determination approaches to permutation matrices, some of them provide two required permutation matrices [12,24,26]. One of the algorithms derived from the characteristic MLS proved to require only single permutation matrix. Recently this algorithm has been extended for the single permutation matrix to work at the same time on a pair of reciprocal MLSs for the simultaneous dual source measurements in acoustic applications [23].

4.2 FFT-Based FMT

Due to odd length $L = 2^n - 1$ of MLS, the periodic cross-correlation expressed in Eq. (14) cannot be carried out simply using the fast Fourier transform (FFT) of 2^n points, since a one-point zero padding will dramatically destroy the nature of MLS (see also Chapter 5 by Müller). The augmented MLS by one zero or any value will totally destroy the PACF as shown in Figure 4, therefore totally destroy the flatness of the power spectra [27].

Recently, a dedicated algorithm has been reported for the FFT to work on the cross-correlation of coded signals not having the required length of 2^n [28]. Using this method the FMT can be performed economically with a computational efficiency of $O[2(n+1)2^{(n+1)}]$ with the following three major steps:

- Augmentation of the MLS $\{m_i\}$ of degree n to $\{\hat{m}_i\}$ and the response to the MLS $\{y_i\}$ to $\{\hat{y}_i\}$ by 2^n+1 zeros, respectively.
- Cross-correlation between two augmented sequences of length $M=2^{(n+1)}$ using the FFT_M (M explicitly indicates $2^{(n+1)}$ points)

$$\hat{h}_i = \text{IFFT}_M\{\text{FFT}_M(\hat{m}_i)\text{FFT}_M^*(\hat{y}_i)\}. \tag{18}$$

- Addition of the first $L=2^n-1$ and the last L points in $\hat{h}_i$ to form the resulting h_i of L points

$$h_i = \hat{h}_i + \hat{h}_{i+M-L}, \quad 0 \le i < L. \tag{19}$$

The algorithm summarized above is straightforwardly generalized to any periodic-coded signals whose period length is not a required length of p^n for FFT [28] with p being a prime, such as inverse repeat sequences, Legendre sequences as described in Section 6. The FFT-based FMT algorithm is computationally more expensive than the HT-based FMT, $O[2(n+1)2^{(n+1)}]$ multiplications and additions vs. $O[n2^n]$ additions [28]. However, the additional computation load will be less and less noticeable with the modern, fast developing computer technology. On the other hand, the FFT-based FMT is much easier to be implemented than the HT-based FMT. The FFT-based FMT is of practical significance when digital sequences other than MLSs have to be applied in acoustical experiments as the next section elaborates.

5 • Gold and Kasami Sequences

A binary Gold sequence $\{G_\tau(i)\}$ can be generated by combining a preferred or a reciprocal binary MLS pair $[\{a_i\},\{b_i\}]$ as $G_\tau(i)=a_i\oplus b_{i+\tau}$, with $\oplus$ denoting addition modulo 2 [22]. In stepping τ point-by-point, a large number of Gold sequences result including the preferred MLS pair to yield a finite set $\{\{a\},\{b\},\{G_0\},\{G_1\},\ldots\}$ containing exactly $L+2$ sequences. Sequences in this finite set possess preferred PCCF properties. PACFs of bipolar Gold sequences associated with $\{G_\tau\}$ are peaked at zero lag, but they also possess small-valued side lobes. The PCCF of Gold sequences possesses the same bound value as that in the side lobes of the corresponding PACF. In fact, the bound values are the same as those of PCCF of preferred and reciprocal MLS pairs from which the Gold sequences are derived.

For even-numbered degree n, a decimation from an MLS $\{a_i\}$ with factor $\lambda(n) = 2^{n/2} + 1$ can lead to a pair $\{a_i\}$ and $\{e_i\}$ with $e_i = a_{\lambda i}$. In this case, $\{e_i\}$ is not an MLS of degree n. A binary Kasami sequence $\{K_\tau(i)\}$ can be generated by combining $\{a_i\}$ and $\{e_i\}$ as $K_\tau(i) = a_i \oplus e_{i+\tau}$. In similar fashion, a set of Kasami sequences $\{\{a\}, \{K_0\}, \{K_1\}, \ldots\}$can be constructed [22] containing $2^{n/2}$ sequences. The side lobes of PACFs and the amplitude of PCCFs among these sequences present even lower bounds, approximately half of those of preferred pairs and Gold sequences:

$$l_k(n) = \frac{\lambda(n)}{2^n - 1}, \tag{20}$$

where n must be a positive even number. Table 1 also lists the bound values of Kasami sequences along with those of preferred and reciprocal MLS pairs for comparison. The power spectral density functions of Gold and Kasami sequences are the same as that of MLSs, being of broadband nature and covering the entire frequency range. In order to demonstrate the excellent correlation properties of Gold and Kasami sequences, Figure 8 shows

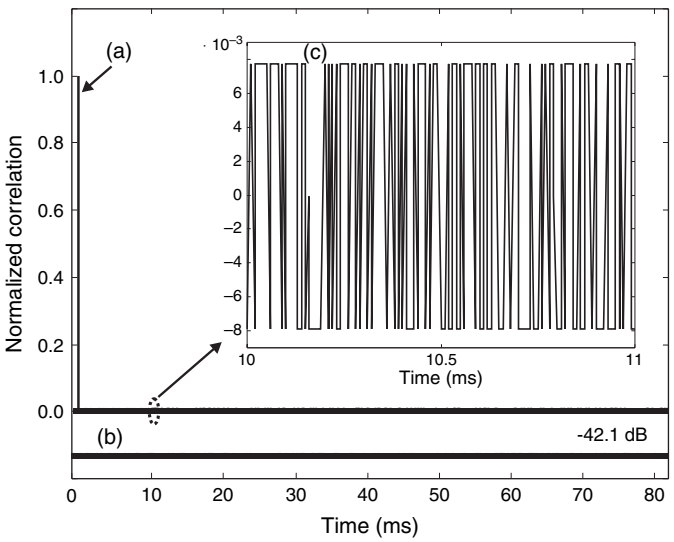


FIGURE 8 Normalized correlation functions of Kasami sequences of degree 14 generated at the sampling frequency of 200 kHz. (**a**) Periodic autocorrelation function (PACF) of one sequence. (**b**) Periodic cross-correlation function (PCCF) between two Kasami sequences (shifted downward beneath the autocorrelation function for a convenient comparison). The bound value of the PCCF is 42.1 dB lower than that of PACF. (**c**) Zoomed presentation of a segment from (**a**) showing three-valued side lobes of the PACF. Their peak values are the same as those of the PCCF in (**b**).

the PACF and PCCF of a pair of Kasami sequences, derived from an MLS of degree 14 decimated by a factor 129. In the example, the PCCF values and the side lobes of the PACF are bounded by 0.00787 or 42.1 dB below the peak value of the individual PACF. Unlike MLSs, the PACFs of Gold and Kasami sequences are pulse-like functions with small-valued side lobes. They are not MLSs any more by definition, but equivalent to the MLSs drawn from MCSs with regard to their usefulness in the simultaneous multiple acoustic source measurement technique [21], particularly when a large number of simultaneous acoustic sources have to be used. Kasami sequences have even lower bound values than that of Gold sequences, however, Kasami sequences are only available for even-numbered degrees. Moreover, the correlation measurement technique using these combined MLS sequences could not process the cross-correlation efficiently; only very recently, an efficient deconvolution algorithm [28] as described in Section 4.2 has become available.

6 • Legendre Sequences

Legendre sequences have been introduced as acoustic excitation signals for measurements of room impulse responses [9] since the late 1970s. Other applications have also been reported, such as measurements of brain-stem auditory-evoked responses (BAERs) [10].

Legendre sequences (LGSs) are periodic pseudorandom signals of period length L, where L is a prime of the form $4m-1$, with m being a positive integer. LGSs $\{a_i\}$ are defined as follows: $a_i = 1$, if i is a quadratic residue modulo L; $a_i = -1$, if i is a quadratic non-residue modulo L [9]. a_0 can be chosen arbitrarily as 0, 1 or -1. LGSs $\{a_i\}$ can also be generated by

$$a_i = i^{(L-1)/2}, \quad 0 < i < L, \tag{21}$$

where the result is taken as the least absolute remainder modulo L. LGSs are bipolar (binary) sequences for $|a_0| = 1$, while they are actually ternary sequences for $a_0 = 0$.

The PACF of LGSs normalized by its period length L is also a two-valued function

$$\phi(i) = \frac{L + |a_0|}{L}\delta(i) - \frac{1}{L}, \tag{22}$$

like that of MLSs as shown in Figure 4. The power spectrum of LGSs is therefore flat. It is this autocorrelation property that makes LGSs also useful for the correlation measurement

technique. There exists a fast prime-length Fourier transform [29] so that the PCCFs can be efficiently processed. The algorithm described in Section 4.2 is also suitable for Legendre sequences for an efficient correlation processing.

Every LGS remains decimation invariant by a quadratic residue. Taking an example of $L = 7$, the LGS

$$a_i = 0, 1, 1, -1, 1, -1, -1; 0, \ldots. \tag{23}$$

A decimation by 2, being a quadratic residue given $L = 7$, results in,

$$a_{2i} = 0, 1, 1, -1, 1, -1, -1; 0, \ldots, \tag{24}$$

which exactly equals a_i. The invariant decimation can be exploited for interleaved sampling (in Section 3.3), a technique being applied in ultrasonic measurement techniques [18].

7 • Application Remarks

LGSs possess a number of intriguing properties, such as two-valued PACF, flat power spectra, pseudorandomness and invariant decimation, similar to MLSs. In addition, there are more period lengths of LGSs in the form of prime $4k - 1$ to choose from in practice than those of MLS-related sequences or Golay codes.

For LTI system identifications, both LGSs and MLS-related signals have a signal-to-noise ratio improvement proportional to their period length, since they possess a low peak factor, or high crest factor when compared to periodic impulsive excitations. In practice, these pseudorandom signals have to pass low/bandpass anti-aliasing filters, resulting in a crest factor of 10–11 dB [30,31] or lower depending on the filter characteristics. They can be used in noisy environments or at low excitation level when, e.g., the dynamic range of the SUT is severely limited or there are practical reasons for wanting a low excitation level [9], but with sufficient averaging the results can still be of high quality.

Due to excellent number-theory properties of MLSs, the MLS measurement technology has received wide acceptance in acoustic applications for over two decades. Relatively high noise immunity and an efficient correlation algorithm have significantly improved room impulse response measurements [9], acoustic impedance-tube measurements [32] and measurements of absorption coefficients in reverberation chamber even with a rotating diffuser [33]. In situ measurements of sound insolation of noise barriers [34], in situ measurements of acoustic surface

reflection functions [35,36] both indoors and outdoors, measurements of head-related transfer functions [37] and hearing aids [31], applications in underwater acoustics [27] and acoustic technology of landmine detection [38] have been documented. Commercially available measurement systems based on MLS have already been serving audio engineers, acoustics researchers and consultants for quite some years. They are finding even more applications in architectural acoustics, audio engineering, physical acoustics, underwater acoustics and outdoor sound propagation.

The interest in MLS measurement technology has also stimulated investigations on its advantages, such as high-transient noise immunity [39]. However, pseudorandom signals in general, and MLSs in particular, are sensitive to time-variance components in the SUT [15,40] leading to low time-variance immunity. While some other methods with associated signals have been emerging for acoustical system identifications and characterizations with certain competing features (see Chapter 5 by Müller), such as even higher time-variance and nonlinear immunity, MLS-related signals and other digital sequences with their distinct advantages will still retain their practical significance in a wide variety of applications, particularly in nonlinearity detection/identification, and in challenging tasks using simultaneous multiple sources.

References

1. Golay, M. J. E.: "Complementary series", *IRE Trans. Inform. Theory* **7** (1961), pp. 82–87.
2. Foster, S.: "Impulse response measurement using Golay codes", *IEEE 1986, Conference on Acoustics, Speech, and Signal Processing* **2** (IEEE New York, 1986), pp. 929–932.
3. Zhou, B., Green D. M. and Middlebrooks J. C.: "Characterization of external ear impulse responses using Golay codes", *J. Acoust. Soc. Am.* **92** (1992), pp. 1169–1171.
4. Zahorik, P.: "Limitations in using Golay codes for head-related transfer function measurement", *J. Acoust. Soc. Am.* **107** (2000), pp. 1793–1796.
5. Burgess, J.: "Chirp design for acoustical system identification", *J. Acoust. Soc. Am.* **91** (1992), pp. 1525–1530.
6. Schroeder, M. R.: *Number Theory in Science and Communication*, 2nd enl. ed. Springer-Verlag, Berlin (1991).
7. Colomb, S. W.: *Shift Register Sequences*, Aegean Park Press, Laguna Hills, CA (1982).
8. MacWilliams, F. J. and Sloane, N. J.: "Pseudo-random Sequences and Arrays", *Proc. IEEE*, **64** (1976), pp. 1715–1729.

9. Schroeder, M. R.: "Integrated-impulse method measuring sound decay without using impulses", *J. Acoust. Soc. Am.* **66** (1979), pp. 497–500.
10. Burkard, R., Shi, Y. and Hecox K. E.: "A comparison of maximum length and Legendre sequences for the derivation of brain-stem auditory-evoked responses at rapid rates of stimulation", *J. Acoust. Soc. Am.* **87** (1990), pp. 1656–1664.
11. Lüke, H. D.: *Korrelationssignale*, Springer-Verlag, Berlin, New York (1992).
12. Borish, J. and Angell, J. B.: "An efficient algorithm for measuring the impulse response using pseudorandom noise", *J. Audio Eng. Soc.* **31** (1983), pp. 478–488.
13. Shi, Y. and Hecox, K. E.: "Nonlinear system identification by m-pulse sequences: application to brainstem auditory evoked responses", *IEEE Trans. Biomed Eng.*, **38** (1988) pp. 834–845.
14. Dunn, Ch. and Hawksford, M. O.: "Distortion immunity of MLS-derived impulse response measurements", *J. Audio Eng. Soc.* **41** (1993), pp. 314–335.
15. Vanderkooy, J.: "Aspects of MLS measuring systems", *J. Audio Eng. Soc.* **42** (1994), pp. 219–231.
16. Ream, N.: "Nonlinear identification using inverse repeat sequences", *J. Acoust. Soc. Am.* **76** (1984), pp. 475–478.
17. Xiang, N. and Genuit, K.: "Characteristic maximum-length sequences for the interleaved sampling method", *ACUSTICA* **82** (1996), pp. 905–907.
18. Mommertz, E. and Bayer, G.: "PC-based high frequency range M-sequence measurements using an interleaved sampling method", *ACUSTICA*, **81** (1995), pp. 80–83.
19. Simon, M. K., Omura, J. K., Scholtz, R. A. and Levitt, B. K.: *Spread Spectrum Communications Handbook*, McGraw-Hill (1994).
20. Wilson, D. K., Ziemann, A., Ostashev, V. E. and Voronovich, A. G.: "An overview of acoustic travel-time tomography in the atmosphere and its potential applications", *ACUSTICA* **87** (2001), pp. 721–730.
21. Xiang, N., Daigle, J. N. and Kleiner, M.: "Simultaneous acoustic channel measurement via maximal-length-related sequences", *J. Acoust. Soc. Am.* **117** (2005), pp. 1889–1894.
22. Sarwate, D. V. and Pursley, M. B.: "Cross-correlation properties of pseudorandom and related sequences", *Proc. IEEE* **68** (1980), pp. 593–619.
23. Xiang, N. and Schroeder, M. R.: "Reciprocal maximum-length sequence pairs for acoustical dual source measurements", *J. Acoust. Soc. Am.*, **113** (2003), pp. 2754–2761.
24. Cohn, M. and Lempel, A.: "On fast M-sequences transforms", *IEEE Trans. Inform. Theory* **23** (1977), pp. 135–137.
25. Lempel, A.: "Hadamard and M-sequences transforms are permutationally similar", *Appl. Optics*, **19** (1979), pp. 4064–4065.
26. Sutter, E. E.: "The fast m-transform: a fast computation of cross-correlations with binary m-sequences", *SIAM J. Comput.* **20** (1991), pp. 686–694.
27. Birdsall, T. G. and Metzger Jr., K.: "Factor inverse matched filtering", *J. Acoust. Soc. Am.* **79** (1986), pp. 91–99.

28. Daigle, J. N. and Xiang, N.: "A specialized fast cross-correlation for acoustical measurements using coded sequences", *J. Acoust. Soc. Am.* **119** (2006), pp. 330–335.
29. Rader, C. M.: "Discrete Fourier transforms when the number of data samples is prime", *Proc. IEEE* **56** (1999), pp. 1107–1108.
30. Rife, D. D. and Vanderkooy, J.: "Transfer-function measurement with maximum-length sequences", *J. Audio Eng. Soc.* **37** (1989), pp. 419–444.
31. Schneider, T. and Jamieson D. G.: "A dual-channel MLS-based test system for hearing-aid characterization", *J. Audio Eng. Soc.* **41** (1993), pp. 583–594.
32. Chu, W. T.: "Architectural acoustic measurements using periodic pseudo-random sequences and FFT", *J. Acoust. Soc. Am.* **76** (1984), pp. 475–478.
33. Chu, W. T.: "Room response measurements in a reverberation chamber containing a rotating diffuser", *J. Acoust. Soc. Am.* **77** (1985), pp. 1252–1256.
34. Garai, M. and Guidorzi P.: "European methodology for testing the airborne sound insulation characteristics of noise barriers in situ: experimental verification and comparison with laboratory data", *J. Acoust. Soc. Am.* **108** (2000), pp. 1054–1067.
35. Mommertz, E.: "Angle-dependent in-situ measurements of reflection coefficients using a subtraction technique", *Appl. Acoust.* **46** (1995), pp. 251–263.
36. Li, J. F. and Hodgson M.: "Use of pseudo-random sequences and a single microphone to measure surface impedance at oblique incidence", *J. Acoust. Soc. Am.*, **102** (1997), pp. 2200–2210.
37. Schmitz, A. and Vorländer, M.: "Messung von Aussenohrstossantworten mit Maximalfolgen-Hadamard-Transformation und deren Anwendung bei Inversionversuchen", *ACUSTICA*, **71** (1990), pp. 257–278.
38. Xiang, N. and Sabatier J. M.: "Laser-Doppler vibrometer-based acoustic landmine detection using the fast M-sequence transform", *IEEE Trans. Geosci. Remote Sens.* **1** (2004), pp. 292–294.
39. Ciric, D. G. and Milosevic M. A.: Transient noise influence in MLS measurement of room impulse response, *ACUSTICA*, **91** (2005), pp. 110–120.
40. Svensson, U. P. and Nielsen, J. L.: "Errors in MLS measurements caused by time variance in acoustic systems", *J. Audio Eng. Soc.* **47** (1999), pp. 907–927.

7

Filters

Ole-Herman Bjor

Norsonic AS, Lierskogen, Norway, e-mail: ohbjor@norsonic.com

1 • Introduction

Filters and frequency weightings are used to give emphasis or de-emphasis to parts of a signal based on its frequency content. The frequency range for filters is often separated in one or more passbands and one and more stopbands. The stopband consists of the band of those frequencies where we want the filter to remove the signal energy (high attenuation); the passband consists of the band of those frequencies where we want a low attenuation. An example is an octave-band filter where the passband is one octave wide. The filter will have low attenuation at frequencies within the octave band and high attenuation for signals outside the band.

The term "frequency weighting" is used when we want a more gradual emphasis/de-emphasis. An example is the A-weighting used in sound level meters introduced to mimic the variation in the sensitivity for the human auditory organ to sound with different frequencies.

The concept of frequency is strongly related to sinusoidal signal. Signals with a sinusoidal shape have a very special position in signal analysis. The reason is that this signal is the only signal that always retains its shape through a linear system. If a sinusoidal signal is used as an input to a linear system, the output will always be a signal with the same shape after the transient part due to the onset of the signal has decayed. The amplitude and phase may change, but not the sinusoidal shape and the frequency. This makes it easy to calculate the final response when

a sinusoidal signal passes a sequence of linear elements. The link between time domain and frequency domain, as given by the Fourier transformation, gives one of more good reasons why we want to describe the signal by its frequency content.

Parseval's theorem states that [1]

$$\int_{-\infty}^{\infty} x^2(t)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(\omega)|^2\,\mathrm{d}\omega, \tag{1}$$

where $x(t)$ is a signal as a function of time (assumed time limited) and $X(\omega)$ its Fourier transform, and $\omega = 2\pi f$ is the angular frequency. The theorem tells that the time integral of the squared signal may be obtained by integrating the squared magnitude of its Fourier-transformed function along the frequency axis. The integral on the right-hand side may be written as the sum of contiguous frequency bands:

$$\begin{aligned}\int_{-\infty}^{\infty} x^2(t)\mathrm{d}t &= \frac{1}{2\pi}\sum_n \left\{ \int_{-\omega_{n+1}}^{-\omega_n} |X(\omega)|^2\,\mathrm{d}\omega + \int_{\omega_n}^{\omega_{n+1}} |X(\omega)|^2\,\mathrm{d}\omega \right\} \\ &= \sum_n \int_{-\infty}^{\infty} x_{w_n}^2(t)\mathrm{d}t,\end{aligned} \tag{2}$$

where the last equation is done by inverse Fourier transformation, and

$$x_{w_n}^2(t)$$

is the signal $x(t)$ filtered, so that only the frequencies between ω_n and ω_{n+1} are present. The equation tells that the total energy can be found as the sum of the energy in the different bands. It also tells that the phase response of the weighting is of no concern when we want to calculate the signal strength (energy) since the equation contains only the magnitude of the Fourier transformation.

2 • Frequency Weightings

2.1 Weightings for Sound Measurements

The human response to a continuous sound depends on the frequency. Traditionally, sound in the frequency range 20 Hz–20 kHz is regarded as audible. However, in reality there is a gradual change between audible and non-audible sound. A tone with a high or a low frequency outside this range may be audible if it has a sufficient strength, and a tone with a frequency inside this range may be inaudible if the strength is too low [2,3].

A network with a frequency-dependent gain may be used to mimic the frequency response of the human auditory organ. The most common network is the A-weighting. Historically, there were three common weightings for sound level measurements: A, B and C, to mimic the auditory response at low, medium and high sound levels. Today, we have better methods for describing the auditory response, and only A- and C-weightings are in general use. The A-weighting is mainly applied for general sound level measurement and to access the risk for hearing impairment due to loud noise. C-weighting is mainly used to access the risk of hearing impairment due to short time/high amplitude or peak values in the sound. Since the C-weighting is relatively flat within the normal audible frequency range, it is often used for indicating the non-weighted response. Recently, the Z-network has been introduced for this purpose.

The International standard for sound level meters IEC 61672-1 defines three frequency weightings for the measurement of sound level: A, C and Z [4]. The standard specifies the frequency response of the weightings as well as the applicable tolerances for two accuracy classes: Class 1 and Class 2. The design goal is similar for both classes, but Class 1 has the lowest tolerances.

A sound level meter consists in general of a sensor for the sound (microphone), a network for frequency weighting and a detector with an indicating device for measuring the strength of the sound. The detector shall indicate the square root of the time-averaged squared value of the frequency-weighted signal, abbreviated root mean square or rms value. As can be seen from the Parseval's theorem, the phase response of the network will not influence the measured rms value. However, other indicators of the signal strength like the peak value will be dependent on the phase.

The A-weighting is given by the following magnitude function:

$$w_{\mathrm{A}}(f) = K_{\mathrm{A}} \frac{f/f_1{}^2}{1+(f/f_1)^2} \frac{f/f_2}{\sqrt{1+(f/f_2)^2}} \frac{f/f_3}{\sqrt{1+(f/f_3)^2}} \frac{1}{1+(f/f_4)^2}, \quad (3)$$

where K_{A}, f_1, f_2, f_3 and f_4 are given by the following approximate values:

$$K_{\mathrm{A}} = 1.258905$$
$$f_1 = 20.60\mathrm{Hz}$$
$$f_2 = 107.7\mathrm{Hz}$$
$$f_3 = 737.9\mathrm{Hz}$$
$$f_4 = 12194\mathrm{Hz}$$

The C-weighting is defined by the equation

$$w_C(f) = K_C \frac{(f/f_1)^2}{1+(f/f_1)^2} \frac{1}{1+(f/f_4)^2}, \tag{4}$$

where f_1 and f_4 are as defined above and

$$K_C = 1.007152$$

Both A- and C-weightings are defined to have unity gain at 1 kHz.

The Z-weighting is a weighting with the same tolerances as A- and C-weighting, but with a desired magnitude equal to one:

$$w_Z(f) = 1. \tag{5}$$

The corresponding gain/attenuation in decibel may be obtained by

$$L_{weight}(f) = 20 \log(w(f)), \tag{6}$$

where $w(f)$ are any of the weightings above (Table 1).

The responses at different frequencies are shown in Figure 1, plotted as the gain in decibel versus frequency in logarithmic scale.

TABLE 1 A-, C- and Z-weighting

Nominal frequency (Hz)	Frequency weightings (L_{weight}) (dB)		
	A	C	Z
10	−70.4	−14.3	0.0
12.5	−63.4	−11.2	0.0
16	−56.7	−8.5	0.0
20	−50.6	−6.2	0.0
25	−44.7	−4.4	0.0
31.5	−39.4	−3.0	0.0
40	−34.6	−2.0	0.0
50	−30.2	−1.3	0.0
63	−26.2	−0.8	0.0
80	−22.5	−0.5	0.0
100	−19.1	−0.3	0.0
125	−16.1	−0.2	0.0
160	−13.4	−0.1	0.0
200	−10.9	0.0	0.0
250	−8.6	0.0	0.0
315	−6.6	0.0	0.0
400	−4.8	0.0	0.0
500	−3.2	0.0	0.0
630	−1.9	0.0	0.0

800	−0.8	0.0	0.0
1,000	0	0	0
1,250	+0.6	0.0	0.0
1,600	+1.0	−0.1	0.0
2,000	+1.2	−0.2	0.0
2,500	+1.3	−0.3	0.0
3,150	+1.2	−0.5	0.0
4,000	+1.0	−0.8	0.0
5,000	+0.5	−1.3	0.0
6,300	−0.1	−2.0	0.0
8,000	−1.1	−3.0	0.0
10,000	−2.5	−4.4	0.0
12,500	−4.3	−6.2	0.0
16,000	−6.6	−8.5	0.0
20,000	−9.3	−11.2	0.0

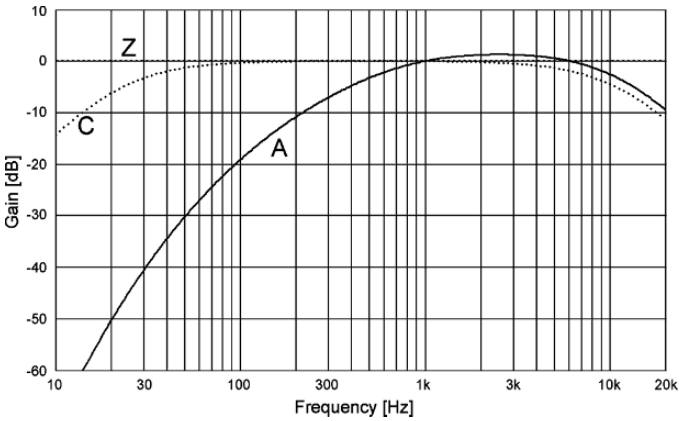


FIGURE 1 Frequency response of A-, C- and Z-weighting.

The weighting may be obtained by an analogue network or by digital signal processing. If a digital IIR network is used, some care is normally needed since the lowest characteristic frequency for the A-weighting is far below the typical sampling frequency [5]. Figure 2 shows a general second-order section for such applications with low sensitivity to truncation errors.

2.2 Networks for Vibration

A network with a gain proportional to $1/(2\pi f)$ may be used for transforming an acceleration signal to a signal proportional to the velocity. This is equivalent to an integration of the signal. This relationship can most easily be seen if an acceleration signal is assumed to be sinusoidal with amplitude A_0

$$a(t) = A_0 \sin(2\pi f t). \tag{7}$$

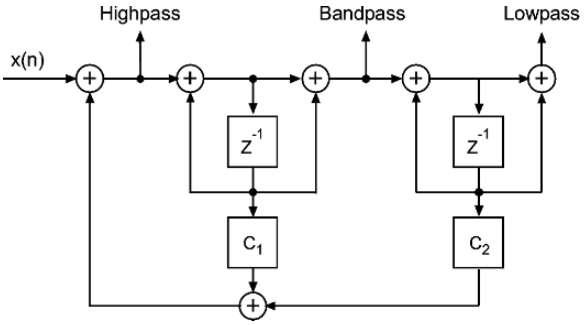


FIGURE 2 General second-order network with simultaneous outputs for lowpass, highpass and bandpass function. The network is well suited for filters with characteristic frequencies far below the sampling frequency (Courtesy: Norsonic AS).

The velocity may be obtained by time integration of the acceleration signal:

$$v(t) = \int_{-\infty}^{t} a(\tau)\mathrm{d}\tau = -\frac{A_0}{2\pi f}\cos(2\pi f t) = V_0\cos(2\pi f t). \qquad (8)$$

The change from sine to cosines indicates a 90° phase shift, which we have omitted in the network description since it will not influence the magnitude of the signal. The network may further be modified to take care of a requested frequency weighting like the weighting for whole body or hand-arm vibration. In a similar way, the displacement may be obtained from a velocity or an acceleration signal. Transformation in the opposite way is also possible by a suitable weighting.

3 • Octave- and Fractional-Octave-Bandpass Filters

For a long time, octave- and fractional-octave-band filters have been used in acoustics to measure the distribution of energy at different frequencies. The filters may be characterized as a constant relative bandwidth filter. This means that the bandwidth is a constant fraction of the midband or centre frequency of the filter. For octave filters, the upper bandedge frequency is an octave above the lower bandedge frequency.[1] Performance requirements for octave- and fractional-octave-band filters for use in acoustic analysis are specified in the International standard IEC61260 [6].

[1] See below for a definition of an octave in this context.

Octave- and fractional-octave-band filters are normally used in a set to cover a certain frequency range. This means that the bandedge frequency of neighbouring filters coincides so the total frequency range is covered by a number of adjacent filter bands. The levels in each band may then be used to describe the frequency distribution of the signal. Modern analyzers apply a set of parallel filters. Alternatively, one filter may sequentially be tuned to measure the level for each band. Figure 3 shows a typical set of octave-band filters with midband frequencies from 31.5 Hz to 16 kHz.

The most common bandwidths of these filters are octave and one-third octave, but 1/12 and 1/24 octave filters are also in common use.

The relation between the midband frequency and the bandwidth is chosen to make it possible to combine results obtained by a narrow filter set to results corresponding to a broader bandwidth. Three values based on 1/3-octave analysis may be converted to a 1/1-octave value. The accuracy is normally high for broadband signals, but for pure tone signals, some deviation may be observed in the transition range between the bands of the more narrow filters. Figure 4 shows the response of a typical octave-band filter and a filter synthesized by the summation of three bands with 1/3-octave width.

Although the name is octave- or fractional-octave-band filters, the characteristic frequencies for the filter are adjusted to fit the decade scale. If f_m is the midband frequency of a filter with a fractional-octave-bandwidth designator, b (b is whole number

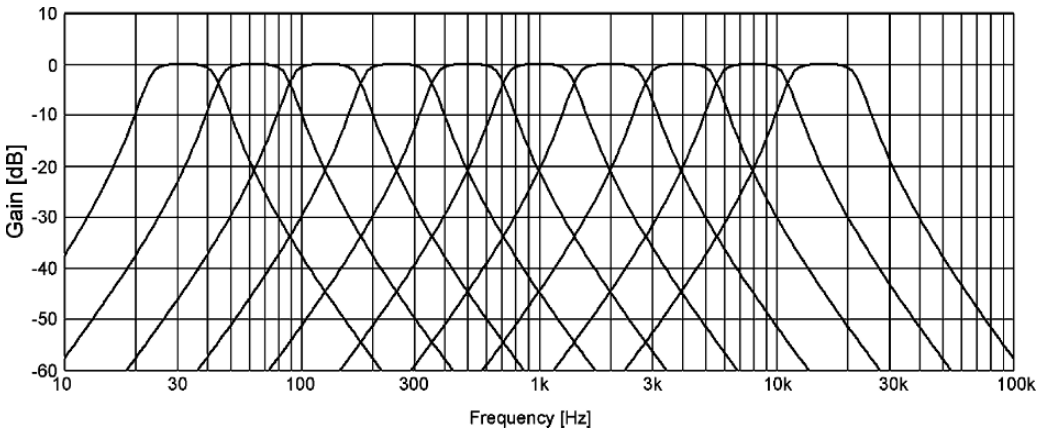


FIGURE 3 Frequency response of a typical set of octave-band filters in the range 31.5 Hz to 16 kHz. Filters may have steeper skirts in the transition between passband and stopband.

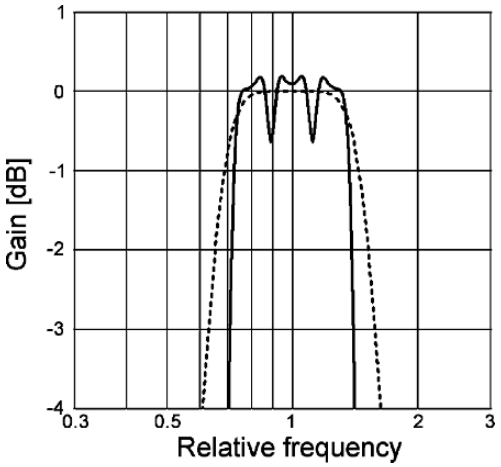


FIGURE 4 Comparison of a one-octave-band filter (- - - -) and a filter obtained as the sum of three one-third octave filters (——). Note that the sum is the sum of the squared signals (power) and thus the response is determined from the magnitude alone and is independent of the phase response of the filters. The diagram shows a typical response and may be different for other filter designs.

1, 2 or 3, etc.), the lower bandedge frequency f_L and the upper bandedge frequency f_H is given by

$$f_L = f_m \cdot 10^{3/20 \cdot b}, \tag{9a}$$

$$f_H = f_m \cdot 10^{-3/20 \cdot b}. \tag{9b}$$

For octave filters ($b = 1$), we have

$$\frac{f_H}{f_L} = 10^{3/10} = 1.995\ldots \approx 2, \tag{10}$$

which explains the designation for this filter since "octave" is used by musicians to describe tones for which the ratio between the pitches is two.

The midband frequency f_m is determined by the equation:

$$f_m = f_R \cdot 10^{3x/10 \cdot b} \quad \text{if } b \text{ is odd}, \tag{11a}$$

$$f_m = f_R \cdot 10^{3(2x+1)/20 \cdot b} \quad \text{if } b \text{ is even}, \tag{11b}$$

where f_R is the reference frequency 1,000 Hz, and x is any positive or negative whole number including zero. Rounded values for the

TABLE 2 Nominal and exact (rounded) frequencies in the range 1 kHz to 10 kHz for one-third octave-band filters

Designation (kHz)	x	Exact (kHz)
1	0	1.000000
1.25	1	1.258925
1.6	2	1.584893
2	3	1.995262
2.5	4	2.511886
3.15	5	3.162278
4	6	3.981072
5	7	5.011872
6.3	8	6.309573
8	9	7.943282
10	10	10.000000

frequency are in common use for designating the different filters for octave- and one-third octave-band filters. Table 2 shows the exact midband frequencies rounded to six digits together with the nominal values used for designation. The table applies to one-third octave-band filters in the frequency range from 1 kHz to 10 kHz. Similar figures, where the decimal point is shifted as appropriate, are used for any decade. A selection of every third number, starting at 1 kHz, is used for octave-band filters.

The relative bandwidth $(f_H - f_L)/f_m$ will be 70.5 % for an octave filter, 23.1 % for one-third octave and 2.88 % for a 1/24th octave.

3.1 Power Spectral Density

A pure tone and a noise signal behave differently when they are analyzed with a filter with a certain bandwidth. The level of the pure tone will be almost independent of the bandwidth as long as the frequency is within the passband, whereas the level of a noise signal will be proportional to the square root of the bandwidth. Most commercial octave- and fractional-octave analyzers have a noise bandwidth close to the nominal value. The power spectral density (PSD) of the signal may be obtained by dividing the signal strength with the square root of the bandwidth [7]. Thus, if the fractional-octave value of the sound pressure is p, the mean PSD for the band will be

$$\mathrm{PSD} = \frac{p}{\sqrt{f_m \cdot \left(10^{3/20 \cdot b} - 10^{-3/20 \cdot b}\right)}}, \tag{12}$$

where f_m and b are as defined above.

3.2 Temporal Distribution

General signal theory states that the product of frequency resolution Δf and time resolution Δt, when describing the signal strength, cannot be selected below a certain limit [8]:

$$\Delta f \cdot \Delta t \geq 2\pi. \tag{13}$$

The implication of this is that the frequency resolution should not be too high (narrow bandwidth) if we want to have a high temporal resolution.

An application in acoustics where we often are restricted by this phenomenon is the measurement of short reverberation times. The reverberation time of a room, T_{60}, is the time, in seconds, that would be required for the sound pressure level to decrease by 60 dB after the sound source has stopped. Any operation to restrict the frequency content will create a virtual reverberation time. If this virtual reverberation time is not considerably less than the reverberation time for the room, the virtual reverberation will destroy the measurement accuracy.

Analogue and digital IIR fractional-octave-band filters with a bandwidth, B, in hertz, can normally be used for measurement of reverberation time without significant error if $T_{60} \geq 16/B$ [9]. However, if the impulse response of the analyzing filter is time reversed – as may be done by storing the signal – this limit can be reduced to one-fourth [10]. See Figure 5. The values are typical for filters found in commercial sound levels meters but will vary with the filter design. Filters with high selectivity may have even longer decay time. Digital FIR filters may have a limit close to the limit set by the time-reversed technique. In these cases, the lower limit is strongly dependent on which part of the decay is used for the extraction of the reverberation time. In general, the first part of the decay – the initial reverberation time or the early decay – is more influenced by the filter than the later part of the decay.

3.3 Weighted Response from Bandpass Filtered Values

It is often required to calculate weighted levels like the A-weighted level from the result of a fractional-octave-band analysis. This is done by obtaining the weighted sum of the power in each band. If the sound level in the fractional-octave-band number n with midband frequency $f_{\mathrm{m}}(n)$ is L_n the weighted level is given by

$$L_{\mathrm{w}} = 10\log\left\{\sum_n w^2[f_{\mathrm{m}}(n)]\cdot 10^{L_n/10}\right\} = 10\log\left\{\sum_n 10^{(L_{\mathrm{n}}+L_{\mathrm{weight}}(f_{\mathrm{m}}(n)))/10}\right\}, \tag{14}$$

where $w[f_{\mathrm{m}}(n)]$ may be any of the previously defined weightings A, C or Z.

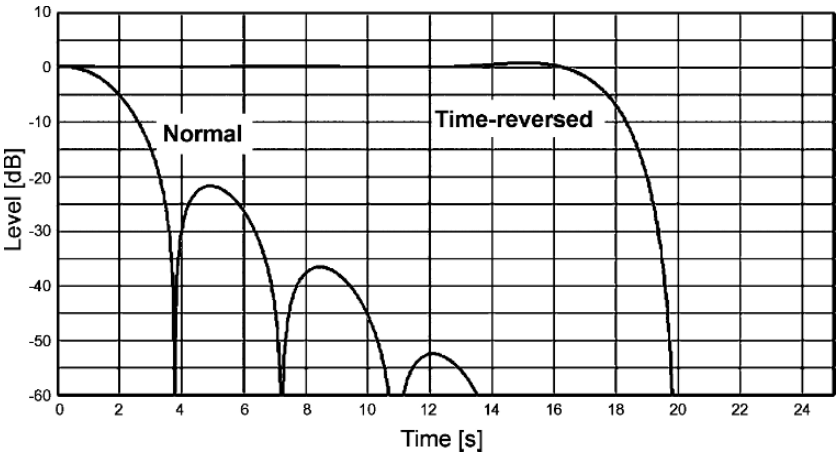


FIGURE 5 The envelope of the output from a 6-pole, Butterworth bandpass filter when a sinusoidal input signal with a frequency equal to the midband frequency of the filter is suddenly switched off is shown. The decay marked "Normal" is from an ordinary analogue filter (IIR). The decay marked "Time reversed" is when the impulse response for the "Normal" filter is time reversed. Note the delay needed to make the filter realizable. The difference in the decay rate is barely different in the first part of the decay (e.g. early decay time), but the rate is significantly higher for a larger dynamic range for the time-reversed filter. The "Normal" and the "Time-reversed" filter will have equal frequency response but a different phase response.

The accuracy of the calculation will in general be better the higher the resolution in the frequency analysis. One-third octave or more narrow analysis will normally be adequate for calculating the A- and C-weighted levels corresponding to the requirements for a Class 1 sound level meter. For sinusoidal signals, the largest deviations will typically be close to the transition frequencies between adjacent bands. The use of octave analysis for such calculation should not be recommended for other purposes than indication of weighted levels. Figure 6 shows the effective frequency response function obtained as the sum of typical octave-band analysis. Different filter shapes may make the deviation between the weighted sum and the ideal A-weighted response even larger.

3.4 Fractional-Octave-Band Analyzers

Fractional-octave-band analyzers contain a set of fractional-octave-band filters with a detector for each filter. The most common are octave-band and one-third octave-band analyzers. The signal is normally analyzed in real time, and such analyzers are therefore often called "real-time analyzers". Figure 7 shows

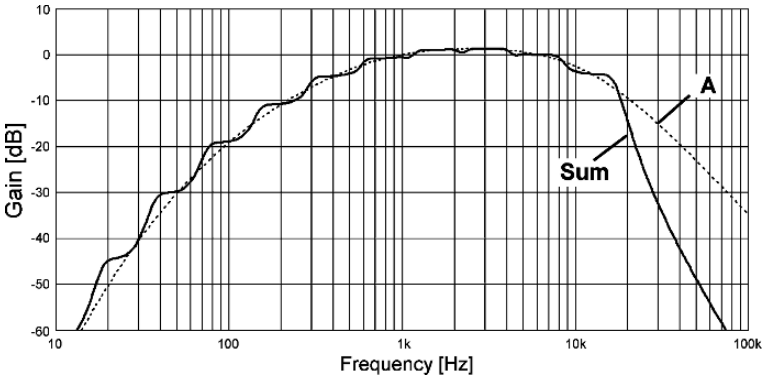


FIGURE 6 Comparison of the ideal A-weighting and the effective weighting obtained by summation of the response from a typical 1/1-octave-band analysis. For most filter designs, the result will not satisfy the requirement in the standard for sound level meters IEC 61672-1.

a typical display from a digital hand-held sound level meter with integrated octave band analyzer. The indicated level may be the instantaneous value like the *F*- or *S*-time weighted level or the mean level (equivalent value) for a specified measurement period.

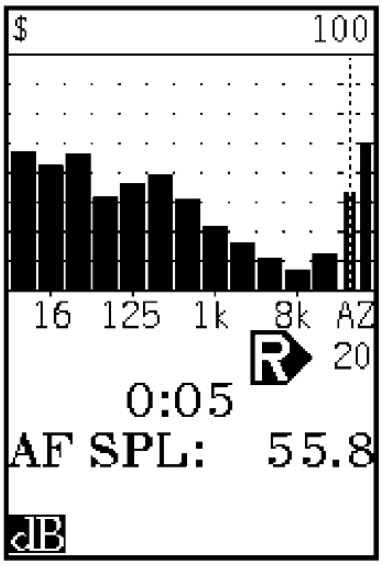


FIGURE 7 Display from a sound level meter with 1/1-octave-band analysis (Courtesy: Norsonic AS).

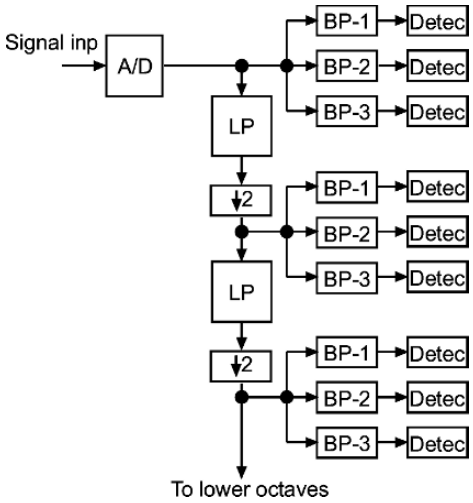


FIGURE 8 Analyzer for one-third octave. The analogue-to-digital converter converts the analogue input to a digital signal. Sampling frequency is 64 kHz. Each octave is analyzed by three 1/3-octave-band filters. The signal to the next lower octave is lowpass filtered, and the sampling is decimated to use every second sample. The block diagram corresponds to the filter implementation in the analyzer Nor830 (Courtesy: Norsonic AS).

Some digital analyzers reduce the sampling frequency for the lower octaves. According to the sampling theorem, each lower octave can use the half of the sampling frequency compared to the octave above. If a digital processor has adequate capacity for analyzing two octaves with the highest sampling frequency, it will have sufficient capacity to analyze all lower octaves if the sampling frequency is lowered for each octave. This can be shown from the sum property of geometric series (full capacity, half capacity, etc.):

$$C\left(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots\right)=2C. \tag{15}$$

The block diagram in Figure 8 shows a third-octave analyzer based on this method.

4 • Phase Linear Filters

Filters discriminate signals due to their frequency contents. A certain length of the signal is needed to determine the frequency, and the length needs to be longer for better frequency resolution.

All filters operating in real time therefore delay the signal. Most filters will have a delay that varies with the frequency. Filters having a frequency-independent delay are called linear phase filters. If the phase shift (in radians) through a filter is $\phi(f)$, the group delay is given by the equation [1]

$$\tau_{\text{group}} = -\frac{1}{2\pi}\frac{\partial\varphi(f)}{\partial f}. \tag{16}$$

The group delay specifies the delay at frequencies around f.

Only finite impulse response filters (FIR) can have a true linear phase. The most common type is a digital filter with a structure as shown in Figure 9. Symmetry in the coefficients, a_n, around, a_0, will guarantee a symmetrical impulse response around the delay N.

Analogue, continuous time filters and digital IIR filters cannot have a true linear phase, but the response can approximate a linear phase response to any degree of accuracy.

For some application, it will be important that more filters have the same delay. A typical example is filters used for dividing the signal in different sub-band for later reconstruction. Applications may be transmission of the sub-bands in different transmission channels or in perceptual coding of sound.

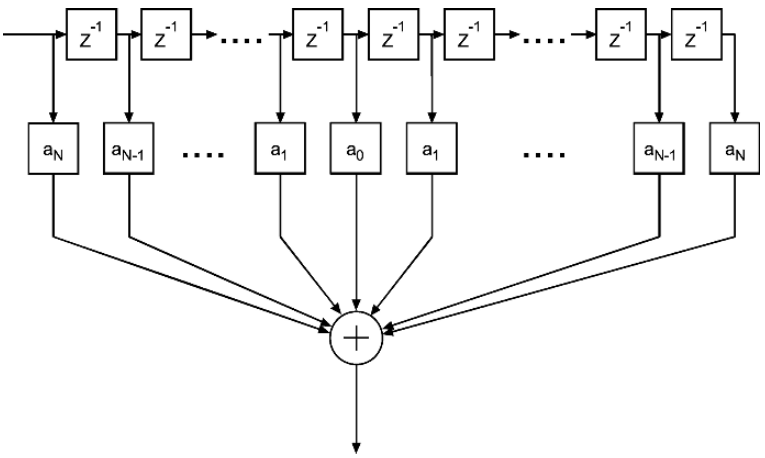


FIGURE 9 Finite impulse response (FIR) digital filter. Due to the symmetry in the filter coefficients, the filter will have a linear phase and a delay of N samples.

5 • In-Phase/Quadrature Filters

In-phase/quadrature filters are a special case of a polyphase filter. One application is to build narrow band filters like filters used in lock-in amplifiers. This is illustrated with an example in Figure 10. The signal to be filtered is modulated with two sinusoidal signals with angular frequency, ω_0, where the phase is in quadrature like a sine and a cosine signal. The spectrum of the signal will be frequency translated with the modulation frequency ω_0. The spectrum centred at ω_0 will, after modulation, be centred at the frequencies 0 (zero) and $2\omega_0$. In a similar way the spectrum centred at $-\omega_0$ will, after modulation, be centred at the frequencies 0 (zero) and $-2\omega_0$. A lowpass filter will remove the signal energy for those frequencies that are not close to zero. After a new modulation, the spectrum is again transferred to frequencies around ω_0. This will happen only for those frequency components that were not removed by the lowpass filtering. Thus the lowpass filters are performing an operation like a narrow bandpass filter.

Since the modulated signal around zero frequency is the result from the sum of two complex signals – positive and negative frequencies – at least two channels (paths) are needed to separate the two spectral components after the second modulation. If only one channel is used, the result will depend on the phase difference between the input signal and the modulation.

In a digital realization, the lowpass operation can often be performed at a reduced sampling frequency and thus reduce the needed capacity for processing power. A further decrease

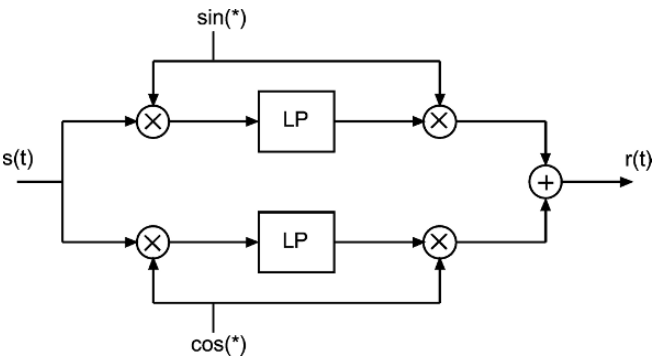


FIGURE 10 In-phase/quadrature filter where two lowpass filters are used to perform a bandpass operation.

in processing power can be gained in situation where the modulation can be substituted with two square waves in quadrature. In a digital design, this means that every second sample is routed to each of the channels.

In cases where only the filtered signal power shall be measured, the second modulation in each channel is not needed. The power can be obtained as the sum of the signal power at the output from each lowpass filter.

6 • Perceptual Masking Filters

Most of the sounds we observe have an envelope that changes slowly compared to the change in the sound pressure due to the frequency of the sound. If we sample such a signal, the information in the signal will then be less than if each of the samples were independent of the adjacent sample. This allows in most cases the signal to be coded in a way that reduces the needed capacity for transmission without loss of information (loss-less coding). A further reduction in the needed transmission capacity can be obtained if the perceptual limitation of our auditory system is utilized (perceptual coding). Perceptual coding is lossy; this means the original signal cannot be completely regenerated.

Sound with a level below a certain frequency-dependent limit cannot be heard. Furthermore, in the human auditory system, a strong frequency component will mask a neighbouring frequency component with a lower level. In particular, lower frequencies tend to mask higher frequencies. Spectral components that cannot be heard can be removed from the transmission without deteriorating the perceived sound quality. Use of this phenomenon is called perceptual coding. Masking can occur in the time domain (temporal masking) or in the frequency domain (frequency masking). A filter set is a major part of a perceptual coding algorithm [11]. The filter set may have a constant bandwidth or a bandwidth corresponding to the critical bandwidth of the auditory system. The critical bandwidth is assumed to be around 100 Hz for frequencies up to 500 Hz and more like a constant percentage bandwidth for higher frequencies (bandwidth about 15 % of frequency). Examples of common perceptual coding formats are MP3, WMA and ATRAC, but more than 100 different formats are currently in use. The capacity for transmission or storing is typically reduced to 5–10 % of the capacity needed for storing or transmitting the un-coded signal.

References

1. H.J. Blinchikoff, A.I. Zverev: *Filtering in the Time and Frequency Domains*. Wiley, 1976, ISBN 0-471-98679-8.
2. L.E. Kinsler, A.R. Frey: *Fundamental of Acoustics*. 2 edition. Wiley, 1962.
3. ISO 226 (2003): *Acoustics – Normal Equal-Loudness-Level Contours*.
4. IEC 61672-1 (2002–05): *Electroacoustics – Sound level meters – Part 1: Specifications*.
5. S.K. Mitra, J.F. Kaiser: *Handbook for Digital Signal Processing*. Wiley, 1993, ISBN 0-471-61995-7.
6. IEC 61260 (1995–08): *Electroacoustics – Octave-Band and Fractional-Octave-Band filters*.
7. J.S. Bendat, A.G. Piersol: *Engineering Applications of Correlation and Spectral Analysis*. Wiley, 1980, ISBN 0-471-05887-4.
8. J.B. Thomas: *An Introduction to Statistical Communication Theory*. Wiley, 1969.
9. F. Jacobsen: A note on acoustic decay measurements. *Journal of Sound and Vibration* (1987) **115**(1), 163–170.
10. F. Jacobsen, J.H. Rindel: Time reversed decay measurements. *Journal of Sound and Vibration* (1987) **117**(1), 187–190.
11. K. Brandenburg: *MP3 and AAC explained. AES* 17th International Conference on High Quality Audio Coding.

8

Adaptive Processing

Thorkild Find Pedersen

Brüel & Kjær, Sound & Vibration Measurements A/S, Nærum, Denmark

1 • Introduction

In acoustic signal processing, adaptive methods have found practical use in different applications, e.g. echo and noise cancelation, signal enhancement, system identification, adaptive equalization and speech coding.

Figure 1 shows the structure of an adaptive transversal filter. It is one of the simplest adaptive filter structures [1,2]. The filter consists of a Finite Impulse Response (FIR) filter of the order N with the adjustable coefficients $\mathbf{w} = [w_0 w_1 \cdots w_N]^{\mathrm{T}}$.

The objective of the adaptive algorithm is to adjust the filter coefficients such that some predefined cost function is minimized. Typically, the cost function is the mean square error between the desired output $d(n)$ and $y(n)$ being the filtered reference signal $u(n)$. Using matrix notation, $y(n)$ is computed as the dot product of $\mathbf{w}$ and the reference vector $\mathbf{u}(n) = [u(n)u(n-1) \cdots u(n-N)]^{\mathrm{T}}$.

In the ideal case, the filter output $y(n)$ will be the part of the reference signal that correlates with the desired output $d(n)$. The adaptive filter will thus attempt to cancel out the interference of the reference signal from the desired output resulting in the error signal $e(n)$. The adaptive filter is therefore suitable for interference canceling when a reference to the disturbing noise in the desired output signal is obtainable.

Adapting the filter coefficients is an iterative process where the coefficient $\mathbf{w}$ converges to an optimal value as the number of available samples increases. The notation $\mathbf{w}(n)$ is used to indicate

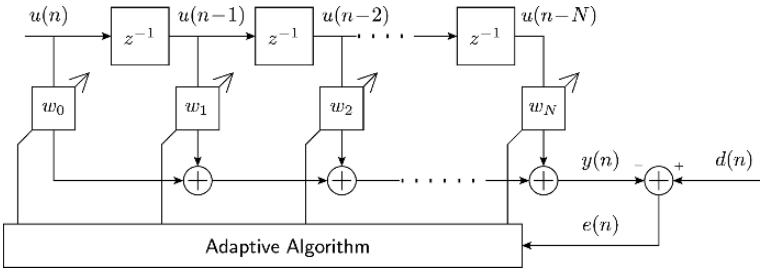


FIGURE 1 Adaptive transversal filter.

the estimates of the coefficients after the nth sample. There are numerous algorithms for adjusting the filter coefficients of which the Least Mean Square (LMS) algorithm described below is the most famous. The LMS algorithm uses the method of steepest descent to optimize the filter weights and converges relatively slowly. Faster convergence rates are obtainable with other algorithms, e.g. the Recursive Least Squares (RLS) algorithm but at the expense of higher computational complexity.

When comparing adaptive algorithms the important factors to consider are *rate of convergence*, the number of iterations it takes for the algorithm to converge to the optimum solution for a stationary input; *misadjustment*, the bias and variance of the error signal, i.e. the accuracy of filter; *tracking*, the ability of the filter to track nonstationary signals; and *robustness*, the sensitivity of the algorithm to internal or external disturbances and ability to avoid local optimum solutions.

2 • The LMS Algorithm

A good first choice algorithm for investigating adaptive processing is the LMS or the Normalized LMS (NLMS) algorithm. It is among the simplest adaptive algorithms and easy to implement. The general equations for the adaptive transversal filter is shown in (1) where the first step is to compute the difference of the current desired output $d(n)$ and the filtered output $y(n)$ based on the last estimate of the filter coefficients $\mathbf{w}(n-1)$. Depending on the algorithm type, an update step $\boldsymbol{\Delta}(n)$ is computed for updating the filter coefficients.

$$\begin{aligned} e(n) &= d(n) - \mathbf{w}^{\mathrm{T}}(n-1)\mathbf{u}(n) \\ \mathbf{w}(n) &= \mathbf{w}(n-1) + \boldsymbol{\Delta}(n)e(n) \end{aligned} \tag{1}$$

For the LMS algorithm, the update step $\boldsymbol{\Delta}(n)$ is given by

$$\boldsymbol{\Delta}(n) = \mu\mathbf{u}(n) \tag{2}$$

where μ is a step-size parameter used to tune the adaption rate of the filter. The LMS algorithm converges for wide sense stationary signals when μ satisfies ($E\{\cdot\}$ is the expectation operator)

$$0 < \mu < \frac{2}{\sum_{k=0}^{N} E\{|u(n-k)|^2\}} \tag{3}$$

In the standard LMS algorithm, the update of the filter coefficients, $\mathbf{w}(n)$, is proportional to the input vector $\mathbf{u}(n) = [u(n)\ u(n-1)\cdots u(n-N)]^{\mathrm{T}}$.

A problem with the LMS algorithm is *gradient noise amplification* caused by the fact that noise in the reference signal propagates directly into the filter coefficients. This is avoided in the NLMS where the update step is made time dependent by normalizing it with the reference vector, i.e.

$$\boldsymbol{\Delta}(n) = \frac{\mu \mathbf{u}(n)}{\delta + \|\mathbf{u}(n)\|^2} \tag{4}$$

where δ is a small positive constant to avoid numerical problems for small input signals. Generally, the NLMS algorithm converges faster than the LMS algorithm.

For large numbers of filter coefficients, it is beneficial to use the *Fast LMS* also called *Frequency Domain LMS* algorithm. The Fast LMS algorithm performs the filtering in the frequency domain using the fast Fourier transformation. The input data is therefore processed in blocks equivalent to the required FFT length, and the filter coefficients are only updated after each FFT block.

Table 1 summarizes the LMS, NLMS and RLS algorithms with respect to computing the update step in (1).

TABLE 1 Adaption algorithms for computing the update step in (1). For LMS and NLMS, the adaption rate is controlled by the step-size parameter μ. δ is a small positive constant. For RLS, it is controlled by the forgetting factor λ, ($0 \ll \lambda < 1$). The closer λ is selected to 1, the slower the RLS algorithm adapts and also the more accurate it will be for stationary signals

Algorithm	Update step	Initialization
LMS	$\boldsymbol{\Delta}(n) = \mu \mathbf{u}(n)$	$\mathbf{w}(0) = \mathbf{0}$
NLMS	$\boldsymbol{\Delta}(n) = \frac{\mu \mathbf{u}(n)}{\delta + \|\mathbf{u}(n)\|^2}$	$\mathbf{w}(0) = \mathbf{0}$
RLS	$\boldsymbol{\Delta}(n) = \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{u}(n)}{1+\lambda^{-1}\mathbf{u}^{\mathrm{T}}(n)\mathbf{P}(n-1)\mathbf{u}(n)}$	$\mathbf{w}(0) = \mathbf{0}$
	$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\boldsymbol{\Delta}(n)\mathbf{u}^{\mathrm{T}}(n)\mathbf{P}(n-1)$	$\mathbf{P}(0) = \delta\mathbf{I}$

3 • Adaptive Applications

Figure 2 shows four typical applications of the adaptive filter: (a) *system identification*, (b) *inverse modeling*, (c) *prediction* and (d) *interference canceling*.

When used for system identification, the filter is placed parallel with the unknown plant to be identified. The filters impulse response is then optimized to match that of the plant by using the driving signal of the plant as input and the plants output as the desired output.

For inverse modeling, the adaptive filter is placed in series with the output from the plant and uses the driving signal to the plant as the desired output. The adaptive filter will therefore attempt to cancel out the effect of the plant, hence the filters impulse response is optimized to match the inverse impulse response of the plant.

For prediction, the adaptive filter is optimized to predict future samples of the input signal. If the input signal is periodic with additive noise, then the periodic part can be separated from the noise if the delay M is larger than the correlation length of the noise. This use of the adaptive filter is also known as the Adaptive Line Enhancer (ALE) or Self-Adaptive Noise Cancelation (SANC). It is possible to cascade ALEs with different properties. In [3], Lee and White describe a two-stage ALE for enhancing impulsive noise from rotating and reciprocating machines. The first ALE

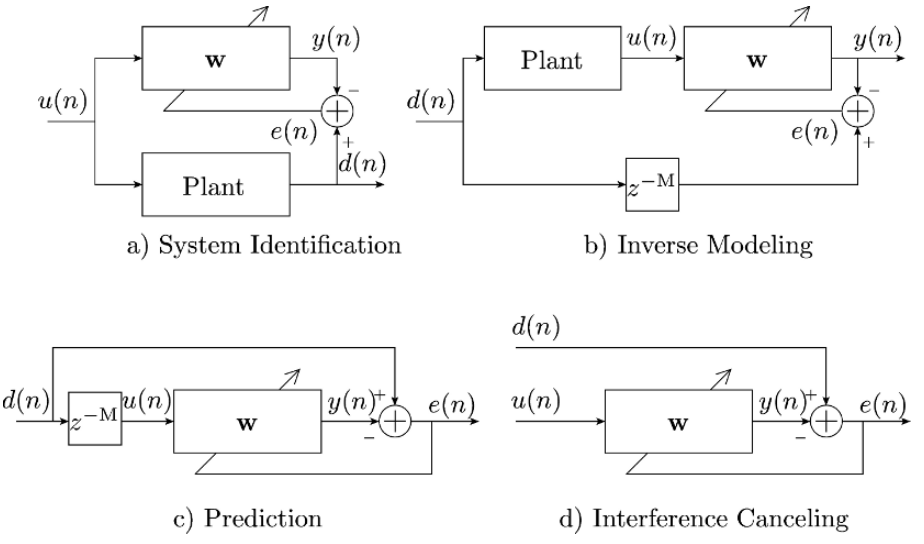


FIGURE 2 Applications of the adaptive filter.

serves to remove a strong periodic component from the signal using a relatively long delay and a transversal filter with many coefficients; the second ALE uses a shorter delay and less filter coefficients to enhance the transients in the error signal from the first ALE.

For interference canceling, the adaptive filter is used as mentioned previously, where the desired reference signal in some way is correlated with the additive noise in the desired output signal $d(n)$.

References

1. B. Widrow, S.D. Stearns: *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ 1985).
2. S. Haykin: *Adaptive Filter Theory* 3rd edition (Prentice-Hall, Englewood Cliffs, NJ, 1996).
3. S.K. Lee, P.R. White: The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery, *Journal of Sound and Vibration* (1998) **217**(3) 485–505.

9

Beamforming and Wavenumber Processing

Jørgen Hald

Brüel & Kjær, Sound & Vibration Measurement A/S, Nærum, Denmark

1 • Planar Nearfield Acoustical Holography

Planar nearfield acoustical holography (NAH) is a powerful experimental tool for studying and diagnosing noise radiation problems [1–11]. Based on measurements of typically sound pressure over a plane near the source, all sound field parameters can be calculated in any plane parallel with the measurement plane. In particular, the sound field can be mapped closer to the source than the measurement plane, which can provide very high spatial resolution of the source distribution. Since the calculation plane cannot in theory approach the source further than the nearest point on the source, planar NAH works best for but is not restricted to sources with a planar geometry. NAH formulations exist also for cylindrical, spherical and conformal geometries [3,5,6], but such formulations will not be treated here.

The geometry of the measurement problem to be considered is illustrated in Figure 1. All sound sources are assumed to be in the region $z < -d$, and the remaining half space $z \geq -d$ contains a homogeneous fluid with no viscosity. In that region,

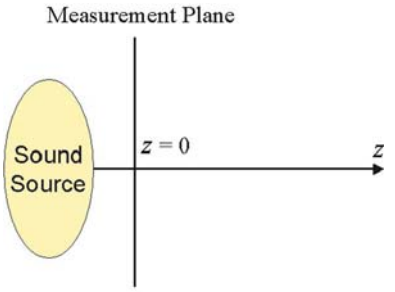


FIGURE 1 Measurement geometry.

the complex time-harmonic sound pressure $p(x, y, z)$ fulfils the Helmholtz equation [3],

$$\nabla^2 p + k^2 p = 0, \quad z \geq -d, \tag{1}$$

where $k = \omega/c = 2\pi/\lambda$ is the acoustic wavenumber, ω is the angular frequency, λ is the wavelength, c is the propagation speed of sound and ∇^2 is the Laplace operator. The implicit time factor is $e^{j\omega t}$.

For any given z-coordinate, $z \geq -d$, we now introduce the following Fourier transform pair of the sound pressure in the two dimensions (x, y):

$$P(k_x, k_y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y, z) e^{j(k_x x + k_y y)} dx\ dy, \tag{2}$$

$$p(x, y, z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(k_x, k_y, z) e^{-j(k_x x + k_y y)} dk_x\ dk_y. \tag{3}$$

Here, the spatial angular frequencies k_x and k_y in the x- and y-directions, respectively, are also denoted as directional wavenumbers. Through insertion of Eq. (3) in Helmholtz equation (1) it can be verified that the following relation must hold:

$$P(k_x, k_y, z) = P(k_x, k_y, z_0)\ G(k_x, k_y, z - z_0) \quad z \geq -d, \quad z_0 \geq -d, \tag{4}$$

where

$$G(k_x, k_y, z) \equiv e^{-jk_z z} \tag{5}$$

with $k_x^2 + k_y^2 + k_z^2 = k^2$, and more precisely because all sources are in the region $z < -d$,

$$k_z \equiv \begin{cases} \sqrt{k^2 - (k_x^2 + k_y^2)} & \text{for} \quad k_x^2 + k_y^2 \leq k^2 \\ -j\sqrt{(k_x^2 + k_y^2) - k^2} & \text{for} \quad k_x^2 + k_y^2 > k^2 \end{cases}. \tag{6}$$

Thus, if the sound pressure p is known on the entire measurement plane $z = 0$, the wavenumber spectrum $P(k_x, k_y) \equiv P(k_x, k_y, 0)$ can be calculated using Eq. (2), and the pressure in any plane parallel with the measurement plane can then be obtained using a combination of Eqs. (3) and (4):

$$p(x, y, z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(k_x, k_y) \mathrm{e}^{-\mathrm{j}(k_x x + k_y y + k_z z)} \mathrm{d}k_x \, \mathrm{d}k_y \quad z \geq -d. \quad (7)$$

For $k_x^2 + k_y^2 \leq k^2$, i.e. inside the so-called radiation circle in the (k_x, k_y) wavenumber plane, the function $\mathrm{e}^{-\mathrm{j}(k_x x + k_y y + k_z z)}$ in Eq. (7) represents a plane wave propagating in the direction of the 3D wavenumber vector $\mathbf{k} \equiv (k_x, k_y, k_z)^{\mathrm{T}}$, see Figure 2. For $k_x^2 + k_y^2 > k^2$, the same function represents a so-called evanescent wave that propagates along the xy-plane in the direction $(k_x, k_y, 0)^{\mathrm{T}}$, but decays exponentially in the positive z-axis direction, i.e. away from the sources. Through the spatial Fourier transform in Eq. (2) we have therefore obtained a 3D representation of the sound pressure field in terms of plane propagating and plane evanescent waves. Based on the 3D pressure distribution in Eq. (7), Euler's

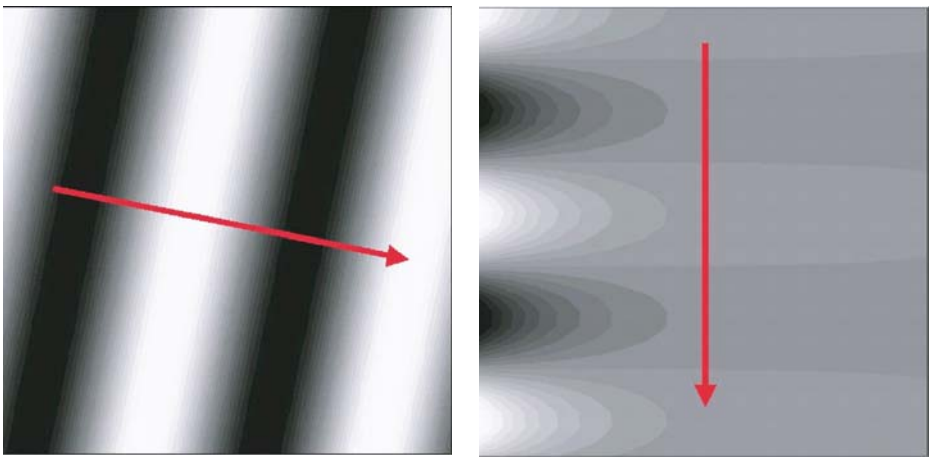


FIGURE 2 The two figures illustrate a plane propagating wave (*left*) and an evanescent wave with $k_x^2 + k_y^2 = (1.1k)^2$ (*right*). An arrow indicates the direction of wave propagation. The "source plane" $z = -d$ constitutes the leftmost limitation in both figures.

equation leads to the following expression for the particle velocity vector **u** [3]:

$$\mathbf{u}(x,y,z) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(k_x,k_y)\frac{\mathbf{k}}{\rho_0 ck}\mathrm{e}^{-\mathrm{j}(k_x x+k_y y+k_z z)}\mathrm{d}k_x\ \mathrm{d}k_y, \qquad (8)$$

where **k** is the wavenumber vector defined above and ρ_0 is the density of the homogeneous fluid. In particular, the normal component of the particle velocity, u_z, can be estimated on the source plane $z = -d$, and from the pressure and the particle velocity the sound intensity can be obtained.

1.1 Regularization

For $z \geq 0$ the sound field reconstruction based on Eqs. (7) and (8) is a forward simulation of wave propagation, which is numerically stable. But for $z < 0$ the reconstruction is an unstable inverse calculation, in which the evanescent wave components with $k_x^2 + k_y^2 > k^2$ must be exponentially amplified according to Eqs. (4)–(6). Measurement inaccuracies will therefore be amplified and totally dominate the reconstruction, unless some type of regularization is introduced.

The standard approach for regularizing the solution is to use a spatial low-pass filter in the form of a window function in the wavenumber plane (k_x, k_y). To prevent exponential amplification of error components, the window must cut off faster than the exponential amplification of the inverse calculation in Eq. (4). A widely used window function is the exponential window Π, defined as [3]

$$\Pi(k_x,k_y,k_\mathrm{c}) \equiv \begin{cases} 1-\frac{1}{2}\mathrm{e}^{-(1-k_\mathrm{r}/k_\mathrm{c})/\alpha} & k_\mathrm{r} \equiv \sqrt{k_x^2+k_y^2} \leq k_\mathrm{c} \\ \frac{1}{2}\mathrm{e}^{(1-k_\mathrm{r}/k_\mathrm{c})/\alpha} & k_\mathrm{r} \equiv \sqrt{k_x^2+k_y^2} > k_\mathrm{c}, \end{cases} \qquad (9)$$

where k_r is the radial wavenumber, k_c is the radial cut-off wavenumber and α is a positive constant that determines how sharp the window cuts off. When α decreases towards zero, the window function Π approaches a rectangular window (Figure 3). Typically $\alpha = 0.1$ is used. A key problem is the identification of a cut-off wavenumber k_c that will prevent significant influence of measurement errors, but allow the largest possible part of the evanescent waves to be reconstructed. This can of course be done through a trial-and-error process, but typically it is done in the following way through specification of a signal-to-noise ratio *SNR* (dB): Considering reconstruction of the sound pressure in the source plane $z = -d$, the wavenumber spectrum of the pressure in that plane is assumed to be roughly constant, $|P(k_x, k_y, -d)| \approx P_\mathrm{S}$.

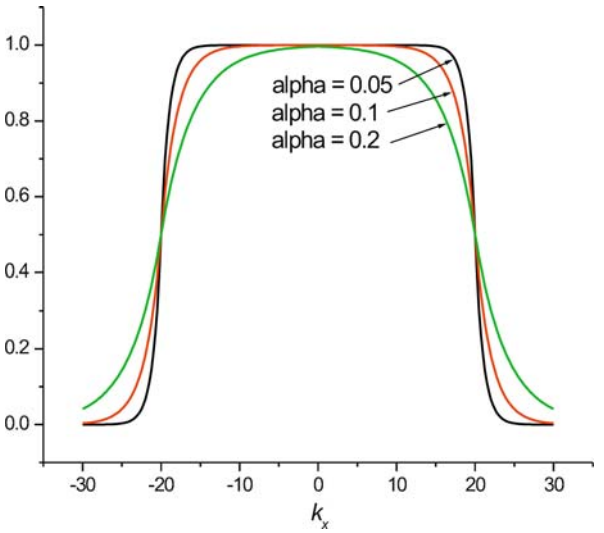


FIGURE 3 The exponential window function Π of Eq. (9) for $k_c = 20$, with α equal to 0.05, 0.1 and 0.2.

In the measurement plane, the propagating waves with $k_x^2 + k_y^2 \leq k^2$ will maintain the same amplitude P_S, while the evanescent waves will be attenuated by a factor $|G(k_x, k_y, d)|$. The second assumption is that the errors on the measured data will add a rather constant wavenumber spectrum, the level of which is approximately *SNR* (dB) lower than the level P_S of the plane wave spectrum. This assumption will hold if the measurement errors are spatially incoherent and incoherent with the sound sources. The radial cut-off wavenumber k_c is then chosen at the point where the level of the evanescent waves equals the level of the error spectrum:

$$e^{-\sqrt{k_c^2 - k^2}d} = 10^{-\mathrm{SNR}/20} \quad \Rightarrow$$

$$k_c = \sqrt{k^2 + \left(\frac{\mathrm{SNR}}{20d\log_{10}(e)}\right)^2} \simeq \sqrt{k^2 + \left(\frac{\mathrm{SNR}}{8.7d}\right)^2}. \qquad (10)$$

Since data will be dominated by errors for radial wavenumbers larger than k_c, the filter can be used for any type of calculation.

The exponential window (also called the exponential filter) requires manual selection of the parameters α and k_c. A couple of other k-space filters have been described in the literature, for which there exist methods to determine the filter coefficients from a set of noisy measurement data: the standard Tikhonov filter and the high-pass Tikhonov filter [4,5].

1.2 Spatial Resolution

The spatial low-pass filter needed to regularize the reconstruction for $z < 0$ will of course introduce a smoothing and thereby limit the spatial resolution. The radial cut-off wavenumber k_c introduces a lower limiting wavelength of $\lambda_c = 2\pi/k_c$ on the reconstruction plane, implying that resolution cannot be better than approximately R_c given by [3]

$$R_c = \frac{\lambda_c}{2} = \frac{\pi}{k_c} \simeq \frac{1}{\sqrt{(2/\lambda)^2 + (\mathrm{SNR}/27.3d)^2}}, \tag{11}$$

where use has been made of Eq. (10). At high frequencies, where the wavelength λ is small, the resolution is approximately half a wavelength, but at low frequencies it approaches a constant equal to $27.3d/\mathrm{SNR}$. So with 27.3 dB dynamic range the low-frequency resolution is approximately equal to the measurement distance.

1.3 Discrete Measurements

Until now we have assumed the pressure to be known at all points on the measurement plane $z = 0$. In practice measurements must be taken at discrete points over a finite area. Provided the pressure transducer is ideal, i.e. it measures the sound pressure at a point without disturbing the sound field, the discretization corresponds to a spatial sampling. In order to allow the spatial Fourier transforms to be evaluated using DFT algorithms, the measurement points must constitute a regular rectangular grid with constant spacing in the x- and y-directions. We will assume the same spacing a in both directions. As mentioned above, the propagation from the source plane $z = -d$ to the measurement plane $z = 0$ acts as a spatial low-pass filter, such that above the radial wavenumber k_c data will be dominated by errors. Since there is no reason to take into account the flat error spectrum in the spatial sampling, we just need at least two spatial samples per period up to the wavenumber k_c. This leads to the criterion

$$a \leq \lambda_c/2 = R_c, \tag{12}$$

where R_c is the resolution defined in Eq. (11). In some practical cases the sample spacing (the microphone spacing) is given by a specific hardware configuration. Equations (11) and (12) can then be used to derive a lower limit on the measurement distance d. But notice that this minimum measurement distance and the maximum sample spacing in Eq. (12) are strict requirements based on the assumption that all wavenumber components have almost the same level in the source plane $z = -d$. If the pressure distribution on the source plane is "smooth", the requirements can be relaxed.

1.4 Finite Measurement Area

So far we have assumed the measurement grid to cover the entire measurement plane. The spatial sampling corresponds to multiplication with a 2D spatial comb function, which will introduce periodic replica in the wavenumber domain, but still the periodic wavenumber spectrum is in principle continuous. If we can assume the pressure samples to be almost zero outside a certain rectangular area, then we need only measure that area, and also the Fourier integral (2) needs only cover that area. If, however, the pressure is not zero outside the measurement window, then by restricting the area of the measurement and of the Fourier integration we implicitly assume the pressure to equal zero outside. The use of such a rectangular window function in the (x, y) plane corresponds to convolution with a sinc function in the wavenumber domain, which will cause components of the periodic repetitions to leak into the "baseband" and appear as aliasing. Another effect of the wavenumber leakage is that propagating waves will leak into the evanescent wave area and vice versa. The wavenumber leakage can be reduced through the use of a smooth window function instead of the rectangular window. To avoid loss of information over the central parts of the measurement area, a Tukey window is often used [3,4].

The fact that serious errors ("spatial window effects") occur if the measurement window does not extend into regions where the pressure has dropped off significantly is in many cases a serious limitation. A couple of the so-called *Patch holography* methods have been introduced to avoid this requirement. One method is to extrapolate outside the measurement area [6]. This method uses the measured data inside the measurement window and estimates the pressure in a region outside through an iterative process. Another method named SONAH (statistically optimal NAH) completely avoids the use of spatial Fourier transforms, and therefore the implicit assumption of zero pressure outside the measurement area is avoided [7,8]. In addition, since spatial DFT processing is not used, SONAH can operate with irregular measurement grids.

1.5 Spatial DFT Processing

With discrete measurement points and a finite measurement area it is attractive to perform the spatial Fourier transform with a DFT, which can be efficiently calculated using an FFT. The main implication on the result is associated with the fact that now data will be discrete in the wavenumber domain also, which is of course necessary for practical reasons: The discrete data set in the wavenumber domain represents a periodic pressure distribution in the measurement plane. When the discrete wavenumber data set is used to calculate for example the pressure in a

parallel plane, then the periodic replica of the measurement area will contribute to the calculated pressure in principle over the entire calculation area. Errors introduced by these replicated measurement areas are called wrap-around errors [1]. The standard way to reduce wrap-around errors is to perform a zero-padding on the measured data before the spatial DFT. Each of the periodic repetitions in the represented pressure will then include both the measurement area and the zero-padding, which will push the first replica of the measurement area further away from the real measurement area.

Having described all the basic considerations and operations on the data, we can formulate mathematically the DFT-based holography calculation. The following formulation is very similar to the one given in [3]. The input data we will assume for the formulation is the measured pressure data after multiplication with a possible smooth window function and after zero-padding. We assume an even number N_x and N_y of data points in the x- and y-directions, respectively, located at positions $(m_x a, m_y a)$, where $-N_x/2 \le m_x \le N_x/2-1$ and $-N_y/2 \le m_y \le N_y/2-1$. The corresponding data set in the wavenumber domain has the same number of data points at the positions $(n_x \Delta k_x, n_y \Delta k_y)$, where $-N_x/2 \le n_x \le N_x/2-1$ and $-N_y/2 \le n_y \le N_y/2-1$ and where

$$\Delta k_x = \frac{2\pi}{N_x a}, \quad \Delta k_y = \frac{2\pi}{N_y a}. \tag{13}$$

Writing the discrete pressure data set in the measurement plane as $p_{\mathrm{D}}(m_x, m_y)$ and the corresponding wavenumber data set as $P_{\mathrm{D}}(n_x, n_y)$, the forward DFT transform corresponding to Eq. (2) becomes

$$P_{\mathrm{D}}(n_x, n_y) = a^2 \sum_{m_x} \sum_{m_y} p_{\mathrm{D}}(m_x, m_y) \mathrm{e}^{2\pi \mathrm{j}\left[\frac{n_x m_x}{N_x} + \frac{n_y m_y}{N_y}\right]}. \tag{14}$$

To calculate the pressure wavenumber data set in a parallel plane with $z \neq 0$, we must multiply with the transfer function G of Eq. (5), and to minimize the effects of measurement errors we multiply with the regularization window function Π of Eq. (9):

$$P_{\mathrm{D}}(n_x, n_y, z) = P_{\mathrm{D}}(n_x, n_y) \cdot G(n_x \Delta k_x, n_y \Delta k_y, z) \cdot \Pi(n_x \Delta k_x, n_y \Delta k_y, k_{\mathrm{c}}), \tag{15}$$

where k_{c} is given in Eq. (10). The pressure data set in the calculation plane can now be obtained through inverse DFT:

$$p_{\mathrm{D}}(m_x, m_y, z) = \frac{1}{N_x N_y a^2} \sum_{n_x} \sum_{n_y} P_{\mathrm{D}}(n_x, n_y, z) \mathrm{e}^{-2\pi \mathrm{j}\left[\frac{n_x m_x}{N_x} + \frac{n_y m_y}{N_y}\right]} \tag{16}$$

The particle velocity in the same plane can be obtained by reusing Eqs. (15) and (16), just with a different transfer function G, which can be found from Eq. (8).

A lot of implementation details and considerations are given in [2,4].

1.6 Adaptations for Practical Application

So far we have considered only time-harmonic representation in the frequency domain. In practical measurement situations some effort is often needed to obtain such a representation of the sound field. As a first example, assume a rather large noise source radiating stationary, but only partially coherent noise. Partial coherence occurs because there are several independent (uncorrelated) excitation mechanisms inside the source. If the source is too large for a realistic microphone array to cover the necessary measurement area, then a smaller array must be scanned over that area. To obtain consistent phase information across the entire area, a reference signal is needed to provide a phase reference. But when averaging of typically cross-spectra is performed using the chosen reference signal, then only the part of the sound field coherent with that reference will be included in the averaged sound field representation. To obtain a complete representation we need a number of references at least equal to the number of independent source mechanisms, and from the measurement we then extract the same number of mutually incoherent time-harmonic sound fields that together represent the total sound field [9,10]. If the source is so small, however, that all positions can be measured simultaneously, then a time-harmonic representation can be obtained simply by doing a Fourier transform (using a full-record FFT) of the simultaneously recorded time data [11].

2 • Beamforming

Based on a planar measurement, NAH can reconstruct the near field of a sound source over a 3D region. By measuring a planar area that completely covers the sound source and by measuring with sufficient spatial sampling density (less than half a wavelength), all details of the sound field can basically be reconstructed, except for details that are lost in measurement errors. For large sources and/or at high frequencies, however, the required measurement area and sampling density make the method impractical. Here, beamforming is an attractive alternative based on a far-field array measurement: From typically a simultaneous recording with all array microphones the method determines, with a certain angular resolution, how

much is incident from all directions. So where NAH provides a calibrated map of selected sound field parameters near the source, beamforming basically provides a directional map of contributions at the array position. The present section will give only a short introduction to the very basic delay-and-sum beamformer. A broad description of beamforming techniques is given in [12].

2.1 Introduction to Delay-and-Sum Beamforming

As illustrated in Figure 4, we consider a planar array of M microphones at locations $\mathbf{r}_m (m = 1, 2, \ldots, M)$ in the xy-plane of our coordinate system. When such an array is applied for delay-and-sum beamforming, the measured pressure time signals $\tilde{p}_m$ are individually weighted and delayed, and then all signals are averaged [12]:

$$\tilde{b}(\boldsymbol{\kappa}, t) = \frac{1}{M} \sum_{m=1}^{M} w_m \tilde{p}_m (t - \Delta_m(\boldsymbol{\kappa})). \tag{17}$$

The individual time delays Δ_m are chosen with the aim of achieving selective directional sensitivity in a specific direction, characterized here by a unit vector $\boldsymbol{\kappa}$. This objective is met by adjusting the time delays in such a way that signals associated with a plane wave, incident from the direction $\boldsymbol{\kappa}$, will be aligned

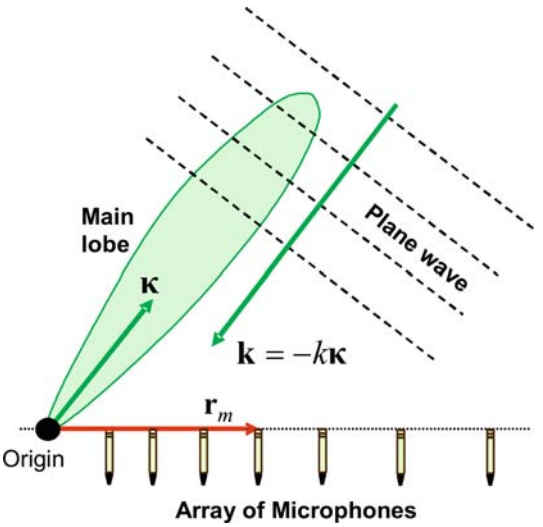


FIGURE 4 Illustration of a microphone array, a directional sensitivity represented by a mainlobe and a plane wave incident from the direction of the mainlobe.

in time before they are averaged. Geometrical considerations (see Figure 4) show that this can be obtained by choosing

$$\Delta_m = \frac{\boldsymbol{\kappa} \cdot \mathbf{r}_m}{c}. \tag{18}$$

Signals arriving from other far-field directions will not be aligned before the averaging, and therefore they will not add up coherently. Thus, we have obtained a directional sensitivity as illustrated by the *mainlobe* in Figure 4.

The frequency domain version of expression (17) for the delay-and-sum beamformer output is

$$b(\boldsymbol{\kappa}, \omega) = \frac{1}{M} \sum_{m=1}^{M} w_m \, p_m(\omega) \mathrm{e}^{-\mathrm{j}\omega\Delta_m(\boldsymbol{\kappa})} = \frac{1}{M} \sum_{m=1}^{M} w_m \, p_m(\omega) \mathrm{e}^{\mathrm{j}\mathbf{k}\cdot\mathbf{r}_m}, \tag{19}$$

where $\mathbf{k} \equiv -k\boldsymbol{\kappa}$ is the wavenumber vector of a fictitious plane wave incident from the direction $\boldsymbol{\kappa}$ in which the array is focused, see Figure 4.

Through our choice of time delays $\Delta_m(\boldsymbol{\kappa})$, or equivalently of the "preferred" wavenumber vector $\mathbf{k} \equiv -k\boldsymbol{\kappa}$, we have "tuned" the beamformer on the far-field direction $\boldsymbol{\kappa}$. Ideally we would like to measure only signals arriving from that direction, in order to get a perfect localization of the sound source. To investigate how much "leakage" we will get from plane waves incident from other directions, we assume now a plane wave incident with a wavenumber vector $\mathbf{k}_0$ different from the preferred $\mathbf{k} \equiv -k\boldsymbol{\kappa}$. The pressure measured by the microphones will then ideally be

$$p_m(\omega) = p_0 \mathrm{e}^{-\mathrm{j}\mathbf{k}_0\cdot\mathbf{r}_m}, \tag{20}$$

which according to Eq. (19) will give the following output from the beamformer:

$$b(\boldsymbol{\kappa}, \omega) = \frac{p_0}{M} \sum_{m=1}^{M} w_m \mathrm{e}^{\mathrm{j}(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{r}_m} \equiv p_0 W(\mathbf{k} - \mathbf{k}_0). \tag{21}$$

Here, the function W,

$$W(\mathbf{K}) \equiv \frac{1}{M} \sum_{m=1}^{M} w_m \mathrm{e}^{\mathrm{j}\mathbf{K}\cdot\mathbf{r}_m}, \tag{22}$$

is the so-called *array pattern*. It has the form of a 2D spatial Fourier transform of a weighting function w, which consists of delta functions at the microphone positions, refer Eq. (2). Because the microphone positions $\mathbf{r}_m$ have z-coordinate equal to zero, the array pattern is independent of K_z. We therefore

consider the array pattern W only in the (K_x, K_y) plane, and when it is used as in Eq. (21) the 3D wavenumber vector is projected onto the (K_x, K_y) plane. In that plane, W has an area with high values around the origin with a peak at $(K_x, K_y) = (0, 0)$. According to Eq. (21), this peak represents the high sensitivity to plane waves coming from the direction $\boldsymbol{\kappa}$, in which the array is focused. Figure 4 contains an illustration of that peak, which is called the *mainlobe*. Other directional peaks, which are called *sidelobes*, will cause waves from such directions to leak into the measurement of the mainlobe direction $\boldsymbol{\kappa}$, creating the so-called *ghost sources* or *ghost images*.

From Eqs. (21) and (22) it is clear that the array pattern contains information about the sidelobe pattern for all frequencies. When the frequency increases, the wavenumber vectors $\mathbf{k}$ become longer, meaning that a larger area of the array pattern function W becomes "visible". For practical application of an array it is important to have low sidelobe levels and therefore good suppression of ghost sources over the frequency range of interest. This leads to the idea of optimizing an array design by adjusting the geometry and perhaps the weights w_m with the aim of minimizing the sidelobes of the array pattern over the area that is visible within a given frequency range [13]. In general, letting the weights go to zero towards the edges of the array can decrease the sidelobe level. This is, however, at the expense of a broadening of the mainlobe, which means poorer angular resolution [12]. The weights are therefore often kept equal to one, which is referred to as *uniform shading*. Irregular geometries are typically used, because they can provide fairly low sidelobe level over very broad frequency bands, avoiding the so-called *grating lobes* associated with strong spatial aliasing effects of regular geometries. Low sidelobe levels (at least 10 dB lower than the mainlobe) can then be maintained up to frequencies where the average element spacing in the array is several wavelengths, typically 3–4 wavelengths.

So far we have considered only resolution of incoming plane waves, corresponding to an infinite focus distance. To focus the array at a point (at a finite distance), the delays in Eq. (18) have to be changed in such a way that signals associated with a spherical wave radiated from the focus point are aligned in time before the averaging. But such a spherical wave will exhibit varying amplitudes across the array, which introduces the need to do amplitude compensation [12,13]. Focusing at a finite distance works well in general and does not significantly increase the sidelobe level if the measurement distance is not smaller than array diameter.

2.2 Resolution and Size of Mapping Area

The size and shape of the mainlobe of the array pattern is determined almost entirely by the size of the array [12]. The microphones just need to have a fairly uniform distribution across the array area, not for example being concentrated at the centre. Clearly, the mainlobe width will determine the angular resolution of the array: A broad mainlobe has the effect that a single incident plane wave will contribute to the beamformed map over a wide angular range. Resolution is here defined in connection with two incoherent monopoles of equal strength on a distant plane parallel with the array. It is defined as the smallest separation that allows them to be distinguished in a beamformed map. For an array with uniform shading that covers a circular area with diameter D, the resolution can be shown to be approximately

$$R(\theta) \simeq \frac{1.22}{\cos^3(\theta)} \frac{L}{D} \lambda, \tag{23}$$

where L is the distance to the mapping plane and θ is the off-axis angle (from the z-axis) [13]. The resolution is seen to degrade very quickly beyond approximately 30° off-axis angle, so in practice mapping will be restricted to that angular range. Close to the axis, the best-case resolution is around one wavelength. This resolution is obtained when the array is used at the smallest recommended measurement distance, which is equal to the array diameter or slightly smaller. If this best possible resolution is required, then because of the 30° maximum off-axis angle the mapping area will be of approximately the same size as the array. As mentioned above, this area can be mapped with an array that has average microphone spacing up to 3–4 wavelengths. Compared to that, NAH requires less than half wavelength spacing. So to map a quadratic area of linear dimension one wavelength, NAH requires more than 4 measurement positions, whereas beamforming can do with down to 0.1 position.

The advantages of NAH are that scaled maps of any acoustic quantity are obtained, that the level of ghost sources can be kept considerably lower and that the resolution is half a wavelength or better, refer Eq. (11). If we look at low frequencies in particular, NAH can provide a resolution approximately equal to the typically very small measurement distance, whereas beamforming can still not do better than around one wavelength, which is very poor at low frequencies.

So at low frequencies, NAH is often required to get acceptable resolution, whereas at high frequencies beamforming is often required to map a given source area with an acceptable number of measurement positions. Reference [14] describes a solution where NAH and beamforming measurements can be performed

with the same irregular array. The array is then optimized to provide low sidelobe level with beamforming, and at the same time it has a sufficiently uniform element density that it applies at low frequencies with the SONAH holography algorithm presented in [8]. Reference [14] also describes a method to scale the beamformed maps as sound intensity, which is a typical type of output from NAH.

References

1. J.D. Maynard, E.G. Williams and Y. Lee, "Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH", *J. Acoust. Soc. Am.*, **78(4)**, 1395–1413 (1985).
2. W.A. Veronesi and J.D. Maynard, "Nearfield acoustic holography (NAH) II. Holographic reconstruction algorithms and computer implementation", *J. Acoust. Soc. Am.*, **81(5)**, 1307–1322 (1987).
3. E.G. Williams, *Fourier Acoustics – Sound Radiation and Nearfield Acoustical Holography*, Academic Press, (1999).
4. M.J. Crocker (editor), *Handbook of Noise and Vibration Control*, Chap. 48 by E.G. Wiliams, Wiley, New York (2005).
5. E.G. Williams, "Regularization methods for near-field acoustical holography", *J. Acoust. Soc. Am.*, **110(4)**, 1976–1988 (2001).
6. E.G. Williams, B.H. Houston and P.C. Herdic, "Fast Fourier transform and singular value decomposition formulations for patch nearfield acoustical holography", *J. Acoust. Soc. Am.*, **114(3)**, 1322–1333 (2003).
7. R. Steiner and J. Hald, "Near-field acoustical holography without the errors and limitations caused by the use of spatial DFT", *Proceedings of ICSV6* (1999).
8. J. Hald, "Patch near-field acoustical holography using a new statistically optimal method", *Proceedings of Inter-Noise* (2003).
9. J. Hald, "STSF – a unique technique for scan-based near-field acoustic holography without restrictions on coherence", *B&K Technical Review*, No. 1 (1989).
10. H.-S. Kwon, Y.-J. Kim and J.S. Bolton, "Compensation for source nonstationarity in multireference, scan-based nearfield acoustical holography", *J. Acoust. Soc. Am.*, **113(1)**, 360–368 (2003).
11. J. Hald, "Time domain acoustical holography and its applications", *Sound Vibr.*, 16–25 (February 2001).
12. D.H. Johnson and D.E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall, New Jersey (1993).
13. J.J. Christensen and J. Hald, "Beamforming", *B&K Technical Review*, No. 1 (2004).
14. J. Hald, "Combined NAH and Beamforming Using the Same Microphone Array", *Sound Vibr.*, 18–27 (December 2004).

PART II
AUDITORY SYSTEM AND HEARING

Birger Kollmeier

University of Oldenburg, Oldenburg, Germany

10 Anatomy, Physiology and Function of the Auditory System
Birger Kollmeier 147
External Ear • Middle Ear • Inner Ear • Auditory Pathway

11 Physiological Measures of Auditory Function
Birger Kollmeier, Helmut Riedel, Manfred Mauermann, and Stefan Uppenkamp 159
Head-Related Transfer Functions • Tympanometry • Acoustic Middle Ear Reflex • Otoacoustic Emissions • Evoked Potentials • Functional Imaging of the Auditory System

12 Auditory Processing Models
Torsten Dau 175
External and Middle Ear Transformation • Cochlear Transformation • Examples of Brainstem Processing Models • Models of Auditory Perception • Applications of Auditory Models

13 Speech Intelligibility
Thomas Brand 197
Speech Intelligibility • Measurement Methods • Factors Influencing Speech Intelligibility • Prediction Methods

14 Signal Processing in Hearing Aids

Volker Hohmann 205

Basic Approach to the Rehabilitation of Hearing Deficiencies • Factors of Hearing Impairment and Rehabilitation Strategies • Limitations of Current Rehabilitation Strategies • Conclusions

10

Anatomy, Physiology and Function of the Auditory System

Birger Kollmeier

University of Oldenburg, Oldenburg, Germany

The human ear consists of the outer ear (pinna or concha, outer ear canal, tympanic membrane), the middle ear (middle ear cavity with the three ossicles malleus, incus and stapes) and the inner ear (cochlea which is connected to the three semicircular canals by the vestibule, which provides the sense of balance). The cochlea is connected to the brain stem via the eighth brain nerve, i.e. the vestibular cochlear nerve or nervus statoacusticus. Subsequently, the acoustical information is processed by the brain at various levels of the auditory system. An overview about the anatomy of the auditory system is provided by Figure 1.

1 • External Ear

The pinna and the initial part of the external auditory canal are constructed from elastic cartilage and show a high interindividual variability. The more internal part of the external ear canal is surrounded by bone. In its rest position, the external ear canal is a bit winded so that the ear drum can only be inspected from outside if this winding is removed (e.g. by softly pulling the pinna to the back and upward). This is utilized when performing *otoscopy*: The permeability of the external auditory

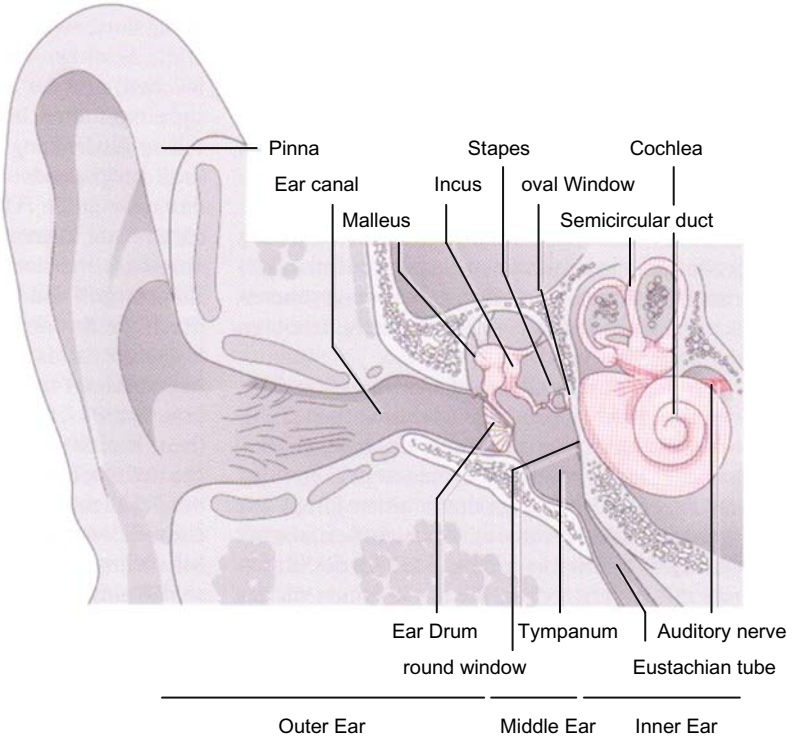


FIGURE 1 Overview of the anatomy of the outer ear, middle ear and inner ear (modified from Kollmeier, 1997).

canal, the shape and consistency of the ear drum, and any suspicious changes in the visible structures are considered by an inspection using a magnifying glass or a microscope. The function of the pinna is to sample the incident sound wave from a larger area into the smaller area of the ear canal (funnel principle, especially for high frequencies). Also, a spectral change (filter) is provided as a function of the incidence direction. The resulting change in timbre as a function of incidental direction can be utilized for localization of the sound source from where the sound is emitted. (Additionally, interaural differences in level and in arrival time occur that are utilized by the brain to perform a very exact localization in the horizontal plane by comparing the input to both ears. However, the interaural information has to be supplemented by the spectral changes provided by the pinna and other structures to avoid front/back confusions.) At low frequencies, the shape of the upper part of the body and the head is primarily responsible for

TABLE 1 Functional organization and possible pathologies of the auditory system as well as the range of audiological diagnostical methods to cover the function of the respective components

Part of the ear	Function	Dysfunction	Audiological diagnostics
Outer ear	Directional filtering	Malformation of pinna	Medical examination
Middle ear	Impedance transformation Air→Water	Conductive hearing loss (e.g., hydrotympanon, otitis media)	Otoscopy Impedance audiometry
Inner ear	Transformation water borne sound → neural excitation, frequency-place transformation	Cochlear hearing loss (e.g., noise-induced, age-related,….)	Tuning fork tests Tone audiogram Otoacoustic emissions
Auditory nerve/ brainstem	Coding of acoustical information, evaluation of interaural differences, modulation processing	Retrocochlear (neural) hearing loss, deterioration of localization	Suprathreshold tests Brain stem audiometry Speech audiometry
Cortex	Speech perception, complex perception	Central hearing disorders (e.g. aphasia	Medical imaging Central speech tests

the exact direction-dependent filtering. The exact shape of the pinna is more relevant at high frequencies where the wavelength has the same order of magnitude as the dimensions of the pinna. These properties are utilized in the dummy-head technique which is termed "virtual acoustics" as soon as a computer gets involved. Basically, the recording can be made either with a dummy head or with a probe microphone in the individual's ear canal. If the sound is reproduced (for instance by equalized headphones) to the same subject, he or she has the impression to be acoustically located in the original recording site.

Most pathological changes of the outer ear and its function can be treated by otolaryngologists (i.e. physicians specialized in ear, nose and throat diseases). These changes can be observed either clinically (i.e. by the previous history and a physical examination) or by otoscopy (Table 1).

2 • Middle Ear

The middle ear is located in the air-filled tympanum which is connected via the Eustachian tube to the nose/throat cavity (Figure 1). This narrow tube can actively be opened by muscle tension when swallowing or yawning. The opening produces a pressure release on both sides of the tube which provides

the same atmospheric pressure both from outside and inside the tympanic membrane. The maintenance of an equilibrium is important for the free movement of the tympanic membrane as well as the successive middle ear ossicles. A strong, static displacement of the tympanic membrane (e.g. by changes of the air pressure) can also be compensated by the joint between malleus head and incus that fixates itself into a new position (Hüttenbrink, 1988; Hudde and Engel, 1997).

The function of the middle ear is the impedance adjustment between the sound propagation in air (low impedance) and the quite high input impedance of the liquid-filled inner ear (high impedance). This is primarily achieved by the ratio between the large area of the tympanic membrane and the small area of the stapes footplate attached to the oval window. An additional contribution arises from the leverage of the long processus of the malleus connected to the ear drum and the shorter processus of the incus connected to the stapes. The total transformation effect produces an approximate increase of the force per unit of area by a factor of 50 at the stapes footplate. Even though the resulting impedance matching is not perfect (a factor of 3,000 would be required for a perfect impedance match between ear and water), it is sufficient to provide an optimum transmission efficiency with a minimum amount of biological effort (since larger membranes for a better impedance match would cause a higher susceptibility to possible damages).

At very high sound pressure levels, the stapedius muscle comes into action which is attached to the head of the stapes and produces a change of middle ear transmission efficiency that can be measured by acoustical impedance measurements (see below). The action of the second middle ear muscle, i.e. the tensor tympani cannot be recorded in audiological measurements.

A pathological change of the middle ear function results in a conductive hearing loss, i.e. an attenuation of an air-borne sound into the inner ear. This may be caused by insufficient ventilation through the Eustachian tube as well as fluid or infections in the middle ear cavity. Even though the type of the pathological change can be very different (i.e. otosclerosis with a fixation of the stapes footplate to the surrounding bone, disruption of the ossicular chain, defect in the tympanic membrane, etc.), the resulting dysfunction is very similar.

For a functional diagnostics of the middle ear, impedance audiometry is utilized. The acoustical reflection at the tympanic membrane is assessed with a microphone in the sealed ear canal in order to estimate the impedance and the sound transmission into the inner ear. Tympanometry tests the static air pressure

delivered to the ear canal that yields the best transition (i.e. maximum acoustical compliance of the ear drum). If a middle ear pathology exists, the optimum static pressure value is displaced or no maximum value is found at all (Chapter 11). From the temporal change of the impedance after stimulating the ear with a high-intensity tone, the presence or the absence of the stapedius muscle reflex can be determined (reflex audiometry). Note that a conductive hearing loss is still connected to an intact bone conduction into the inner ear. This difference between air and bone conduction ("air–bone gap") can be assessed in clinical tuning fork tests (according to Weber and Rinne) and in audiometry with a bone conduction stimulator in comparison to thresholds obtained with an audiometric headphone.

3 • Inner Ear

The inner ear consists of a coiled shell-shaped tube which is embedded in the very hard temporal bone and consists of three compartments (Figure 2). The oval window connects to the scala vestibuli which is connected at the upper end of the cochlea, the so-called helicotrema, with the scala tympani. The intermediate compartment (scala media) is filled with endolymph and is separated from the scala tympani by the basilar membrane (Figure 2). The lateral width of the basilar membrane increases steadily from the oval window to the helicotrema and simultaneously decreases in its stiffness. If a sound excites the oval window, a pressure difference occurs orthogonal to the basilar membrane between the scala vestibuli and the scala tympani which produces a movement of the basilar membrane. Even though the water-borne sound travels with high velocity through all compartments of the cochlea, the location-dependent sound pressure difference across the basilar membrane leads to a peculiar wave pattern on the basilar membrane, the so-called travelling wave. It propagates with a comparatively slow velocity and a very high dispersion along the basilar membrane from stapes toward the helicotrema and displays maximum amplitude for high frequencies close to the stapes and for low frequencies close to the helicotrema (frequency-place transformation). Many theories have been developed in the past to explain this pattern, including a linear theory (de Boer, 1980) and three-dimensional finite element models (Böhnke and Arnold, 1999).

The displacement of the basilar membrane produces a shear movement in relation to the tectorial membrane which is detected by the inner and outer hair cells embedded on top of the basilar membrane (Figure 3). The stereocilia at the end of the hair cells

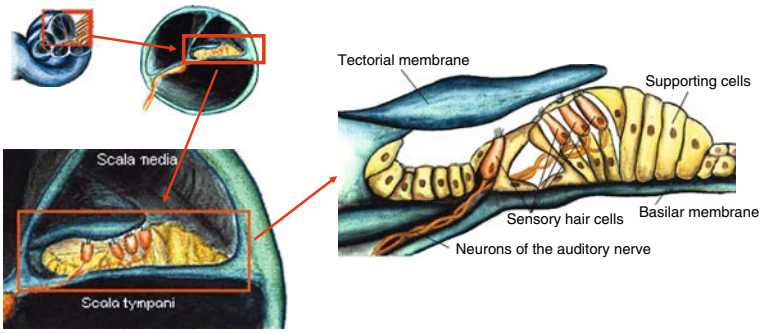


FIGURE 2 Schematical cross-section of the inner ear (modified after Silbernagl and Despopoulos, 2003).

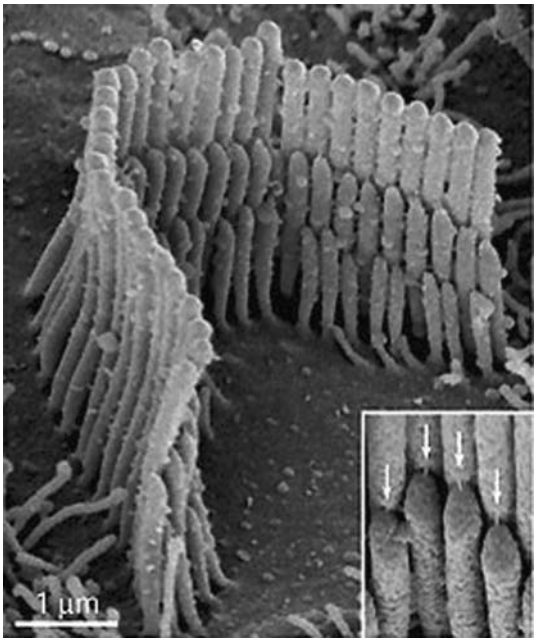


FIGURE 3 Scanning microscope photography of an intact hair cell bundle (stereocilia) located atop a hair cell. Tip links (*arrows*) connect shorter stereocilia to their taller neighbours. Reproduced with permission from Nature Reviews Genetics 5, 489–498, copyright 2004 Macmillan Magazines Ltd.

are connected with each other with tiny links ("tip links") that produce a change in ion permeability, if a mechanical stress is applied (Hudspeth et al., 1998, Dallos et al., 1991). Hence, the hair cells detect a shear movement by a change in their intracellular

potential. While the inner hair cells are each connected to a large number of afferent nerve fibres (that transfer excitation from the receptors to the brain), the three rows of outer hair cells are primarily connected to efferent fibres (i.e. slowly changing adaptation information from the brain to the hair cells). The outer hair cells are capable of actively contracting themselves if an external voltage is applied. This is used to amplify vibrations at low acoustical input levels, so that an active feedback loop simultaneously increases the sensitivity and the frequency specificity of the basilar membrane response. Even though the exact mechanisms of these "active processes" in the inner ear is not yet fully understood (see Dallos et al., 1997; Gummer et al., 2002 for a review), they have a prominent importance for the normal auditory function especially when listening to low acoustical levels.

If the feedback from mechanical/acoustical energy via electrical voltage into the contraction of outer hair cells (which in turn is mechanical/acoustical energy) assumes a loop gain of greater than 1, the system becomes instable and begins to oscillate. These low-energetic oscillations are actually recordable with a sensitive microphone in the sealed ear canal when employing sophisticated noise reduction methods. They are called spontaneous *otoacoustic emissions* (SOAE). Such emissions can also be evoked by a short, transient stimulus (transient-evoked otoacoustic emissions, TEOAE) or with a continuous tone (simultaneous-evoked otoacoustic emissions). They can be applied to objectively test the auditory system (Table 1).

A damage to the inner ear causes a sensorineural hearing loss that may have various reasons (excessive noise exposure, age-related detriment of the hair cells, metabolic diseases, etc.). In most cases, the function of both the inner and the outer hair cells is decreased to a variable degree. The consequence is a reduced sensitivity of the ear for stimuli close to the respective frequency region. If primarily the inner hair cells are damaged, the active processes are still intact that still provide a kind of dynamic compression (i.e. amplification of low levels that saturates at high acoustic levels), while the overall sensitivity is reduced. Conversely, a partial or complete loss of outer hair cells produces both a loss in sensitivity and a loss in compression, i.e. no special adaptation occurs to amplify soft sounds. Consequently, a steep increase of perceived loudness can be observed when the stimulus level is increased (the so-called recruitment phenomenon, see Chapter 14 (Hohmann)).

The function of the inner ear can be tested with a tuning fork or with the audiogram using both bone-conduction transducers in comparison to measuring the sensitivity to air-borne sound with headphones. The presence of otoacoustical emissions

is an indication of a (nearly) normal auditory system, while the threshold test and speech audiometry can be used to assess suprathreshold distortions of the auditory perception due to a hearing loss (i.e. recruitment phenomenon or decreased speech intelligibility in noise). As an "objective" auditory test which tests both the function of the inner ear and parts of the brain, the brain stem audiometry (brain stem-evoked response audiometry, BERA) is available for clinical diagnostics. With this method, the electrical voltage at the surface of the skull is recorded as response to an acoustical stimulus delivered to the patient's ear. To separate the stimulus-related reaction of the brain from the uncorrelated, random noise produced by the brain, an averaging method is used. If the recorded response for a given auditory stimulus complies with the normal value, one can objectively certify a normal functioning auditory system.

4 • Auditory Pathway

The auditory nerve originates in the middle of the cochlea and runs through the meatus acusticus internus (inner ear duct) to the brain stem where it reaches the cochlear nucleus. From this brain nucleus, several links exist to other nuclei in the brain stem (i.e. superior olive, nucleus accessories and to the nuclei of the lateral lemniscus). The auditory pathway goes up through the lateral lemniscus and the medial lemniscus into the inferior colliculus and the medial geniculate body until it finally reaches the primary auditory cortex which is located in area 41of the temporal lobe in the cortex (Figure 4). Several linkages exist at each of these stages to the respective other side.

The function of the auditory nerve and the auditory pathway is the coding and processing of all acoustical information into neural excitation patterns in the appropriate structures that represent auditory information to the brain. The auditory nerve codes acoustical information by the synchronicity in the spiking pattern of different neural fibres as well as increasing the rate of the respective fibres with increasing stimulus level. In the brain stem, already complex auditory features are evaluated. For example, an interaural comparison to detect interaural arrival time and intensity differences is already performed in the superior olive complex as a basis for localization of sound sources (see above). In addition, a modulation frequency analysis is already performed in the inferior colliculus (Langner and Schreiner, 1988). The common modulation of the intensity in different frequency bands seems to be an important feature for

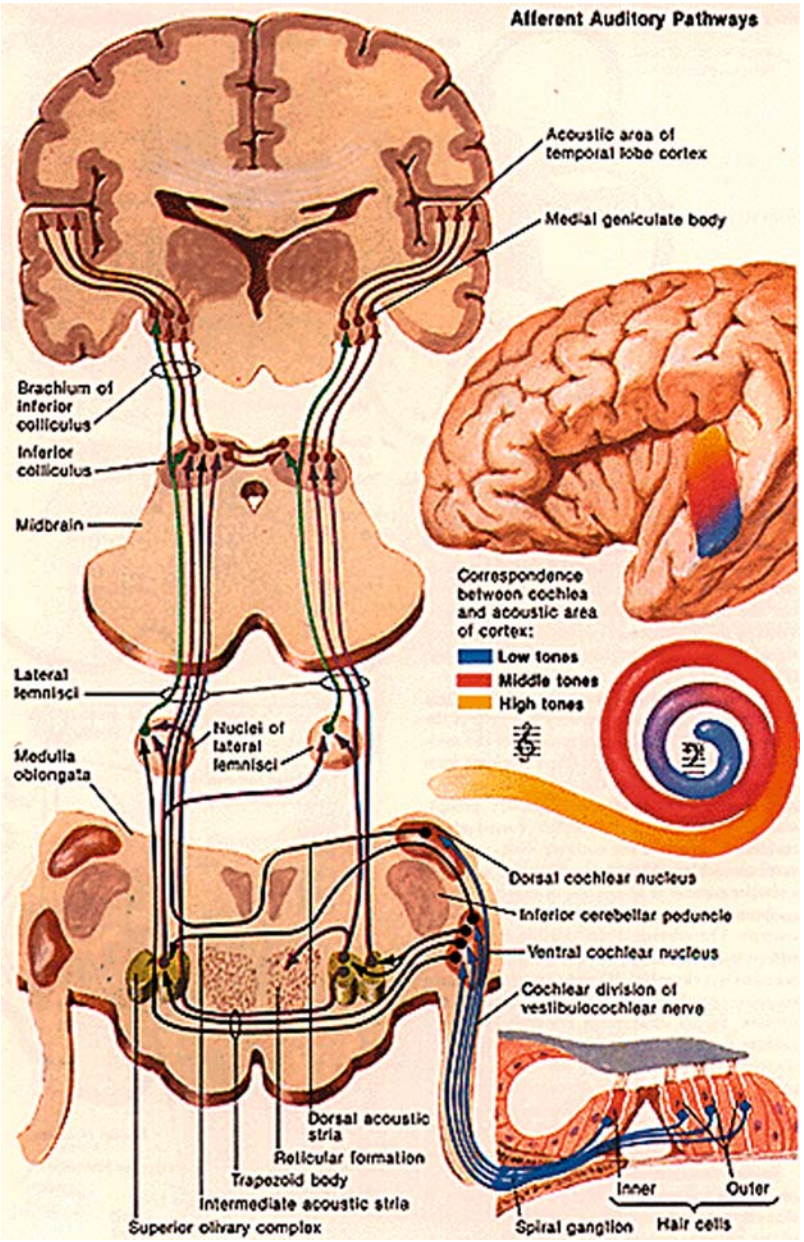


FIGURE 4 Schematical view of the central auditory system with the anatomical view (reprinted from the CIBA Collection of Medical Illustrations, Vol. 1).

an auditory object, so that a modulation analysis is assumed to take place already at a comparatively early processing stage of the auditory system (see Chapter 12 (Dau)). Furthermore, the whole auditory pathway is characterized by its tonotopic organization, i.e. neighbouring acoustical frequencies yield excitations in neighbouring structures within the whole auditory pathway. A similar spatial organization is supposed for the spatial organization of sound sources in space (spatiotopic organization) and a for a regular ordering of the different modulation frequencies in the brain (periodotopic mapping).

Pathological changes of the functions of the auditory nerves and the more peripheral parts of the central auditory system (e.g. acusticus neurinoma that mechanically damages the auditory nerve, or hearing deficits connected with other neurological symptoms as a consequence of a stroke or other neurological disease) are denoted as retrocochlear or neural hearing disorders. They can be detected with brain stem audiometry where a prolonged delay time is observed between the electrical waves originating from the cochlea (i.e. waves I and II) and the more centrally originating parts of the evoked response audiometry (ERA). Alternatively, imaging techniques like functional MR imaging or positron emission tomography (PET) can be used. More centrally located damages to the auditory system (for example due to blood circulation deficits or lesions) can result in various neurological symptoms, among which the aphasic syndromes are most prominent (i.e. the inability to speak correctly). Several central language tests have been developed and applied in order to improve the diagnostic properties here.

In a physical or communication theory sense, signal processing in the ear can be described by a succession of several signal processing stages that model the "effective" processing in the auditory system without taking the actual biological implementation into account. An important element of such a functional description is a filter bank as a first stage that distributes the sound according to its frequency into different frequency bands and hence mimics some aspects of the basilar membrane (Figure 5). It may be preceded by a broad band-pass filter that mimics the "effective" frequency shaping by the middle ear. In each frequency channel, the instantaneous energy is then obtained by extracting the temporal envelope which can be modelled by a half-wave rectifier with successive low-pass filtering. This process and the subsequent adaptation stage should simulate the function of the hair cell and the auditory nerve where the respective sensitivity is actually adapted to

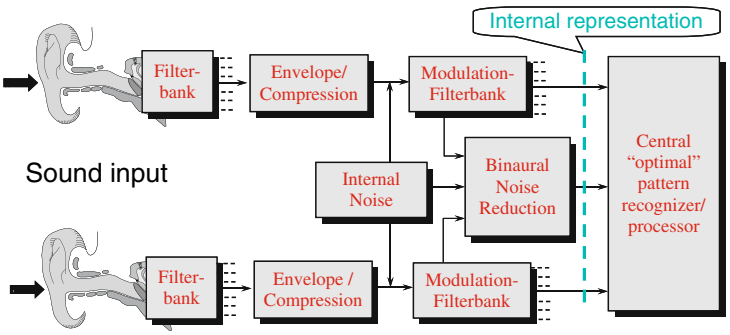


FIGURE 5 Schematical view of the "effective" signal processing in the auditory system.

the average of the respective input signal, and the exact fine structure of the input signal is lost for high-signal frequencies. The adaptation is followed by a binaural comparison as well as a separation into different modulation frequencies by a modulation filter bank. Hence, the output of these basic processing structures can be thought of as a two-dimensional pattern for each point in time that represents audio frequency on one axis and modulation frequency on the other axis. In addition, for each combination of modulation frequency and centre frequency, an interaural relation can be assumed to be present (Kollmeier and Koch, 1994).

The time course of this multidimensional pattern can be thought of as representing the "internal representation" of an acoustical stimulus presented to the auditory system. To account for processing inaccuracies of the auditory system, an additive processing noise ("internal noise") has to be added. Hence, most models of the "effective" signal processing in the auditory system assume that these most important processing steps of the auditory system can be represented with comparatively simple processing circuits that require very few parameters for their functioning. The actual detection of any signals in noise or of any changes in the acoustic input signal can be modeled then by an optimum detector or matched filter (Green and Swets, 1974; Dau et al., 1997). In other words, the current models assume that auditory detection and discrimination is not necessarily limited by the cognitive information processing and the previous knowledge about the stimulus and the training in the auditory task, but rather by any inaccuracies in representing the external acoustical stimulus in the internal representation. A more detailed review is provided in Chapter 12. Even though the current status of such models of the signal processing in the

auditory system does not allow a very detailed description of the auditory system, it appears that a technical oriented way of interpreting the action of the auditory system seems to be very promising in order to "understand" the auditory system better and to derive applications from that (Kollmeier, 2003).

References

Böhnke F, Arnold W (1999) 3D-finite element model of the human cochlea including fluid-structure couplings. ORL 61, 305–310.

Dallos P, Evans BN, Hallworth R (1991) Nature of the motor element in electrokinetic shape changes of cochlear outer hair cells. Nature 350, 155–157.

Dallos P, He DZZ, Lin X, Sziklai I, Mehta S, Evans BN (1997) Acetylcholine, outer hair cell electromotility, and the cochlear amplifier. J Neurosci 17, 2212–2226.

Dau T, Kollmeier B, Kohlrausch A (1997) Modeling auditory processing of amplitude modulation: II. Spectral and temporal integration. J Acoust Soc Am 102, 2906–2919.

de Boer E (1980) Auditory physics. Physical principles in hearing theory I. Phys Rep 62, 87–174.

Green DM, Swets JA (1974) Signal detection theory and psychophysics, 2nd edition. Krieger, New York.

Gummer AW, Meyer J, Frank G, Scherer MP, Preyer S (2002) Mechanical transduction in outer hair cells. Audiol Neuro-Otol 7, 13–16.

Hudde H, Engel A (1997) Measuring and modeling basic properties of the human middle ear and ear canal. Part III: Eardrum impedances, transfer funtions and model calculations. Acustica/Acta Acustica 84, 1091–1108.

Hudspeth AJ, Choe Y, Mehta AD, Martin P (1998) Putting ion channels to work: Mechanoelectrical transduction, adaptation, and amplification by hair cells. Proc Nat Acad Sci USA 97, 11765–11772.

Hüttenbrink KB (1988) Die Mechanik der Gehörknöchelchen bei statischen Drucken, I. Normales Mittelohr Laryng Rhinol Otol 67, 45–52.

Kollmeier B (1997) Grundlagen. In: Kiessling J, Kollmeier B, Diller G, Versorgung und Rehabilitation mit Hörgeräten. Thieme, Stuttgart, 1–47.

Kollmeier B (2003) Auditory principles in speech processing: Do computers need silicon ears? In: Proceedings of the 8th European Conference on Speech Communication and Technology (Eurospeech), ICSA International Speech Communication Association, pp. 5–8.

Kollmeier B, Koch R (1994) Speech enhancement based on physiological and psychoacoustical models of modulation perception and binaural interaction. J Acoust Soc Am 95, 1593–1602.

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. 1. Neuronal mechanisms. J Neurophysiol 60, 1799–1822.

Silbernagl S, Despopoulos A (2003) Taschenatlas der Physiologie, Thieme, Stuttgart.

11

Physiological Measures of Auditory Function

Birger Kollmeier, Helmut Riedel, Manfred Mauermann, and Stefan Uppenkamp

University of Oldenburg, Oldenburg, Germany

When acoustic signals enter the ears, they pass several processing stages of various complexities before they will be perceived. The auditory pathway can be separated into structures dealing with sound transmission in air (i.e. the outer ear, ear canal, and the vibration of tympanic membrane), structures dealing with the transformation of sound pressure waves into mechanical vibrations of the inner ear fluids (i.e. the tympanic membrane, ossicular chain, and the oval window), structures carrying mechanical vibrations in the fluid-filled inner ear (i.e. the cochlea with basilar membrane, tectorial membrane, and hair cells), structures that transform mechanical oscillations into a neural code, and finally several stages of neural processing in the brain along the pathway from the brainstem to the cortex. It is possible to investigate the function at nearly all of these stages by using methods tailored to exploit the particular physiological representation of interest. This chapter gives a brief overview of the main physiological measures that are currently in use to investigate the auditory system.

1 • Head-Related Transfer Functions

The main function of the outer ear and ear canal is to collect sound from the outside world and transmit it to the tympanic membrane. The shape of pinnae shows very large variability between subjects. An individual pinna can be characterized be its head-related transfer function (HRTF). This is generated from the Fourier transform of the impulse response from a sound played away from the listener and recorded at a point close to the tympanic membrane. The shape of the HRTF determines a "coloration" of the power spectrum of any perceived sound. It is mainly dependent on the angle between ear canal axis and the location of the external sound source. A certain amount of directional information is hence acoustically coded and can be centrally decoded at the output of the transduction mechanisms in the auditory system. A convolution of an acoustic signal with pre-recorded HRTFs can be used to simulate spatial listening conditions.

2 • Tympanometry

Functional diagnostics of the middle ear can be performed using a technique known as tympanometry that assesses the acoustic reflection at the tympanic membrane. The required apparatus to perform clinical tympanometry includes (1) a probe with a miniature loudspeaker (or tube connected to an insert earphone), (2) a precision measurement microphone, and (3) a connection to an external air pump. This probe is inserted in the ear canal using a sealing tip to allow pressurization of the ear canal. Using a probe tone typically of 220 Hz delivered to the loudspeaker, the sound pressure level measured with the microphone is kept constant as the static air pressure in the ear canal is varied. From the loudspeaker level that is needed to obtain the constant predefined level, the acoustical compliance of the middle ear can be obtained, i.e. the imaginary part of the inverse of the impedance. A review of models describing the middle ear input impedance, as a function of several structural parameters of the middle ear (including pathological cases), is provided by Hudde and Engel (1998). The compliance is expressed in millilitres of an equivalent air cushion. It usually exhibits a steep maximum if the static air pressure in the ear canal equals the static pressure in the tympanic cavity. This indicates the best sound transmission from the external ear to the ear drum and subsequent structures. If a middle ear pathology exists, the optimum static pressure is displaced or no maximum value is found at all. An overview of

the effect of different middle ear pathologies on tympanometric results for clinical usage is given by Böhme and Welzl-Müller (1998) and Lehnhardt and Laszig (2000).

3 • Acoustic Middle Ear Reflex

The recording and interpretation of the acoustic middle ear reflex (AMER) is a powerful diagnostic technique in clinical audiology (overview in Møller, 2000). The contraction of the stapedius muscle in response to intense sounds decreases the sound transmission to the inner ear. The reflex is elicited in both ears, even with monaural stimulation, although it is slightly stronger in the ipsilateral ear where the threshold is also decreased. The absence of an acoustic reflex can be caused by conductive or sensorineural hearing loss, by disorders in the auditory nerve, or by disorders in the motor neurons in the facial nerve. The most important parameter recorded is the acoustic reflex threshold (ART). This is defined as the lowest stimulus intensity at which a contraction of the stapedius muscle can be detected (Clemis, 1984). The AMER is highly sensitive to middle ear pathology and can help to assess sensorineural hearing impairment. In most cases, however, a sensorineural hearing loss up to 60 dB does not alter the ART. The reflex threshold therefore appears to be more linked to the uncomfortable level (UCL) rather than absolute threshold.

In clinical usage, the acoustic reflex is most commonly detected as a change in the middle ear compliance by means of a low-frequency probe tone (Metz, 1951). A second short tone pulse presented at a high level elicits the reflex. The resulting decrease of acoustic compliance increases the level of the low probe tone, when measured with a microphone in the closed ear canal. When recorded in this way, the ART ranges between 80 and 100 dB hearing level for pure-tone stimulation in normal-hearing listeners. An alternative method to record the AMER was described by Neumann et al. (1996). They used a pair of successive short tone bursts rather than the combination of a single stimulus with the low-frequency probe tone. The method relies on the short latency, or time constant, of the acoustic reflex of the order of 25–100 ms. A change in the compliance of the middle ear results in a slight difference in the measured amplitudes for each of the two tone bursts. This can easily be detected by calculating the difference signal. The threshold criterion is based on the detection of phase coherence between successive measurements. With this method, the ART appears to be substantially lower than previously reported (Neumann et al., 1996). Acoustic reflex thresholds have also been linked to perceived

loudness (Gorga et al., 1980; Kawase et al., 1998), which could substantially increase their clinical potential, e.g. aiding the fit of hearing aids based on physiological rather than psychophysical measures. This is currently still a controversial issue (Margolis and Popelka, 1975; Olsen and Rasmussen, 1999a, b).

4 • Otoacoustic Emissions

Otoacoustic emissions are acoustical signals produced in the cochlea, which can be recorded in the ear canal using a sensitive microphone and some noise cancellation techniques. They can be recorded from nearly all normal-hearing human listeners, and also from many different animal models, including nearly all vertebrates (Manley and Taschenberger, 1993). Otoacoustic emissions (OAE) are a consequence of the non-linear and active preprocessing of sound in the cochlea. People with a cochlear hearing loss in general show reduced or no emissions. This implies that OAE have the potential to act as an objective test of hearing function. The measurement of OAE is essentially the only non-invasive method to investigate physiological signals from the cochlea.

Different OAE types are classified according to the stimulus type and related details of the measurement paradigm. In this terminology, spontaneous OAE, i.e. narrowband signals emitted by about 40–60 % of normal-hearing people at a very low level without any external stimulus, are separated from evoked OAE. These are responses from the cochlea to various types of acoustic stimulation.

In most cases, spontaneous OAE are very weak, with sound pressure levels of less than 5 dB SPL in an occluded ear canal. Usually they are recorded in the frequency domain by averaging several power spectra of the acoustic noise in the closed ear canal. So far, spontaneous OAE have little diagnostic value, as for the majority of people, neither the presence nor the absence of SOAE is linked to any audiological problem. Spontaneous OAE are simply a consequence of active processes in the inner ear. Many properties are in line with the idea of active, non-linear oscillators (e.g. van der Pol oscillator, van Dijk and Wit, 1990; Talmadge et al., 1991).

There are several types of evoked OAE. Transient-evoked emissions are weak sound emissions in response to a short, transient stimulus, like a click, a tone burst, or a chirp (see Figure 1a). In this case, stimulus and emission can be separated in time (Kemp, 1978). Click-evoked OAE are recorded with latencies between 3 and about 30 ms depending on the frequency

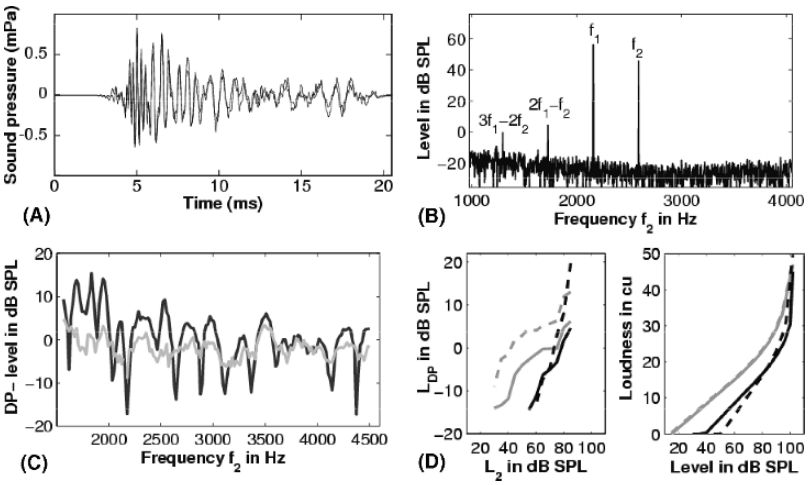


FIGURE 1 (**a**) Click-evoked otoacoustic emission for a normal-hearing listener. The two traces represent averaging in two separate buffers to allow for subsequent cross-correlation. (**b**) Spectrum of a single DPOAE measurement: Note the distortion products at $2f_1 - f_2$ and $3f_1 - 2f_2$ in addition to the primary stimuli at f_1 and f_2. (**c**) High-resolution DP-Gram with and without additional suppressor tone near $2f_1 - f_2$. Black line: DPOAE levels of the $2f_1 - f_2$ distortion product for a fixed frequency ratio f_2/f_1 of 1.2; f_2 was varied in steps of 18 Hz. Grey line: DPOAE levels measured with the same paradigm, but with additional presentation of a suppressor tone 12 Hz below $2f_1 - f_2$ at a level of 50 dB SPL. (**d**) Comparison of DPOAE I/O functions (*left panel*) and loudness growth (*right panel*) from categorical loudness scaling. The grey lines show the results from two normal-hearing subjects; the black lines in both panels show the respective results from two hearing-impaired subjects. Note the similarity of changes in both, indicating the loss of compression in hearing-impaired subjects.

of the transient stimulus. With broadband stimulation, nearly all normal-hearing listeners emit a broadband response. Click-evoked emissions are now widely used for screening of infant hearing, since they can sufficiently differentiate between populations of normal hearing and hearing-impaired subjects. However, the transition between normal hearing and hearing impairment is not very clear, with a large variability across subjects. A quantitative prediction of hearing thresholds for individual listeners based on click-evoked OAE is difficult.

Simultaneous-evoked OAE are emissions in response to stationary external stimulation, typically sinusoids (Kemp, 1979a; Wilson, 1980). One very important type is the distortion product

OAE (DPOAE), which allows a separation of stimulus and response in the frequency domain. DPOAE can be recorded as one or even several combination tones when stimulating with two sinusoids at different frequencies (Kemp, 1979b; Anderson and Kemp, 1979). The most prominent component is the distortion product at the frequency $2f_1 - f_2$, when the stimuli are sinusoids at f_1 and f_2 with an appropriate frequency ratio f_2/f_1 (see Figure 1b). A detailed review of the different types of OAE is given by Probst et al. (1991).

OAE classification related to the different stimulus and measurement paradigms should be clearly distinguished from a classification according to the underlying mechanisms generating the emissions in the inner ear (OAE source types). In general, two types of OAE sources are identified in the literature: (1) the so-called place-fixed or reflection emissions and (2) the wave-fixed or distortion emissions (see, e.g. Kemp, 2002; Shera and Guinan, 1999). Nearly "pure" reflection emissions can be recorded by the measurement of SFOAE or TEOAE at low levels. They are characterized by a strongly rotating phase of the emission across frequencies, i.e. a high group delay. These emissions in primates are assumed to be generated by a reflection, either at periodically distributed spatial irregularities across the cochlear partition (Strube, 1989) or as the coherent interference of wavelets scattered by randomly distributed perturbations located throughout the tall and broad peak of the travelling wave (Zweig and Shera, 1995). The latter is more consistent with physiological findings. Even spontaneous OAE in mammals can be explained as the result of multiple reflections between regions of coherent reflection and the oval window within the cochlea, resulting in amplitude-stabilized standing waves (Shera, 2003).

A non-linear transmission line model of the cochlea, allowing coherent reflection, can explain several properties seen in OAE, like the minimum spacing between SOAE, the fine structure in TEOAE and SFOAE, and also the fine structure observed in hearing thresholds and DPOAE (e.g. Talmadge et al., (1998)). Apical DPOAE ($2f_1 - f_2, 3f_1 - 2f_2, \ldots$, as can be seen in Figure 1b) measured in the sealed ear canal with a primary frequency ratio f_2/f_1 of about 1.2 are considered to be a mixture of two OAE sources with different underlying mechanisms. The first is a distortion source being in the region of maximal interaction of the two primaries close to the characteristic place of f_2. The second is a coherent reflection source at the characteristic place of the distortion product frequency itself at, e.g. $2f_1 - f_2$, excited by the travelling wave initiated by the distortion source. The

distortion source can be extracted from DPOAE measurements either by using suppression and presenting a third tone close to the reflection site, typically $2f_1 - f_2$ (e.g. Heitmann et al., 1998), or by an offline analysis called spectral smoothing (Kalluri and Shera, 2001) or latency windowing (Knight and Kemp, 2001; Mauermann and Kollmeier, 2004). Figure 1c shows one example of a DPOAE fine structure. This appears as a quasiperiodic variation of DPOAE level across frequencies of up to 15 dB, with a spacing between adjacent maxima of about 0.4 bark (see black line in the figure). The fine structure mirrors the interaction of the reflection source with its rotating phase and the distortion source, characterized by an almost constant phase across frequencies. The grey line in Figure 1c gives the almost separated contribution from the distortion source from the same subject. This was obtained by DPOAE measurements with an additional suppressor tone of 50 dB SPL at 15 Hz below the respective $2f_1 - f_2$ frequencies. The pronounced spectral DPOAE fine structure related to the interaction of the two sources with their different phase characteristics is strongly reduced (compare black and grey lines in Figure 1c).

Input/output (I/O) functions for DPOAE, (i.e. the DP level as a function of the level of the primaries) appear to be a useful tool to assess hearing status (Figure 1d). The right panel of Figure 1d shows the results from categorical loudness scaling for two normal-hearing (grey lines) and two hearing-impaired listeners (black lines), indicating a clear recruitment in the hearing-impaired subjects. The psychoacoustically observed loss of compression ("recruitment") is clearly reflected by the I/O functions from DPOAE measurements. The data shown in Figure 1d were obtained by varying the primary level L_1 for a given primary level L_2 to get an optimized distortion product level. Furthermore, a suppressor tone was presented close to $2f_1 - f_2$ to reduce the influence of the reflection source. The source separation appears to be essential for the usage of I/O functions from DPOAE measurements since DPOAE I/O functions are strongly influenced by the interaction of the two DPOAE sources (Mauermann and Kollmeier, 2004).

At present, most OAE recordings described in the literature are a mixture of the different sources, and clinically they mainly serve as an infant or neonatal screening tool. However, since the underlying OAE mechanisms are most probably differently affected by different cochlear pathologies, a clear separation of the contribution from different OAE sources and mechanisms may in future lead to a more detailed insight into cochlear status based on OAE measurements.

5 • Evoked Potentials

The neural activity along the auditory pathway and in the cortex can be observed by the recording of several types of far-field-evoked potentials. This is done using electrodes fixed to the head of the listeners, and utilizing various methods of noise cancellation during data acquisition, particularly averaging a large number of single recordings. When stimulated with a transient like a click, a characteristic pattern of several peaks and valleys is visible in the evoked potential, with latencies between 1 and more than 500 ms (Jewett and Williston, 1971). These potentials are classified according to their latency relative to stimulus onset, auditory brainstem responses (ABR) with latencies between 1 and 10 ms, middle latency auditory evoked potentials (MAEP) with latencies between 10 and approximately 50 ms, and finally late auditory evoked potentials (LAEP) with latencies of 100 ms and above. The most important type of far-field potential for clinical applications is the recording of ABR, usually summarized as brainstem evoked response audiometry (BERA).

The ABR is generated by the synchronized neural activity of many nerve fibres along the auditory pathway. The recorded electrical potential at the scalp represents the superposition of the activity of a very large number of electric dipole sources along all the stages of the pathway from the spiral ganglion up to the inferior colliculus (IC). Nevertheless, it is still possible to describe the peaks of the far-field potential by equivalent electric dipole sources, which can be linked to particular anatomical structures. For example, wave I of the ABR, at a latency of less than 1 ms, is closely related to the compound action potential which can also be recorded from the auditory nerve. Wave V of the ABR, with latencies varying between 5 and 8 ms as a function of stimulus level, is the most robust peak and can be recorded down to the absolute threshold of hearing. This part of the response appears to be generated by structures at the input to the IC (Scherg, 1991).

Electrophysiological responses to clicks are dominated by neural activity in the high-frequency channels. This is because high-frequency channels are synchronized, whereas the low-frequency channels are not, due to dispersion along the basilar membrane. One way to synchronize excitation in all frequency channels for ABR recording was described by Dau et al. (2000). They used optimized chirp stimuli with increasing instantaneous frequency to compensate for dispersion and demonstrated that these chirp signals enhance ABR wave V. This is probably due to increased synchronization of the low-frequency channels. An example of ABR recordings in response to a click and a synchronizing chirp stimulus is given in Figure 2. Two main effects can be observed. First, the amplitude of wave V is considerably increased

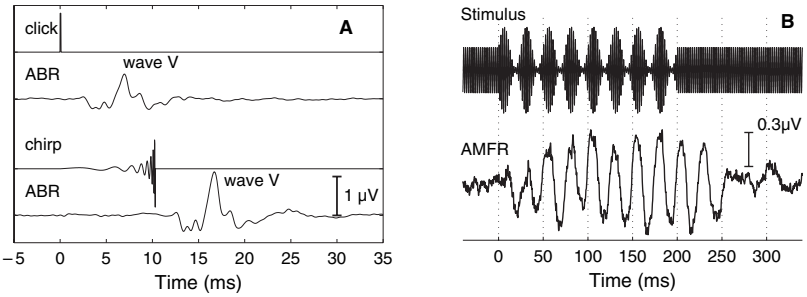


FIGURE 2 (**a**) Auditory brainstem responses to a click stimulus (*top*) and to a chirp stimulus (*bottom*) with increasing instantaneous frequency, designed to compensate for basilar membrane dispersion. The biggest peak, ABR wave V, has a latency of approximately 6 ms relative to stimulus offset and is significantly enhanced during chirp stimulation due to greater synchronization along the cochlear partition. Stimulus level: 40 dB nHL, diotic stimulation, recording bandwidth 100–1500 Hz, 10,000 averages. (**b**) Amplitude modulation following response. Stimulus: 60 dB SL, carrier frequency at 480 Hz, sinusoidally modulated at 40 Hz, monaural stimulation, recording bandwidth 15–1500 Hz. An evoked potential reflecting the amplitude modulation is clearly visible.

for the chirp in comparison to the click. Second, the position of wave V relative to the stimulus onset is delayed by an additional 10.5 ms for the chirp, while the delay relative to stimulus offset is almost identical to the click condition. This illustrates that the intended phase alignment between different BM channels is preserved at the level of IC. In a similar experiment with magnetoencephalography (MEG, i.e., the recording of the very weak magnetic field associated with brain activity), Rupp et al. (2002) have shown that this synchronization is even preserved in the middle latency response generated at the level of primary auditory cortex.

ABR responses and middle latency responses are dominated by cochlear filtering, compression, and temporal smoothing. Late auditory evoked potentials, however, are less stimulus specific. Also at the long latencies associated with cortical activity beyond primary auditory cortex, the response becomes more and more process specific. One famous example is the mismatch negativity response which arises from a perceptual contrast between an individual, deviant sound presented occasionally in a background series of standards that establish a context for the contrast (Näätänen et al., 1978). This response can be recorded reliably with both EEG and MEG.

Another important branch of evoked potentials are auditory steady-state responses (ASSR), such as frequency following

responses (FFR), amplitude modulation following responses (AMFR) and sustained potentials or magnetic fields (Picton et al., 1978, 2003; Hall, 1979). All of these responses reflect one particular perceptual property of a stationary stimulus, like for example pitch, or modulation frequency. For ASSR, the stimulus is usually presented continuously to avoid onset effects. One example of an AMFR is given in the right panel of Figure 2. For a continuous 480 Hz tone, an amplitude modulation is switched on for eight full cycles. The response is dominated by a sinusoidal component at 40 Hz, corresponding to the modulation frequency of the stimulus. This 40 Hz response is assumed to be of cortical origin. It can be recorded at stimulus levels down to the subjective threshold of hearing for all audiometric carrier frequencies, which makes it potentially useful as an objective, frequency-specific audiometric tool.

Both EEG and MEG provide a temporal resolution of one millisecond or less and allow to study very fine details of temporal processing in the auditory pathway. However, the spatial resolution is inferior to the high resolution of modern neuroimaging methods like functional magnetic resonance imaging (see next section). In future, a combination of EEG and MEG with these methods therefore should substantially increase our knowledge of how the auditory brain is working.

6 • Functional Imaging of the Auditory System

The main neuroimaging methods to visualize the function of the brain are positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). With PET, radioactive positron emitting tracer materials are injected into the subject, and the distribution of radiation from the brain tissue is observed, using a ring of gamma coincidence detectors. Tomographic 3D reconstruction of the distribution of activity within the brain is calculated by means of Fourier transformations (Cho et al., 1993), similar to computerized X-ray tomography. PET has been used for a big variety of auditory experiments, ranging from pitch perception to spoken language (overview in Johnsrude et al., 2002). PET has now been overtaken by functional MRI in terms of the number of studies performed, due to the wider distribution of MRI scanners and to the finer spatial and temporal resolution that can be achieved with modern MRI scanners. However, PET still has one crucial advantage, especially for auditory experiments, that is the absence of any loud distracting noise produced by the scanner.

Functional magnetic resonance imaging (fMRI) allows the exploration of neural activity in the brain at a spatial resolution

of a few millimetres and a temporal resolution of several seconds. The combination of functional and anatomical images, acquired in the same listener, enables a direct relation of physiological processes to the underlying anatomical structures in the brain. In recent years, a great deal of progress has been made in utilizing this technique for auditory investigations.

The principle of MRI is based on radio frequency excitation and relaxation of nuclear magnetic moments that have been aligned with a strong external magnetic field. Functional MRI is usually based on the narrow magnetic resonance of proton spins. Spatial encoding is achieved using additional gradient coils that restrict RF excitation to predefined slices through the tissue. Phase and frequency encoding during RF excitation and measurement of the NMR signal are utilized for coding the other two spatial dimensions. The spatial resolution and signal-to-noise ratio are directly linked to the strength of the external magnetic field. For imaging human subjects, field strengths from 0.5 to 7 T are used.

There are several problems with fMRI for auditory experiments, for example, switching of the gradient coils in the strong external magnetic field results in very loud noises. This is due to mechanical vibrations caused by Lorentz forces acting on the coils. Precautions have to be introduced; the first one is to avoid damage to the subject's ears, and the second is to ensure that the recorded brain activity is caused by the actual experimental stimulus, and not by the scanner noise itself. There are now MR compatible, high-fidelity headphones available that have been fitted into conventional ear defenders. These give an attenuation of external noise between 15 and 40 dB. Additional damping of scanner noise can be achieved by sound absorbing material around the subject's head and also inside the walls of the scanner bore. In addition to these passive measures, active noise cancellation (ANC) has been utilized to compensate for the scanner noise (Goldmann et al., 1989; Chen et al., 1999; Chambers et al., 2001). By means of ANC, an additional reduction of the perceived scanner noise of 10–15 dB has currently been achieved.

With fMRI, a secondary, metabolic response is observed rather than the primary neural activation. The actual signal change of the measured NMR signal is based on the change of the oxygenation level of blood due to metabolic processes that are triggered by the neural activity. These changes result in a slight change of the relaxation times in activated regions of the brain and are the basis for an image contrast between different regions (overview in Buxton, 2002). As a consequence, the observed signal change has a comparatively long latency of several seconds relative to stimulus onset. This lack of immediate responses can

be turned into an advantage for auditory fMRI by introduction of the paradigm of "sparse temporal sampling". It allows for a separation of the activation due to scanner noise from the one due to the actual stimulus of interest. Instead of continuous data acquisition during epochs of stimulus and epochs of rest, stimulation and data acquisition are separated in time, with periods of several seconds of stimulus presentation without any scanning (Edmister et al., 1999; Hall et al., 1999). A combination of sparse temporal sampling and cardiac gating, that is synchronization of the MR image acquisition with the cardiac cycle, has been shown to allow for imaging of not only the auditory cortex, but even the brainstem and thalamic structures with fMRI (Griffiths et al., 2001).

Since the observed NMR signal changes in activated regions of the brain are very small (in the order of 1 %), the analysis of fMRI data is heavily based on repeated measures and statistical models that test for significant differences between conditions (Worsley et al., 1992). A widespread approach is the method of statistical parametric mapping based on the theory of Gaussian random fields (Frackowiak et al., 2004; http://www.fil.ion.ucl.ac.uk/spm).

Figure 3 gives one example for the activation of auditory cortex in response to acoustic stimulation using sparse temporal sampling and statistical parametric mapping for data analysis.

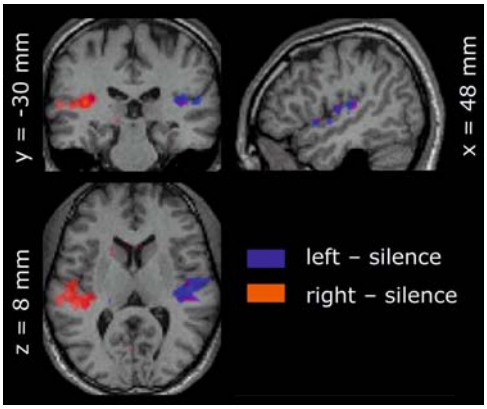


FIGURE 3 Activation of Heschl's gyrus for unilateral presentation of a broadband sound. The activation maps represent the *T*-statistics for significant differences between sound presentation and silence (i.e. no acoustic stimulus). Stimulation of the right ear largely activates the left hemisphere and vice versa. Activation maps are superimposed on an anatomical image of the same listener, in coronal (*top left*), saggital (top right), and axial view (*bottom left*). The coordinates are relative to the origin in commissura anterior.

The map shows significantly activated voxels in response to unilaterally presented broadband sound either to the left ear (in blue) or to the right ear (in red), when contrasted with the baseline condition (no sound presentation). The activation in these "sound–silence" contrasts is largely restricted to Heschl's gyrus in both hemispheres, and the crossing of the major projections on the way from the ears to cortex is clearly visible.

References

Anderson SD, Kemp DT (1979) The evoked cochlear mechanical response in laboratory primates: a preliminary report. Arch Otorhinolaryngol 224, 47–54.

Böhme G, Welzl-Müller K (1998) Audiometrie. Huber, Göttingen.

Buxton RB (2002) Introduction to Functional Magnetic Resonance Imaging. Cambridge University Press, Cambridge.

Chambers J, Akeroyd MA, Summerfield AQ, Palmer AR (2001) Active control of the volume acquisition noise in functional magnetic resonance imaging: method and psychoacoustical evaluation. J Acoust Soc Am 110, 3041–3054.

Chen CK, Chiueh TD, Chen JH (1999) Active cancellation system of acoustic noise in MR imaging. IEEE Transact Biomed Eng 46, 186–191.

Cho ZH, Jones JP, Singh M (1993) Foundations of Medical Imaging. Wiley, New York.

Clemis, JD (1984) Acoustic reflex testing in otoneurology. Otolaryngol Head Neck Surg 92, 141–144.

Dau T, Wegner O, Mellert V, Kollmeier B (2000) Auditory brainstem responses (ABR) with optimized chirp signals compensating basilar-membrane dispersion. J Acoust Soc Am 107, 1530–1540.

van Dijk P, Wit HP (1990) Amplitude and frequency fluctuation of spontaneous otoacoustic emissions. J Acoust Soc Am 88, 1779–1793.

Edmister WB, Talavage TM, Ledden PJ, Weisskoff RM (1999) Improved auditory cortex imaging using clustered volume acquisitions. Hum Brain Mapp 7, 89–97.

Frackowiak RSJ, Friston KJ, Frith CD, Dolan RJ, Price CJ, Zeki S, Ashburner J, Penny W (2004) Human brain function, 2nd edition. Elsevier Academic Press, London.

Goldmann AM, Gossmann WE, Friedlander PC (1989) Reduction of sound levels with antinoise in MR imaging. Radiology 173, 519–550.

Gorga MP, Lilly DJ, Lenth RV (1980) Effect of signal bandwidth upon threshold of the acoustic reflex and upon loudness. Audiology 19, 277–292.

Griffiths TD, Uppenkamp S, Johnsrude I, Josephs O, Patterson RD (2001) Encoding of the temporal regularity of sound in the human brainstem. Nat Neurosci 4, 633–637.

Hall JW (1979) Auditory brainstem frequency following responses to waveform envelope periodicity. Science 205, 1297–1299.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp **7**, 213–223.

Heitmann J, Waldmann B, Schnitzler HU, Plinkert PK, Zenner HP (1998) Suppression of distortion product otoacoustic emissions (DPOAE) near $2f_1 - f_2$ removes DP-gram fine structure – evidence for a secondary generator. J Acoust Soc Am 103, 1527–1531.

Hudde H, Engel A (1998). Acoustomechanical human middle ear properties. Part III: eardrum impedances, transfer functions, and model calculations. ACUSTICA – Acta Acustica 84, 1091–1108.

Jewett DL, Williston JS (1971) Auditory-evoked far-fields averaged from the scalp of humans. Brain 94, 681–696.

Johnsrude IS, Giraud AL, Frackowiak RSJ (2002) Functional imaging of the auditory system: the use of positron emission tomography. Audiol Neurootol 7, 251–276.

Kalluri R, Shera CA (2001) Distortion-product source unmixing: a test of the two-mechanism model for DPOAE generation. J Acoust Soc Am 109, 622–637.

Kawase T, Hidaka H, Ikeda K, Hashimoto S, Takasaka T (1998) Acoustic reflex thresholds and loudness in patients with unilateral hearing losses. Eur Arch Otorhinolaryngol 255, 7–11.

Kemp DT (1978) Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 64, 1386–1391.

Kemp DT (1979a) Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. Arch Otorhinolaryngol 224, 37–45.

Kemp DT (1979b) The evoked cochlear mechanical response and the auditory microstructure – evidence for a new element in cochlear mechanics. Scand Audiol Suppl 9, 35–47.

Kemp DT (2002) Exploring cochlear status with otoacoustic emissions. In: Otoacoustic Emissions – Clinical Applications, 2nd ed., edited by MS Robinette, TJ Glattke, Thieme, New York, pp. 1–47.

Knight RD, Kemp DT (2001) Wave and place fixed DPOAE maps of the human ear. J Acoust Soc Am 109, 1513–1525.

Lehnhardt E, Laszig R (2000) Praxis der Audiometrie. Thieme, Stuttgart.

Manley G, Taschenberger G (1993) Spontaneous otoacoustic emissions from a bird: a preliminary report. In: Biophysics of Hair Cell Sensory Systems, edited by H Duifhuis, JW Horst, P van Dijk, SM van Netten. World Scientific, Singapore, pp. 33–39.

Margolis RH, Popelka GR (1975) Loudness and the acoustic reflex. J Acoust Soc Am 58, 1330–1332.

Mauermann M, Kollmeier B (2004) Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source. J Acoust Soc Am 116, 2199–2212.

Metz O (1951) Studies of the contraction of the tympanic muscles as indicated by changes in the impedance of the ear. Laryngoscope LXVIII(I), 48–62.

Møller AR (2000) Hearing – Its Physiology and Pathophysiology, Chapter 12. Academic Press, San Diego.

Näätänen R, Gaillard AWK, Mäntysalo S (1978) Early selective-attention effect on evoked potential reinterpreted. Acta Psychol 42, 313–329.
Neumann J, Uppenkamp S, Kollmeier B (1996) Detection of the acoustic reflex below 80 dB HL. Audiol Neurootol 1, 359–369.
Olsen SO, Rasmussen AN, Nielsen LH, Borgkvist BV (1999a) The acoustic reflex threshold: not predictive for loudness perception in normally-hearing listeners. Audiology 38, 303–307.
Olsen SO, Rasmussen AN, Nielsen LH, Borgkvist BV (1999b) The relationship between the acoustic reflex threshold and levels of loudness categories in hearing impaired listeners. Audiology 38, 308–311.
Picton TW, Woods DL, Proulx GB (1978) Human auditory sustained potentials. I. The nature of the response. Electroencephalogr Clin Neurophysiol 45, 186–197.
Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses. Int J Audiol 42, 177–219.
Probst R, Lonsbury-Martin BL, Martin GK (1991) A review of otoacoustic emissions. J Acoust Soc Am 89, 2027–2067
Rupp A, Uppenkamp S, Gutschalk A, Beucker R, Patterson RD, Dau T, Scherg M (2002) The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans. Hear Res 174, 19–31.
Scherg M (1991) Akustisch evozierte Potentiale. Kohlhammer, Stuttgart.
Shera CA (2003) Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves. J Acoust Soc Am 114, 244–262.
Shera CA, Guinan Jr. JJ (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am 105, 782–798.
Strube HW (1989) Evoked otoacoustic emissions as cochlear Bragg reflections. Hear Res 38, 35–45.
Talmadge CL, Tubis A, Wit HP, Long GR (1991) Are spontaneous otoacoustic emissions generated by self-sustained cochlear oscillators? J Acoust Soc Am 89, 2391–2399.
Talmadge CL, Tubis A, Long GR, Piskorski P (1998) Modeling otoacoustic emission and hearing threshold fine structure. J Acoust Soc Am 104, 1517–1543.
Wilson JP (1980) Evidence for a cochlear origin for acoustic reemissions, threshold fine-structure and tonal tinnitus. Hear Res 2, 233–252.
Worsley KJ, Evans AC, Marrett S, Neelin P (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. J Cereb Blood Flow Metab 12, 900–918.
Zweig G, Shera CA (1995) The origin of periodicity in the spectrum of evoked otoacoustic emissions. J Acoust Soc Am 98, 2018–2047.

12

Auditory Processing Models

Torsten Dau

Technical University of Denmark, Lyngby, Denmark,
e-mail: tda@elektro.dtu.dk

There are at least two main reasons why auditory processing models are constructed: to represent the results from a variety of experiments within one framework and to explain the functioning of the system. Specifically, processing models help generate hypotheses that can be explicitly stated and quantitatively tested for complex systems. The models can also help determine how a deficit in one or more components affects the overall operation of the system. The development of auditory models has been hampered by the complexity of the individual auditory processing stages and their interactions. This resulted in a multiplicity of auditory models described in the literature. The models differ in their degree of complexity and quantification. Models of auditory processing may be roughly classified into biophysical, physiological, mathematical (or statistical) and perceptual models depending on which aspects of processing are considered. Most of the models can be broadly referred to as functional models, that is, they simulate some experimentally observed input–output behavior of the auditory system without explicitly modeling the precise internal physical mechanisms involved. In the following, several models of auditory signal processing are summarized, including descriptions of the signal processing in the auditory periphery, examples of neural processing in the brainstem and models of auditory perception that often contain some decision

stage after the assumed preprocessing. Some of the models can be useful for technical and clinical applications, such as improved man–machine communication by employing auditory-model-based processing techniques, or new processing strategies in digital hearing aids and cochlear implants.

1 • External and Middle Ear Transformation

The external ear is commonly subdivided into the external ear diffraction system and the pinna-ear canal resonator system. The outer ear enhances the sound pressure at the eardrum for frequencies in the range 1–9 kHz, with a maximum enhancement at 3 kHz of 10–15 dB. The transfer function of the external ear system is highly dependent on the direction of sound incidence (Shaw, 1980). Sound diffraction of the external ear system has been modeled as electrical network by Killion and Clemis (1981) and Giguère and Woodland (1994) using electroacoustic analogies. The sound propagation in the ear canal has been modeled as a transmission line by Gardner and Hawley (1972). The middle ear transfers airborne sound in the ear canal into fluid motion in the cochlea. The transmission in the middle ear is most efficient for midrange frequencies, with a maximum at 1 kHz, and drops off markedly for very low and very high frequencies. Zwislocki (1962) developed the first electroacoustic network model of the middle ear based on anatomical considerations and on input impedance data measured at the eardrum. The network was later adapted by Lutman and Martin (1979) to account for the action of the stapedius muscle. The latter model was later modified by Giguère and Woodland (1994) which allowed them to directly connect their cochlear network to the middle ear. An alternative approach has been provided by Moore et al. (1997) where the transfer of sound through the external and middle ear has been modeled effectively using fixed filters. In a recent implementation of the model (Glasberg and Moore, 2002), the combined effect of the outer and middle ear has been modeled by a single finite impulse response (FIR). The output was considered as representing the sound reaching the cochlea in a model of loudness of time-varying sounds.

2 • Cochlear Transformation

2.1 Basilar Membrane Processing

Stapes vibrations within the oval window set the cochlear fluid into motion. The response of the basilar membrane (BM) is a traveling wave propagating from the base to the apex of the cochlea. Models of the BM have been formulated in many ways. They range from mechanical structures, via electrical networks

consisting of inductors, resistances, capacitors, diodes and amplifiers, to abstract structures that are put into a mathematical form. In biophysical models of BM processing, the network is typically realized as a transmission-line model, whereby the BM is spatially discretized into a series of segments of certain length. The transmission-line network elements are derived from electroacoustic analogies and from the assumption that the natural frequency of the resonant circuit in each segment is equal to the characteristic frequency of the BM at that place. These "classical" point impedance models of the cochlea are reviewed in, e.g., de Boer (1980, 1984, 1991), Giguère and Woodland (1994) and Hubbard and Mountain (1996). One of the characteristics of these (biophysical) models is that their impulse responses reflect not only the local mechanisms at the point of measurement but depend on the mechanics of the entire cochlea (e.g., Zweig, 1991; Neely, 1983; de Boer, 1995; Olson, 1999; Shera, 2001). This is fundamentally different from those models that treat the cochlea essentially as a bank of parallel bandpass filters. These "functional" models do not account for cochlear hydrodynamics; they simulate the response of the BM at a single site and the bandpass filters are instead suggested to originate via the differential build-up or decay of micromechanical resonances (e.g., Patterson et al., 1987; Irino and Patterson, 1997, 2000, 2001; Shekhter and Carney, 1997).

There has been evidence that the BM motion is highly nonlinear and is a major source of level compression (e.g., Rhode, 1971; Ruggero and Rich, 1991). Several functional modeling studies have explored combinations of linear filters and (memoryless) nonlinearities that provide a phenomenological representation of BM responses. Initial models of this type included the "bandpass nonlinearity" models of Pfeiffer (1970) and Duifhuis (1976). Goldstein (1990, 1995), Meddis et al. (2001), Lopez-Poveda and Meddis (2001) and Summer et al. (2003) extended their approach with a multiple bandpass nonlinear model, which included two interacting paths. One path in this dual resonance nonlinear (DRNL) filter configuration contains a lowpass filter followed by a memoryless nonlinearity and the other path contains a bandpass filter. The output of the system is the sum of the outputs of the linear and the nonlinear paths. The latest implementations of the model have been tested in large detail against a wide range of physical observations.

2.2 Transformation into Neural Activity

A strong correlation between the normal function of the outer hair cells and nonlinear BM responses has been demonstrated by Ruggero and Rich (1991). The outer hair cells have been shown to participate in the process of cochlear amplification

as mechanical force generators acting on the cochlear partition (Kim, 1986). The inner hair cells (IHC) transduce the motion of the organ of Corti into firing activity in the auditory nerve (AN). Models of AN responses have been developed studying the representation of sounds at the first level of neural coding in the auditory system. Sachs and Abbas (1974) and Schoonhoven et al. (1997) investigated phenomenological models that explain rate-level functions for AN fibers in terms of the relation between AN threshold and BM compression. Also, detailed models of the IHC–AN synapse (e.g., Meddis, 1986, 1988; Geisler, 1990; Hewitt and Meddis, 1991) provide descriptions of several features of AN rate-level functions. Probably the most detailed descriptions of AN activity are given in the computational AN models of Heinz et al. (2001a,b) and Zhang et al. (2001) which are a modification of an earlier model by Carney (1993). The models describe many of the main response properties associated with the cochlear amplifier, including broadened tuning with increases in level, the associated compressive-magnitude and nonlinear phase responses, as well as two-tone suppression.

3 • Examples of Brainstem Processing Models

The activity of auditory brainstem and central neurons has not been studied in the same detail as the activity in the periphery. This is also reflected in a much smaller number of computational models of brainstem and central auditory processing. An exception might be the modeling of binaural processing and models of coding of amplitude modulation.

3.1 Neural Binaural Models

Several models of binaural interaction in brainstem neurons have been proposed. Most models are based on the coincidence counter hypothesis following an internal delay line as suggested by Jeffress (1948). The physiological basis for such coincidence counters are the excitation–excitation (EE) type cells (e.g., Goldberg and Brown, 1969; Joris and Yin, 1995) that are found in the superior olivary complex (SOC), representing the first neurons in the ascending auditory pathway exhibiting binaural interaction. Their discharge rate in response to binaural stimulation depends on the interaural time difference (ITD) and, at favorable ITDs, exceeds the sum of the responses for either ear alone (Goldberg and Brown, 1969). In models based on an array of EE-type cells with a range of "characteristic delays", the neural discharge rate resulting from the EE interaction is usually modeled as an interaural cross-correlation function (e.g., Jeffress, 1948; Shamma et al., 1989; Colburn et al., 1990). Within these

models, an ITD results in a shift of the cross-correlation function along the interaural delay axis and hence leads to a predicted lateralization.

There has also been physiological evidence for excitation–inhibition type cells; corresponding EI-cells in the SOC and in the inferior colliculus (IC) are excited by the signals from one ear and inhibited by the signals from the other ear (e.g., Kuwada et al., 1984; Joris and Yin, 1995; Palmer et al., 1997; McAlpine et al., 1998). For neurons in the IC the exciting and inhibiting channels are typically reversed (IE-type cells). These observations have led to models consistent with the equalization cancellation (EC) theory by Durlach (1963, 1972). It has been shown in corresponding neural models that both ITD and IID sensitivities may be understood by considering the outcome of a subtractive mechanism for EI/IE-type neurons (e.g., Colburn and Moss, 1981; Zacksenhouse et al., 1992). A detailed review of the physiological binaural models can be found in Colburn (1996).

3.2 Neural Models of Amplitude-Modulation Processing

Another example has been the processing of amplitude modulations (AM). Envelope fluctuations are a common physical feature produced by many natural sound sources. While AM information is encoded in temporal activity patterns of neurons at the early stages of the auditory system, it is generally believed that this temporal periodicity code is transformed into a rate-based code at some higher stage of auditory processing. It has been suggested that arrays of neurons in the brainstem, each tuned to a different modulation rate, can be considered as a filter in the modulation domain, and the array of neurons is known collectively as a modulation filterbank. Neurons with appropriate properties have been found in the cochlear nucleus (Møller, 1976) and the inferior colliculus (Rees and Møller, 1983; Langner and Schreiner, 1988; Lorenzi et al., 1995). Detailed reviews can be found in Palmer (1995) and Joris et al. (2004).

The mechanisms by which the auditory system extracts, encodes and processes AM information are not yet fully understood. Several physiologically motivated models have been suggested, assuming different mechanisms of how temporal AM information can be transformed into a rate-based representation between CN and IC. The neural circuit suggested by Langner (1981) (see also Langner, 1997; Langer et al., 1997) consists of three functional units at the model CN stage, with each of these units being associated with a specific neural response type in the CN. At the model IC stage, a coincidence detector unit is located that represents the model output unit. In the model by Hewitt and Meddis (1994), the CN model stage employs populations

of simulated chopper units that are morphologically associated with stellate cells in the ventral cochlear nucleus (VCN; Rhode et al., 1983; Smith and Rhode, 1989). The model by Hewitt and Meddis (1994) therefore transforms the bandpass-shaped temporal modulation transfer functions (tMTFs) observed in VCN chopper units into bandpass-shaped rate functions (rMTFs). The different best modulation frequencies (BMFs) of different IC coincidence detector units hereby result from different chopping frequencies among different populations of chopper units.

A third model describing the formation of bandpass-shaped rMTFs in IC neurons was suggested by Nelson and Carney (2004). The CN stage of their model employs a VCN bushy cell that receives undelayed excitatory AN input and delayed inhibitory AN input. The model bushy cell projects onto the model IC neuron in the same manner, i.e., it provides undelayed excitatory input and delayed inhibitory input. The interaction of excitatory and inhibitory synaptic inputs results in a bandpass-shaped rMTF at the model IC neuron if the excitatory synapses are assumed to be faster and weaker than the inhibitory synapses. Different BMFs within different IC model neurons are obtained by varying the time courses of the synaptic inputs.

4 • Models of Auditory Perception

Some perception models address the processing of the acoustical stimuli without much attention to underlying physiological processes (black box); others are inspired by neurophysiological findings. Typically, in the latter case, rather than trying to model many physiological details, the approach is to focus on the "effective" signal processing which uses as little physiological and physical parameters as necessary, but tries to predict as many perceptual data as possible. On the one hand, such a modeling strategy will never allow conclusions about the details of signal processing at a neuronal or single-unit level. On the other hand, if the effective model correctly describes the transformation of physical parameters in a large variety of experimental conditions, this strongly suggests certain general processing principles. These, in turn, may motivate the further study search for neural circuits in corresponding physiological studies.

4.1 Physiological Models for Basic Auditory Percepts

Explaining basic auditory perceptual phenomena in terms of physiological mechanisms has a long tradition. A nice review of this modeling approach is given in Delgutte (1996). There have been early systematic attempts at predicting psychophysical performance from the activity of AN fibers (e.g., Siebert, 1965, 1970). The same strategy has been used in a series of later

modeling studies (e.g., Delgutte, 1990; Heinz et al., 2001a,b; Colburn et al., 2003) combining analytical and computational population models of the AN with statistical decision theory to evaluate performance limits in tasks like intensity discrimination and masking. A general result is that models incorporating a realistic cochlear frequency map, cochlear filters, suppression, and rate-level functions are reasonably successful to account for performance in basic detection, discrimination and masking tasks. One of the most recent implementations of this type of model (Heinz et al., 2001a,b, see Figure 1) quantifies the effects of the cochlear amplifier on temporal and average-rate based information given in the AN and evaluates the influences of compression and suppression on psychophysical measures of auditory frequency selectivity. In principle, the approach can also be used for more complex stimuli and tasks, although the increased importance of more central factors might limit the value of models based on AN activity. Figure 1 should be placed before Chapter 4.2 (It belongs to 4.1).

4.2 Binaural Perception Models

Most models of binaural perception are based on the overlapping concepts of interaural coincidence networks and interaural correlation functions applied to individual frequency bands. These models are all fundamentally similar, and the concept of an

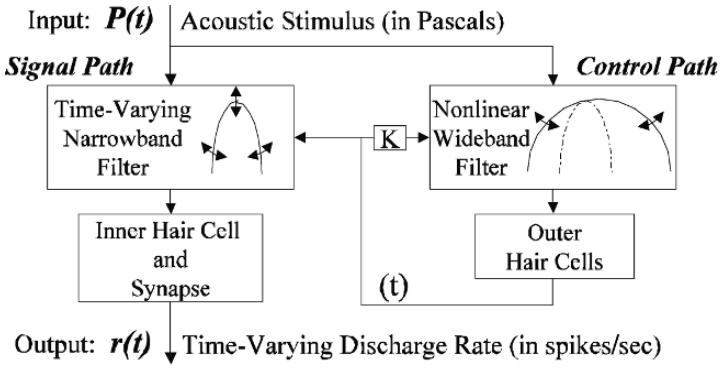


FIGURE 1 Schematical block diagram of the nonlinear auditory nerve (AN) model for a single characteristic frequency (CF) of Heinz et al. (2001a,b). The model uses a nonlinear control path to account for broadened tuning with increases in level, including the associated compressive-magnitude and nonlinear phase responses, suppression and the fast dynamics of these properties. The model was developed for use in modeling basic psychophysical experiments. The approach combines analytical and computational population models of the AN with statistical decision theory to evaluate performance limits in tasks like intensity discrimination and masking.

internal display of activity on a center frequency versus interaural time delay plane is almost universally used for insight about binaural phenomena (for a detailed review see Colburn, 1996). The models differ in their incorporation of interaural intensity differences (IIDs). Models based on interaural coincidence or correlation do not include the effects of IIDs such that additional assumptions and mechanisms need to be included to incorporate these effects. In some cases, the effects of IIDs are incorporated into a common mechanism (e.g., the EC model by Durlach, 1963, 1972; Lindemann, 1986; Gaik, 1993). In other cases (e.g., Jeffress, 1948; Colburn and Durlach, 1978; Colburn et al., 1990; Breebaart et al., 2001a,b,c), IID information is extracted separately and then combined with ITD information. The models of sound localization (e.g., Searle et al., 1976; Middlebrooks, 1992; Blauert, 1997) do not explicitly address mechanisms for the extraction of ITD or IID information; they are mainly concerned with the spectral variation in overall intensity and interaural differences that relate to source position. Concerning the ability of models to describe psychoacoustical data, most models have restricted their application only to a relatively small set of data. The binaural signal detection model by Breebaart et al. (2001a) probably represents the model that has been applied to the largest range of binaural detection conditions. Breebaart et al. (2001b,c) demonstrated that, for many experimental conditions, models based on the EC theory or the cross-correlation give similar results, but that in conditions where predictions differ, an EC-like mechanism may be favored over the cross-correlation.

4.3 Models of Loudness and Pitch

Relatively much is also known about the neural correlates of sensations like loudness and pitch in the early stages of processing. Accordingly, models of loudness and pitch have been commonly associated with mechanisms in these stages. For example, loudness models contain outer- and middle ear transformation, a calculation of the cochlear excitation pattern from the physical spectrum, a transformation of the excitation pattern to a specific loudness pattern and the calculation of the overall loudness. This type of model has been in most cases applied to steady sounds (e.g., Fletcher and Munson, 1937; Zwicker and Scharf, 1965; Zwicker and Fastl, 1999; Moore et al., 1997). Models of loudness of time-varying sounds, where the overall loudness is calculated from the short-term loudness using an averaging mechanism, were suggested by Zwicker (1965, 1977), Fastl (1993), Stone et al. (1997) and Glasberg and Moore (2002). The most recent model (Glasberg and Moore, 2002) predicts the loudness of brief sounds as a function of duration and the loudness of sounds that are amplitude modulated.

Pitch models have been somewhat more complex and none of them has been able to account for all of the experimental data. Detailed reviews about auditory representations of pitch can be found in Lyon and Shamma (1996) and de de Cheveigné (2005). For harmonic complex tones (such as voiced human speech), two main classes of models have been popular. Purely spectral (or spatial) models refer to algorithms that extract pitch from a representation of the spectral profile along the tonotopic axis. All these algorithms require that the input spectral profile be able to resolve (at least partially) several low-order harmonics which are the most important harmonics for pitch perception. There have been hypotheses to explain how pitch can be computed from resolved peaks in the spectral profile (e.g., Goldstein, 1973; Wightman, 1973; Terhardt, 1979; Cohen et al., 1995). A common principle in all these models is a pattern-matching stage which compares the resolved components spectrum to "internally" stored templates of various harmonic series in the central auditory system.

In the other class of models, the representation of pitch is derived exclusively from timing information. The pioneering study of Licklider (1951) about the duplex theory of pitch perception was the basis for later temporal models of pitch perception. For example, in the models of Meddis and Hewitt (1992) and Meddis and O'Mard (1997) the representation of pitch is derived exclusively from the phase-locked pattern of activity in each AN fiber (see Figure 2). It is assumed in the purely temporal models that the central auditory system makes no use of the spatial organization and tonotopic axis. The models "measure" the period or interval of the response of each AN channel and several ways have been suggested how this might be implemented in the central auditory system. Typically, the models form a running autocorrelation function (e.g., Brown and Puckette, 1989; Patterson et al., 1995). This makes it possible to estimate the perceived pitch from the peaks of the autocorrelation functions of the responses, as expressed for instance in a correlogram. An alternative formulation of a temporal mechanism as the basis for pitch has been suggested by de de Cheveigné (1998) where a cancellation is assumed instead of autocorrelation, thus assuming an inhibitory network instead of an excitatory one.

Some models of pitch reflect representations that utilize combinations of spectral and temporal cues. These spectro-temporal models assume, in different ways and implementations, that the tonotopic axis of the cochlea and the specific properties of the cochlear filters (like bandwidth and shape), as well as the information conveyed by phase-locking, are necessary for the representation of pitch (e.g., Shamma and Klein, 2000; Meddis, 1997;

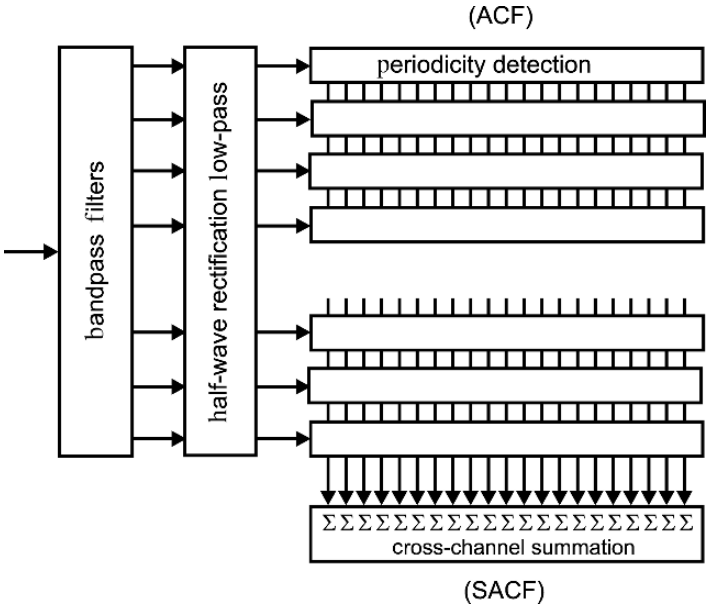


FIGURE 2 Computational model of pitch perception based on autocorrelation (Meddis, 1997). The model essentially consists of four main stages: (1) peripheral (mechanical) bandpass filtering, (2) half-wave rectification and lowpass filtering, (3) within-channel periodicity extraction via the autocorrelation function (ACF) and (4) across-channel aggregation of periodicity estimates via the summary autocorrelation function (SACF). The pitch decision algorithm is then applied to the result of the across-channel aggregation in stage 4.

Moore, 2003). Recently, experimental results by Oxenham et al. (2004) obtained with modulations imposed on high-frequency carrier tones strongly argued in favor of a fundamental role for tonotopic place in the computation of pitch and that, if temporal cues are involved, these must be intrinsically linked to the correct place (see also Shamma, 2004). These experiments provide constraints for future models of the representation of pitch in the auditory system.

4.4 Models of Temporal Resolution

Models of auditory temporal resolution typically have been described as consisting of an initial stage of bandpass filtering, reflecting a simplified action of BM filtering. Each filter is followed by a nonlinear device. In recent models, the nonlinear device typically includes two processes, half-wave rectification and a compressive nonlinearity, resembling the compressive input–output function on the BM (e.g., Ruggero and Rich, 1991; Oxenham and Moore, 1994, 1997; Plack and Oxenham, 1998).

The output is fed to a smoothing device implemented as a lowpass filter (Viemeister, 1979) or a sliding temporal integrator (e.g., Moore et al., 1988). This is followed by a decision device, typically modeled as the signal-to-noise ratio. In the model, the characteristics of the auditory filters are based on auditory filter shapes derived from masking experiments. The characteristics of the nonlinearity and the temporal integrator have been investigated extensively in conditions of nonsimultaneous masking. Forward and backward masking can be accounted for in terms of the build-up and decay processes at the output of the sliding temporal integrator. The same principle model structure has also been suggested to account for other phenomena associated with temporal resolution, such as gap detection and modulation detection (e.g., Viemeister, 1979).

An alternative way of describing forward masking is in terms of neural adaptation (e.g., Jesteadt et al., 1982; Nelson and Swain, 1996). Qualitatively, in an adaptation-based explanation, the portion of the signal furthest from the masker is most important in determining signal threshold, and that portions of the signal closer in time to the masker should contribute less to detection. This is because the response to portions closer in time to the masker will be more adapted, and hence less detectable, than later portions of the signal. Several processing models comprise adaptation and account for various aspects of forward masking (Dau et al., 1996a,b; Buchholz and Mourjoloulus, 2004a,b). For example, the model by Dau et al. (1996a) uses feedback loops to produce an effect similar to adaptation. The dynamic temporal properties of the feedback loops are combined with a logarithmic compression for stationary portions of the stimuli. This allows the model to also account for intensity discrimination data. The model assumes an optimal detector as the decision device, realized as a matched filter applied to the internal representations of the stimuli after the preprocessing stages. Both types of models, temporal integration and adaptation, can lead to similar results even though seemingly conceptually different (Oxenham, 2001).

4.5 Models of Spectro-temporal Processing and Modulation Perception

One of the models that include assumption about the processing in more central stages was suggested by Patterson et al. (1992, 1995). In their model, after the simulation BM motion and conversion into a neural activity pattern in the AN, a form of "strobed" temporal integration is applied to each channel of the neural activity pattern to stabilize any repeating patterns and convert it into a simulation of a stabilized "auditory image" of the sound. Sequences of auditory images were suggested to produce correlates of auditory perceptions that illustrate the dynamic response of the auditory system to everyday sounds.

Another model that includes explicit assumptions about the processing in central stages was suggested by Dau et al. (1997a,b) and in Jepsen et al. (2008). The key stage of these models is a modulation filterbank following the adaptation stage in each peripheral auditory filter (see Figure 3). The filterbank realizes a limited-resolution decomposition of the temporal modulations and was inspired by neurophysiological findings in the auditory brainstem (e.g., Langner and Schreiner, 1988; Palmer, 1995). The information at the output of the processing is combined by an optimal detector assuming independent observations across frequency. The parameters of the filterbank were obtained in simulations of perceptual modulation detection and masking data (e.g., Houtgast, 1989; Bacon and Grantham, 1989; Dau et al.

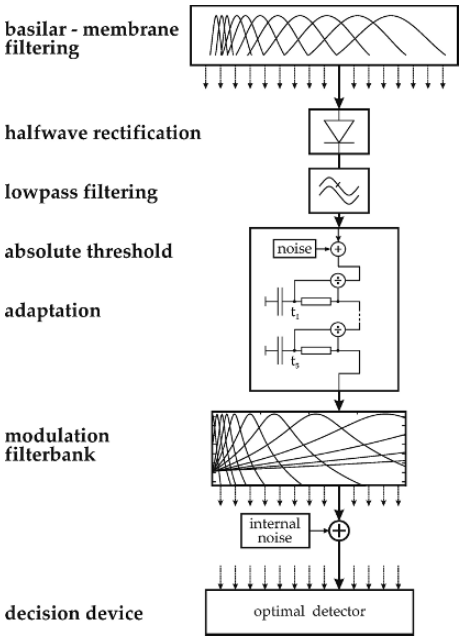


FIGURE 3 Modulation filterbank model by Dau et al. (1997a). The model consists of a simplified action of BM filtering and hair-cell transduction, an adaptation stage and a modulation filterbank. This processing transforms the signals into their internal representations. The signals are finally subjected to an optimal detector as the decision stage. The model describes modulation detection and masking data, spectro-temporal masking and forward masking. A modified version of it has also been used as (monaural) preprocessing for binaural detection (Breebaart et al., 2001a). The internal representation obtained with the model has been tested in several applications, such as in automatic speech recognition and objective speech quality assessment.

1997a,b; Ewert and Dau, 2000). The model has been tested in conditions of simultaneous and nonsimultaneous masking (Dau et al., 1996a,b; Derleth et al., 2001) and spectro-temporal masking (Derleth and Dau, 2000; Verhey, 2002; Verhey et al., 1999; Ewert et al., 2002). The original implementation of the model (assuming a lowpass filter instead of the modulation filterbank) was used as the preprocessing module in the binaural signal detection model by Breebaart et al. (2001a).

A conceptually similar model was presented by Shamma and colleagues (e.g., Chi et al., 1999; Elhilali et al., 2003). They suggested a spectro-temporal analysis of the envelope, motivated by neurophysiological findings in the auditory cortex (Calhoun and Schreiner, 1998; de Charms et al., 1998). In this model, a "spectral" modulation filterbank is integrated with the temporal modulation analysis resulting in two-dimensional. spectro-temporal filters. Thus, in contrast to the current implementation of Dau et al.'s model, where spectral information is integrated subsequent to an independent temporal analysis, Shamma's model also contains joint (and inseparable) spectral–temporal modulations. In conditions where both temporal and spectral features of the input are manipulated, the two models respond differently. Shamma's model has been utilized for the assessment of speech intelligibility (Chi et al., 1999; Elhilali et al., 2003), the prediction of musical timbre (Ru and Shamma, 2003) and the perception of complex sounds (Carlyon and Shamma, 2003).

5 • Applications of Auditory Models

Auditory processing models have been utilized in several applications. For example, systems have been developed for digital recording and transmission of sound that can nearly give transparent high-fidelity reproduction quality but with a largely reduced bit rate (Brandenburg and Stoll, 1994; Gilchrist and Grewin, 1996). These perceptual coders (like, e.g., MP3) rely on the masking properties of the human auditory system, and thus are based on models of auditory signal processing. Listening tests with good coding systems have revealed that the great majority of human listeners cannot distinguish a perceptually coded signal from the original signal (Gilchrist and Grewin, 1996).

Loudness models have been used, for example, for fitting hearing aids. The models have been built into the hearing aids which is possible with modern digital hearing aids. One approach has been to attempt to restore loudness perception to "normal". Hereby, the loudness model helps to determine

the frequency-dependent gain for a given sound such that the loudness perceived by the hearing-impaired listener is the same as would be perceived by a normal-hearing listener without amplification (e.g., Kollmeier and Hohmann, 1995; Launer and Moore, 2002).

In automatic speech recognition (ASR), front ends have been designed to transform the incoming speech signal into a representation which serves as input for later pattern recognition stages. Using the processing model of Dau et al. (1996a) as front end, Tchorz and Kollmeier (1999) and Kleinschmidt et al. (2001) showed that the robustness in noise is higher when compared to "classical" mel-scale cepstral feature extraction. It is mainly the temporal processing in the model that accounts for the robust behavior. The model reflects similar features as RASTA processing of speech (Hermansky and Morgan, 1994). Both techniques perform some kind of envelope bandpass filtering around 4 Hz and hence take roughly 200 ms of integration interval into account for feature calculation. The same preprocessing model has also been utilized for the objective prediction of the transmission quality of low-bit rate speech coding algorithms (Hansen and Kollmeier, 1999).

The computational AN model by Heinz et al. (2001a,b) has been utilized for the interpretation of auditory-evoked brainstem potentials (Dau, 2003). In particular, the model enables studying the effects of nonlinear cochlear processing on the evoked potential patterns generated in the brainstem. This is of interest for the clinical diagnostic of a (cochlear) hearing impairment, particularly in small children and newborns who cannot participate in perceptual listening tests.

References

Bacon SP, Grantham DW (1989) Modulation masking: effects of modulation frequency, depth, and phase. J Acoust Soc Am 85, 2575–2580.

Blauert J (1997) Spatial Hearing: The Psychophysics of Human Sound Localization (MIT, Cambridge, MA).

Brandenburg K, Stoll G (1994) ISO-MPEG-1 audio: a generic standard for coding of high-quality digital audio. J Audio Eng Soc 42, 780–792.

Breebaart J, van de Par S, Kohlrausch A (2001a) Binaural processing model based on contralateral inhibition. I. Model structure. J Acoust Soc Am 110, 1074–1088.

Breebaart J, van de Par S, Kohlrausch A (2001b) Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters. J Acoust Soc Am 110, 1089–1104.

Breebaart J, van de Par S, Kohlrausch A (2001c) Binaural processing model based on contralateral inhibition. III. Dependence on temporal parameters. J Acoust Soc Am 110, 1105–1117.

Brown JC, Puckette MS (1989) Calculation of a "narrowed" autocorrelation function. J Acoust Soc Am 85, 1595–1601.

Buchholz JM, Mourjoloulus J (2004a) A computational auditory masking model based on signal dependent compression. I. Model description and performance analysis. Acust Acta Acust 5, 873–886.

Buchholz JM, Mourjoloulus J (2004b) A computational auditory masking model based on signal dependent compression. II. Model simulations and analytical approximations. Acust Acta Acust 5, 887–900.

Calhoun BM, Schreiner CE (1998) Spectral envelope coding in cat primary auditory cortex: linear and non-linear effects of stimulus characteristics. Eur J Neurosci 10, 926–940.

Carlyon RP, Shamma S (2003) An account of monaural phase sensitivity. J Acoust Soc Am 114, 333–348.

Carney LH (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. J Acoust Soc Am 93, 401–417.

Chi T, Gao Y, Guyton MC, Ru P, Shamma S (1999) Spectro-temporal modulation transfer functions and speech intelligibility. J Acoust Soc Am 106, 2719–2732.

Cohen MA, Grossberg S, Wysell LL (1995) A spectral network model of pitch perception. J Acoust Soc Am 98, 862–879.

Colburn HS (1996) Computational models of binaural processing. In Auditory Computation, edited by HL Hawkins, TA McMullen, AN Popper and RR Fay (Springer, New York), pp 332–400.

Colburn HS, Durlach NI (1978) Models of binaural interaction. In Handbook of Perception, edited by E Carterette and M Friedman (Academic, New York), Vol. IV, pp. 467–518.

Colburn HS, Moss PJ (1981) Binaural interaction models and mechanisms. In Neuronal Mechanisms of Hearing, edited by J Syka and L Aitkin (Plenum Press, New York), pp. 283–288.

Colburn HS, Han YA, Culotta CP (1990) Coincidence model of MSO responses. Hear Res 49, 335–346.

Colburn HS, Carney LH, Heinz MG (2003) Quantifying the information in auditory-nerve responses for level discrimination. J Assoc Res Otolaryngol 4, 294–311.

Dau T (2003) The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. J Acoust Soc Am 113, 936–950.

Dau T, Püschel D, Kohlrausch A (1996a) A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. J Acoust Soc Am 99, 3615–3622.

Dau T, Püschel D, Kohlrausch A (1996b) A quantitative model of the "effective" signal processing in the auditory system. II. Simulations and measurements. J Acoust Soc Am 99, 3623–3631.

Dau T, Kollmeier B, Kohlrausch A (1997a) Modeling auditory processing of amplitude modulation. I. Modulation detection and masking with narrowband carriers. J Acoust Soc Am 102, 2892–2905.

Dau T, Kollmeier B, Kohlrausch A (1997b) Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration in modulation detection. J Acoust Soc Am 102, 2906–2919.
de Boer E (1980) Auditory physics. Physical principles in hearing theory. I. Phys Rep 62, 88–174.
de Boer E (1984) Auditory physics. Physical principles in hearing theory. II. Phys Rep 105, 142–226.
de Boer E (1991) Auditory physics. Physical principles in hearing theory. III. Phys Rep 203, 125–231.
de Boer E (1995) The inverse problem solved for a three-dimensional model of the cochlea. I. Analysis. J Acoust Soc Am 98, 896–903.
de Charms RC, Blake DT, Merzenich MM (1998) Optimizing sound features for cortical neurons. Science 280(5368), 1439–1443.
de Cheveigné A (1998) Cancellation model of pitch perception. J Acoust Soc Am 103, 1261–1271.
de Cheveigné A (2005) Pitch perception models. In Pitch: Neural Coding and Perception, edited by C Plack, AJ Oxenham, AN Popper and RR Fay (Springer, New York).
Delgutte B (1990) Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers. J Acoust Soc Am 87, 791–809.
Delgutte B (1996) Physiological models for basic auditory percepts. In Auditory Computation, edited by HL Hawkins, TA McMullen, AN Popper and RR Fay (Springer, New York), pp. 157–220.
Derleth RP, Dau T (2000) On the role of envelope fluctuation processing in spectral masking. J Acoust Soc Am 108, 285–296.
Derleth RP, Dau T, Kollmeier B (2001) Modeling temporal and compressive properties of the normal and impaired auditory system. Hear Res 159, 132–149.
Duifhuis H (1976) Cochlear nonlinearity and second filter: possible mechanism and implications. J Acoust Soc Am 59, 408–423.
Durlach NI (1963) Equalization and cancellation theory of binaural masking-level differences. J Acoust Soc Am 35, 1206–1218.
Durlach NI (1972) Binaural signal detection: equalization and cancellation theory. In Foundations of Modern Auditory Theory, edited by J Tobias (Academic, New York), pp. 369–462.
Elhilali M, Chi T, Shamma S (2003) A spectro-temporal modulation index (stmi) for assessment of speech intelligibility. Speech Commun 41, 331–348.
Ewert SD, Dau T (2000) Characterizing frequency selectivity for envelope fluctuations. J Acoust Soc Am 108, 1181–1196.
Ewert SD, Verhey JL, Dau T (2002) Spectro-temporal processing in the envelope-frequency domain. J Acoust Soc Am 112, 2921–2931.
Fastl H (1993) Loudness evaluation by subjects and by a loudness meter. In Sensory Research – Multimodal Perspectives, edited by RT Verrillo (Erlbaum, Hillsdale, NJ), pp. 199–210.
Fletcher H, Munson WA (1937) Relation between loudness and masking. J Acoust Soc Am 9, 82–108.

Gaik W (1993) Combined evaluation of interaural time and intensity differences: psychoacoustic results and computer modeling. J Acoust Soc Am 94, 98–110.

Gardner MB, Hawley MS (1972) Network representations of the external ear. J Acoust Soc Am 52, 1620–1628.

Geisler CD (1990) Evidence for expansive power functions in the generation of the discharges of 'low- and medium-spontaneous' auditory-nerve fibers. Hear Res 44, 1–12.

Giguère C, Woodland PC (1994) A computational model of the auditory periphery for speech and hearing research. I. Ascending path. J Acoust Soc Am 95, 331–342.

Gilchrist N, Grewin C (1996) Collected Papers on Digital Audio Bit Rate Reduction (Audio Engineering Society, New York).

Glasberg BR, Moore BCJ (2002) A model of loudness applicable to time-varying sounds. J Audio Eng Soc 50, 331–341.

Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J Neurophysiol 32, 613–636.

Goldstein JL (1973) An optimum processor theory for the central formation of the pitch of complex tones. J Acoust Soc Am 54, 1496–1516.

Goldstein JL (1990) Modeling rapid waveform compression on the basilar membrane as multiple-bandpass-nonlinearity filtering. Hear Res 49, 39–60.

Goldstein JL (1995) Relations among compression, suppression, and combination tones in mechanical responses of the basilar membrane: data and MBPNL model. Hear Res 89, 52–68.

Hansen M, Kollmeier B (1999) Continuous assessment of time-varying speech quality. J Acoust Soc Am 106, 2888–2899.

Heinz MG, Zhang X, Bruce IC, Carney LH (2001a) Auditory-nerve model for predicting performance limits of normal and impaired listeners. ARLO 5(3), 91–96.

Heinz MG, Colburn HS, Carney LH (2001b) Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. Neural Comput 13, 2273–2316.

Hermansky H, Morgan N (1994) Rasta processing of speech. IEEE Trans Speech Audio Process 2, 578–589.

Hewitt MJ, Meddis R (1991) An evaluation of eight computer models of mammalian inner hair-cell function. J Acoust Soc Am 90, 904–917.

Hewitt MJ, Meddis R (1994) A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus. J Acoust Soc Am 95, 2145–2159.

Houtgast T (1989) Frequency selectivity in amplitude-modulation detection. J Acoust Soc Am 85, 1676–1680.

Hubbard AE, Mountain DC (1996) Analysis and synthesis of cochlear mechanical function using models. In Auditory Computation, edited by HL Hawkins, TA McMullen, AN Popper and RR Fay (Springer, New York), pp. 62–120.

Irino T, Patterson RD (1997) A time-domain, level-dependent auditory filter: the gammachirp. J Acoust Soc Am 101, 412–419.

Irino T, Patterson RD (2000) A gammachirp perspective of cochlear mechanics that can also explain human auditory masking quantitatively. In Recent Developments in Auditory Mechanics, edited by H Wada, T Takasaka, K Ikeda, K Ohyama and T Koike (World Scientific, Singapore), pp. 230–236.

Irino T, Patterson RD (2001) A compressive gammachirp auditory filter for both physiological and psychophysical data. J Acoust Soc Am 109, 2008–2022.

Jeffress LA (1948) A place theory of sound localization. J Comp Physiol Psychol 41, 35–39.

Jepsen ML, Ewert SD, Dau T (2008) A computational model of human auditory signal processing and perception. J Acoust Soc AM in press, expected in issue 124(1) July 2008.

Jesteadt W, Bacon SP, Lehman JR (1982) Forward masking as a function of frequency, masker level, and signal delay. J Acoust Soc Am 71, 950–962.

Joris PX, Yin TCT (1995) Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. J Neurophysiol 73, 1043–1062.

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84, 541–577.

Killion MC, Clemis JD (1981) An engineering view of middle ear surgery. J Acoust Soc Am 69 (Suppl 1), S44.

Kim DO (1986) Active and nonlinear biomechanics and the role of the outer-hair-cell subsystem in the mammalian auditory system. Hear Res 22, 105–114.

Kleinschmidt M, Tchorz J, Kollmeier B (2001) Combining speech enhancement and auditory feature extraction for robust speech recognition. Speech Commun 32, 75–91.

Kollmeier B, Hohmann V (1995) Loudness estimation and compensation employing a categorical scale. In Advances in Hearing Research, edited by GA Manley, GM Klump, C Köppl, H Fastl and H Öckinghaus (World Scientific, Singapore).

Kuwada S, Yin TC, Syka J, Buunen TJ, Wickesberg RE (1984) Binaural interaction in low-frequency neurons in inferior colliculus of the cat. IV. Comparison of monaural and binaural response properties. J Neurophysiol 51, 1306–1325.

Langner G (1981) Neuronal mechanisms for pitch analysis in the time domain. Exp Brain Res 44, 450–454.

Langner G (1997) Neural processing and representation of periodicity pitch. Acta Otolaryngol Suppl 532, 68–76.

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J Neurophysiol 60, 1799–1822.

Langner G, Sams M, Heil P, Schulze H (1997) Frequency and periodicity are represented in orthogonal maps in the human auditory

cortex: evidence from magnetoencephalography. J Comp Physiol 181, 665–676.
Launer S, Moore BCJ (2002) Use of a loudness model for hearing aid fitting. V. On-line gain control in a digital hearing aid. Int J Audiol 42, 262–273.
Licklider JC (1951) A duplex theory of pitch perception. Experientia 7, 128–134.
Lindemann W (1986) Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. J Acoust Soc Am 80, 1608–1622.
Lopez-Poveda EA, Meddis R (2001) A human nonlinear cochlear filterbank. J Acoust Soc Am 110, 3107–3118.
Lorenzi C, Micheyl C, Berthommier F (1995) Neuronal correlates of perceptual amplitude-modulation detection. Hear Res 90, 219–227.
Lutman ME, Martin AM (1979) Development of an electroacoustic analogue model of the middle ear and acoustic reflex. J Sound Vibr 64, 133–157.
Lyon R, Shamma S (1996) Auditory representations of timbre and pitch. In Auditory Computation, edited by HL Hawkins, TA McMullen, AN Popper and RR Fay (Springer, New York), pp. 221–270.
McAlpine D, Jiang D, Shackleton TM, Palmer AR (1998) Convergent input from brainstem coincidence detectors onto delay-sensitive neurons in the inferior colliculus. J Neurosci 18, 6026–6039.
Meddis R (1986) Simulation of mechanical to neural transduction in the auditory receptor. J Acoust Soc Am 79, 702–711.
Meddis R (1988) Simulation of auditory-neural transduction: further studies. J Acoust Soc Am 83, 1056–1063.
Meddis R, Hewitt MJ (1992) Modeling the identification of concurrent vowels with different fundamental frequencies. J Acoust Soc Am 91, 233–245.
Meddis R, O'Mard L (1997) A unitary model of pitch perception. J Acoust Soc Am 102, 1811–1820.
Meddis R, O'Mard LP, Lopez-Poveda EA (2001) A computational algorithm for computing nonlinear auditory frequency selectivity. J Acoust Soc Am 109, 2852–2861.
Middlebrooks JC (1992) Narrow-band sound localization related to external ear acoustics. J Acoust Soc Am 92, 2607–2624.
Møller AR (1976) Dynamic properties of excitation and two-tone inhibition in the cochlear nucleus studied using amplitude-modulated tones. Exp Brain Res 25, 307–321.
Moore BCJ (2003) An Introduction to the Psychology of Hearing, 5th Ed. (Academic, Amsterdam).
Moore BCJ, Glasberg BR, Plack CJ, Biswas AK (1988) The shape of the ear's temporal window. J Acoust Soc Am 83, 1102–1116.
Moore BCJ, Glasberg BR, Baer T (1997) A model for the prediction of thresholds, loudness, and partial loudness. J Audio Eng Soc 45, 224–240.

Neely ST (1983) The cochlear amplifier. In Mechanics of Hearing, edited by E de Boer and MA Viergever (Martinus Nijhoff, The Hague), pp. 111–118.
Nelson PC, Carney LH (2004) A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. J Acoust Soc Am 116, 2173–2186.
Nelson DA, Swain AC (1996) Temporal resolution within the upper accessory excitation of a masker. Acust Acta Acust 82, 328–334.
Olson ES (1999) Direct measurement of intracochlear pressure waves. Nature (London) 402, 526–529.
Oxenham AJ (2001) Forward masking: adaptation or integration? J Acoust Soc Am 109, 732–741.
Oxenham AJ, Moore BCJ (1994) Modeling the additivity of nonsimultaneous masking. Hear Res 80, 105–118.
Oxenham AJ, Moore BCJ (1997) Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing. In Modeling Sensorineural Hearing Loss, edited by W Jesteadt (Erlbaum, Hillsdale, NJ).
Oxenham AJ, Bernstein JG, Penagos H (2004) Correct tonotopic representation is necessary for complex pitch perception. Proc Natl Acad Sci 101, 1421–1425.
Palmer AR (1995) Neural signal processing. In Hearing, edited by BCJ Moore (Academic, San Diego).
Palmer AR, McAlpine D, Jiang D (1997) Processing of interaural delay in the inferior colliculus. In Acoustical Signal Processing in the Central Auditory System, edited by J Syka (Plenum, New York), pp. 353–364.
Patterson RP, Nimmo-Smith I, Holdsworth J, Rice P (1987) An efficient auditory filterbank based on the gammatone function. In a paper presented at a meeting of the IOC Speech Group on Auditory Modeling at RSRE.
Patterson RD, Robinson K, Holdsworth J, McKeown D, Zhang C, Allerhand M (1992) Complex sounds and auditory images. In Auditory Physiology and Perception, Proceedings of the 9th International Symposium on Hearing, edited by Y Cazals, L Demany and K Horner (Pergamon, Oxford), pp. 429–446.
Patterson RD, Allerhand MH, Giguère C (1995) Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. J Acoust Soc Am 98, 1890–1894.
Pfeiffer RR (1970) A model for two-tone inhibition of single cochlear-nerve fibers. J Acoust Soc Am 48, 1373–1378.
Plack CJ, Oxenham AJ (1998) Basilar-membrane nonlinearity and the growth of forward masking. J Acoust Soc Am 103, 1598–1608.
Rees A, Møller AR (1983) Responses of neurons in the inferior colliculus of the rat to AM and FM tones. Hear Res 10, 301–330.
Rhode WS (1971) Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. J Acoust Soc Am 49, 1218–1231.
Rhode WS, Oertel D, Smith PH (1983) Physiological response properties of cells labeled intracellularly with horseradish peroxidase in cat ventral cochlear nucleus. J Comp Neurol 213, 448–463.

Ru P, Shamma S (2003) Presentation of musical timbre in the auditory cortex. J New Music Res 26, 154–169.
Ruggero MA, Rich NC (1991) Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane. J Neurosci 11, 1057–1067.
Sachs MB, Abbas PJ (1974) Rate versus level functions for auditory nerve fibers in cats: tone burst stimuli. J Acoust Soc Am 81, 680–691.
Schoonhoven R, Prijs VF, Frijns JH (1997) Transmitter release in inner hair cell synapses: a model analysis of spontaneous and driven rate properties of cochlear nerve fibres. Hear Res 113, 247–260.
Searle CL, Braida LD, Davis MF, Colburn HS (1976) Model for auditory localization. J Acoust Soc Am 60, 1164–1175.
Shamma S, Klein D (2000) The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. J Acoust Soc Am 107, 2631–2644.
Shamma SA (2004) Topographic organization is essential for pitch perception. Proc Natl Acad Sci 101, 1114–1115.
Shamma SA, Shen NM, Gopalaswamy P (1989) Stereausis: binaural processing without neural delays. J Acoust Soc Am 86, 989–1006.
Shaw EAG (1980) The acoustics of the external ear. In Acoustical Factors Affecting Hearing Aid Performance, edited by GA Studebaker and IH Hochberg (University Park, Baltimore), pp. 109–125.
Shekhter I, Carney LH (1997) A nonlinear auditory nerve model for CF-dependent shift in tuning with sound level. Assoc Res Otolaryngol Abs 20, 670.
Shera CA (2001) Frequency glides in click responses of the basilar membrane and auditory nerve: their scaling behavior and origin in traveling-wave dispersion. J Acoust Soc Am 109, 2023–2034.
Siebert WM (1965) Some implications of the stochastic behavior of primary auditory neurons. Kybernetik 2, 206–215.
Siebert WM (1970) Frequency discrimination in the auditory system: place or periodicity mechanism. Proc IEEE 58, 723–730.
Smith PH, Rhode WS (1989) Structural and functional properties distinguish two types of multipolar cells in the ventral cochlear nucleus. J Comp Neurol 282, 595–616.
Stone MA, Moore BCJ, Glasberg BR (1997) A real-time DSP-based loudness meter. In Contributions to Psychological Acoustics, edited by A Schick and M Klatte (BIS Universität, Oldenburg), pp. 587–601.
Summer CJ, O'Mard LP, Lopez-Poveda EA, Meddis R (2003) A nonlinear filter-bank model of the guinea-pig cochlear nerve: Rate responses. J Acoust Soc Am 113, 3264–3274.
Tchorz J, Kollmeier B (1999) A model of auditory perception as front end for automatic speech recognition. J Acoust Soc Am 106, 2040–2050.
Terhardt E (1979) Calculating virtual pitch. Hear Res 1, 155–182.
Verhey JL (2002) Modeling the influence of inherent envelope fluctuations in simultaneous masking experiments. J Acoust Soc Am 111, 1018–1025.
Verhey JL, Dau T, Kollmeier B (1999) Within-channel cues in comodulation masking release (CMR): experiments and model predictions using a modulation-filterbank model. J Acoust Soc Am 106, 2733–2745.

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 66, 1364–1380.
Wightman FL (1973) The pattern-transformation model of pitch. J Acoust Soc Am 54, 407–416.
Zacksenhouse M, Johnson DH, Tsuchitani C (1992) Excitatory/inhibitory interaction in the LSO revealed by point process modeling. Hear Res 62, 105–123.
Zhang XD, Heinz MG, Bruce IC, Carney LH (2001) A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. J Acoust Soc Am 109, 648–670.
Zweig G (1991) Finding the impedance of the organ of Corti. J Acoust Soc Am 89, 1229–1254.
Zwicker E (1965) Temporal effects in simultaneous masking and loudness. J Acoust Soc Am 38, 132–141.
Zwicker E (1977) Procedure for calculating loudness of temporally variable sounds. J Acoust Soc Am 62, 675–682.
Zwicker E, Fastl H (1999) Psychoacoustics – Facts and Models, 2nd Ed. (Springer, Berlin).
Zwicker E, Scharf B (1965) A model of loudness summation. Psychol Rev 72, 3–26.
Zwislocki JJ (1962) Analysis of the middle-ear function. Part I: input impedance. J Acoust Soc Am 34, 1514–1523.

13

Speech Intelligibility

Thomas Brand

University of Oldenburg, Oldenburg, Germany

1 • Speech Intelligibility

1.1 Definition

Speech intelligibility (SI) is important for different fields of research, engineering and diagnostics in order to quantify very different phenomena like the quality of recordings, communication and playback devices, the reverberation of auditoria, characteristics of hearing impairment, benefit using hearing aids or combinations of these things. A useful way to define SI so that it is directly and quantitatively measurable is: *"Speech intelligibility SI is the proportion of speech items (e.g., syllables, words or sentences) correctly repeated by (a) listener(s) for a given speech intelligibility test."*

1.2 Intelligibility Function and Speech Reception Threshold (SRT)

The Intelligibility function describes the listener's speech intelligibility SI as a function of speech level L (in dB), which may either refer to the sound pressure level of the speech signal or to the speech-to-noise ratio (SNR) if the test is performed with interfering noise (Figure 1).

In most cases it is possible to fit the function SI(L) to the empirical data using

$$\mathrm{SI}(L) = \frac{1}{A}\left(1 + \mathrm{SI}_{\max}\frac{A-1}{1+\exp\left(-\frac{L-L_{\mathrm{mid}}}{s}\right)}\right) \tag{1}$$

with

L_{mid}: speech level of the midpoint of the intelligibility function.
s: slope parameter. The slope at L_{mid} is given by $\frac{\mathrm{SI}_{\max}(A-1)}{4\,A\,s}$.

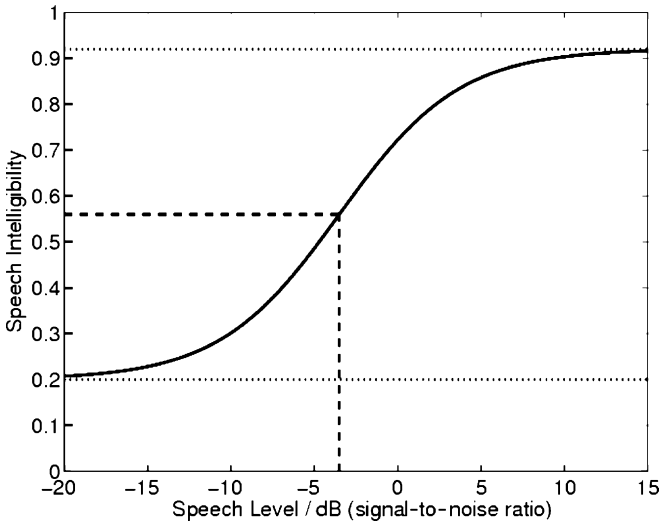


FIGURE 1 Typical example of SI function (*solid line*) for word intelligibility test (closed response format with five response alternatives). The *dashed line* denotes L_{mid}. The *dotted lines* denote the lower limit ($1/A$) and the upper limit (SI_{asymp}) of the SI function. Parameters: $L_{\text{mid}} = -3.5$ dB SNR, $\text{SI}_{\text{max}} = 0.9$ ($\text{SI}_{\text{asymp}} = 0.92$), $A = 5$, slope $= 0.05$/dB($s = 3.6$ dB).

SI_{max}: parameter for maximum intelligibility. This can be smaller than 1 in some cases (e.g., distorted speech signals or listeners with hearing impairment). The asymptotic maximum of SI is $\text{SI}_{\text{max}} + (1 - \text{SI}_{\text{max}})/A$.

A: Number of response alternatives. For example, $A = 10$ when the listener should respond in a closed response format using digits between "0" and "9". In SI tests with "open response format", like word tests without limiting the number of response alternatives, A is assumed to be infinite, which means

$$\text{SI} = \text{SI}_{\text{max}} \frac{1}{1 + \exp\left(-\frac{L - L_{\text{mid}}}{s}\right)} \quad \text{and slope} = \frac{\text{SI}_{\text{max}}}{4\,s}. \qquad (2)$$

The primary interest of many applications is the speech reception threshold (SRT), which denotes the speech level (measured in dB) for a given intelligibility (e.g., SI = 0.5 or 0.7).

1.3 Accuracy of SI Measurements

The accuracy of SI measurements is given by the binomial distribution. Consequently, the standard error se(SI) of a SI estimate which is based on n items (e.g., words) is

$$\mathrm{se(SI)} = \sqrt{\frac{\mathrm{SI}(1-\mathrm{SI})}{n}}. \tag{3}$$

A further increase of this standard error is caused by the fact that SI tests consist of several items (e.g., 50 words) which unavoidably differ in SI. Therefore, SI tests should be constructed in a way that the SI of all items is as homogeneous as possible.

To a first approximation, the standard error of the SRT is equal to $\mathrm{se(SI_{SRT})}$ (the standard error of the SI estimate at the SRT) divided by the slope of the intelligibility function at the SRT:

$$\mathrm{se(SRT)} = \frac{\mathrm{se(SI_{SRT})}}{\mathrm{slope(SRT)}}. \tag{4}$$

2 • Measurement Methods

2.1 Speech Materials

A speech material (i.e., a set of speech items like words or sentences) is suitable for SI tests when different requirements are fulfilled: the different speech items have to be homogeneous in SI in order to yield high measurement accuracy and reproducibility in a limited measuring time. The distribution of phonemes should represent the respective language. Only speech materials which have been optimized properly by a large number of evaluation measurements can fulfill these requirements.

A large number of SI tests using different materials are available for different languages. An overview of American SI tests can be found in Penrod (1994). There are different formats, i.e., nonsense syllables, single words and sentences. Sentences represent a realistic communication situation best. Nonsense syllables and words allow for assessing confusion matrices and for analyzing transmission of information. Furthermore, the intelligibility functions of most sentence tests (Plomp and Mimpen, 1979; Hagerman, 1982; Nilsson et al., 1994; Kollmeier and Wesselkamp, 1997; Wagener et al., 1999, 2003) show slopes between 0.15 and 0.25 per dB which are considerably steeper than the values obtained with nonsense syllables or single-word tests.

Since the standard deviation of SRT estimates is inversely proportional to the slope of the intelligibility function, these sentence tests are better suited for efficient and reliable SRT measurements than single-word tests.

2.2 Presentation Modes

Signals can be presented via either loudspeakers (free field condition) or headphones. The free field condition is more natural but it has the drawbacks of requiring a larger experimental effort, being difficult to calibrate, and, especially for spatial speech/noise situations, small movements of the listeners head may influence the result of the SI measurement.

The advantages of presentation via headphones are: very good reproducibility for each individual listener; smaller experimental effort; and spatial speech/noise conditions can be easily realized using virtual acoustics. Drawbacks are: the individual calibrations are complicated because headphones produce different sound pressures in different ears; and measurements with hearing aids are not possible.

Adaptive procedures (Chapter 15) can be used to yield SRT estimates efficiently by concentrating the presentation levels near to the SRT. In sentence tests, each word can be scored independently which allows adaptive procedures to be designed which converge more efficiently than adaptive procedures usually used in psychoacoustics (Brand and Kollmeier, 2002).

3 • Factors Influencing Speech Intelligibility

3.1 Measuring Method

The different speech materials mentioned above generate different results. Therefore, only standardized speech materials or speech materials with well-known reference intelligibility functions should be used.

3.2 Noise and Room Acoustics

Noise and reverberation reduce SI. Therefore, if the intention is to measure SI in silence, environmental noise and reverberation have to be minimized using, for example, sound-insolated cabins and damping headphones. On the other hand, measurements can be used to specifically investigate the influence of different noises and room acoustics on SI, which is important for the so-called "cocktail party phenomenon".

3.3 Cocktail Party Phenomenon

The human auditory system has very impressing abilities in understanding a target talker even if maskers, i.e., competitive sound sources like different talkers, are present at the same time. This is the so-called "cocktail party phenomenon" and an interesting review of research on this topic can be found in Bronkhorst (2000). The SI in these multi-talker conditions is influenced by many masker properties like sound pressure level, frequency spectrum, amplitude modulations, spatial direction and number of maskers. The spatial configuration of target speaker and masker plays a very important role. Binaural hearing (hearing

with both ears) enables a very effective release from masking (improvement of the SRT) of up to 12 dB compared to monaural hearing (hearing with one ear) (Bronkhorst, 2000).

3.4 Hearing Impairment

An introduction to SI in clinical audiology can be found in Penrod (1994). Hearing impairment can lead to an increase of the SRT, a decrease of the maximum reachable intelligibility SI_{asymp} and a flatter slope of the intelligibility function. The most difficult situations for hearing impaired listeners are noisy environments with many interfering sound sources ("cocktail party situation"). Therefore, SI tests on noise are important diagnostic tools for assessing the daily life consequences of a hearing impairment and the benefit of a hearing aid. SI plays a very important role for the research on and the fitting of hearing aids (see Chapter 14).

4 • Prediction Methods

4.1 Articulation Index (AI), Speech Intelligibility Index (SII) and Speech Transmission Index (STI)

The most common methods for the prediction of speech intelligibility are the Articulation Index (AI) (French and Steinberg, 1947; Fletcher and Galt, 1950; ANSI, 1969), which was renamed Speech Intelligibility Index (SII) (ANSI, 1997), and the Speech Transmission Index (STI) (Houtgast and Steeneken, 1985; IEC, 1998) (Table 1). The strength of these models is the large amount of empirical knowledge they are based on. All these models assume that speech is coded by several frequency channels that carry independent information. This can be expressed by

$$\mathrm{AI} = \sum_i \mathrm{AI}_i, \tag{5}$$

with AI denoting the cumulative Articulation Index of all channels and AI_i denoting the Articulation Index of a single channel.

AI and SII are derived from the speech signal by calculating the signal-to-noise ratio (SNR) in the different frequency channels:

$$\mathrm{AI} = \sum_i \frac{W_i(\mathrm{SNR}_i + 15)}{30}, \tag{6}$$

with W_i denoting a frequency channel weighting factor and SNR_i denoting the signal-to-noise ratio in channel i. W_i depends on the speech material used and takes into account the fact that high frequencies are more important for the recognition of consonants than for the recognition of meaningful sentences. The main differences between the different versions of AI and SII are the way they include non-linearities like distortion, masking and broadening of frequency bands.

TABLE 1 Examples of methods for the prediction of speech intelligibility (SI)

Method (reference)	Signal parameters	Comments
Articulation Index, AI (French and Steinberg, 1947)	Levels and frequency spectra of speech and noise, kind of speech material	Macroscopic model that describes the influence of the frequency content of speech on intelligibility
Articulation Index, AI (Fletcher and Galt, 1950)	Levels and frequency spectra of speech and noise, kind of speech material	More complex than French and Steinberg version, describes more non-linear effects, seldom used
Articulation Index, AI (ANSI S3.5-1969)	Levels and frequency spectra of speech and noise	Simplified version based on French and Steinberg (1947), not in use anymore
Speech Intelligibility Index, SII (ANSI S3.5-1997)	Levels and frequency spectra of speech and noise, kind of speech material, hearing loss	Revision of ANSI S3.5-1969, includes spread of masking, standard speech spectra, relative importance of frequency bands
Speech Transmission Index, STI (IEC 60268-16)	Modulation transfer function	Predicts the change of intelligibility caused by a speech transmission system (e.g., an auditorium) based on the modulation transfer function of the system
Speech recognition sensitivity model, SRS (Müsch and Buus, 2001a,b)	Levels and frequency spectra of speech and noise, number of response alternatives	Alternative to SII, handles frequency band interactions and is better suited for unsteady frequency spectra
Holube and Kollmeier (1996)	Speech and noise signals, hearing loss	Microscopic modeling of signal processing of auditory system combined with simple automatic speech recognition

The Speech Transmission Index (STI) uses the modulation transfer function instead of the SNR and is especially successful for predicting SI in auditoria and rooms, because it explicitly takes into account the flattening of the information-carrying speech envelopes due to reverberation.

The transformation of AI, SII or STI, respectively, to speech intelligibility requires a non-linear transformation that has to be fitted to empirical data. The transformation depends on the kind

of speech material used and is usually steeper for materials with context (e.g. sentences) than for materials with single words.

4.2 Statistical Methods

The assumption of independent information in different frequency channels does not hold in all situations because synergetic as well as redundant interactions between different channels occur. The speech recognition sensitivity model (Müsch and Buus, 2001a,b) takes these interactions into account using statistical decision theory in order to model the linguistic entropy of speech.

4.3 Functional Method

The methods mentioned above are based on relatively rough parameterizations of speech, such as long-term frequency spectrum or the modulation transfer function. The method proposed by Holube and Kollmeier (1996), however, is based on physiological and psychoacoustical data and is a combination of a functional model of the human auditory system {Chap. 12} and a simple automatic speech recognition system {Chapter 29}. A drawback of this approach is that there is still a large gap between recognition rates of humans and automatic speech recognition systems (for a review, see Lippmann, 1997).

References

ANSI (1969) Methods for the Calculation of the Articulation Index, ANSI S3.5-1969, American National Standards Institute, New York.

ANSI (1997) Methods for Calculation of the Speech Intelligibility Index, ANSI S3.5-1997, American National Standards Institute, New York.

Brand T, Kollmeier B (2002) Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests, J Acoust Soc Am **111**(6), 2801–2810.

Bronkhorst A (2000) The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions, Acustica **86**, 117–128.

Fletcher H, Galt RH (1950) The perception of speech and its relation to telephony, J Acoust Soc Am **22**, 89–151.

French NR, Steinberg JC (1947) Factors governing the intelligibility of speech sounds, J Acoust Soc Am **19**, 90–119.

Hagerman B (1982) Sentences for testing speech intelligibility in noise, Scand Audiol **11**, 79–87.

Holube I, Kollmeier B (1996) Speech intelligibility prediction in hearing-impaired listeners based on a psychoacoustically motivated perception model, J Acoust Soc Am **100**, 1703–1716.

Houtgast T, Steeneken HJM (1985) A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am **77**, 1069–1077.

IEC (1998) Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index, INTERNATIONAL STANDARD 60268-16 Second edition 1998–03.

Kollmeier B, Wesselkamp M (1997) Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment, J Acoust Soc Am **102**, 2412–2421.

Lippmann R (1997) Speech recognition by machines and humans, Speech Commun **22,** 1–15.

Müsch H, Buus S (2001a) Using statistical decision theory to predict speech intelligibility. I. Model structure, J Acoust Soc Am **109**, 2896–2909.

Müsch H, Buus S (2001b) Using statistical decision theory to predict speech intelligibility. II. Measurement and prediction of consonant-discrimination performance, J Acoust Soc Am **109**, 2910–2920.

Nilsson M, Soli SD, Sullivan JA (1994) Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise, J Acoust Soc Am **95**(2), 1085–1099.

Penrod JP (1994) Speech threshold and word recognition/discrimination testing. In Handbook of Clinical Audiology, edited by Katz J, Williams & Wilkins, Baltimore, MD, 4th ed., Chap. 10, 147–164.

Plomp R, Mimpen A (1979) Improving the reliability of testing the speech-reception threshold for sentences, Audiology **18**, 43–52.

Wagener K, Kühnel V, Kollmeier B (1999) Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests (Development and evaluation of a German sentence test I: Design of the Oldenburg sentence test), Z Audiol **38**, 4–15.

Wagener K, Josvassen JL, Ardenkjaer R (2003) Design, optimization and evaluation of a Danish sentence test in noise, Int J Audiol **42**(1), 10–17.

14

Signal Processing in Hearing Aids

Volker Hohmann

University of Oldenburg, Oldenburg, Germany

Since the introduction of the first commercial digital hearing aid in 1996, the possibilities of digital signal processing (DSP) have increasingly been exploited in hearing aids. DSP allows for the implementation of signal processing schemes (i.e., "algorithms") that have no counterpart in the analog domain and thus offers new, interesting perspectives for the rehabilitation of hearing impairment, only parts of which have been realized to date. This section reviews the processing schemes currently used in digital hearing aids and outlines the main lines of research toward improved processing schemes.

1 • Basic Approach to the Rehabilitation of Hearing Deficiencies

Processing schemes divide into two general categories. *Compensation schemes* compensate for altered aspects of auditory perception by preprocessing the acoustic signals in such a way that the combination of preprocessing and altered perception by the hearing impaired leads to an "aided" perception which is as close to the normal perception as possible. Amplification, dynamic range compression and several speech enhancement schemes fall into this category. A key problem is that the alteration of perception has to be described by nonlinear processes in many cases, which are not

invertible in general: A combination of two nonlinear systems does not necessarily mimic the behavior of another (linear or nonlinear) system. Compensation schemes explicitly or implicitly consider a model of the altered perception (auditory processing models or perceptual models, see Chapter 12 and 15).

Substitution schemes, on the other hand, aim at substituting lost aspects of auditory perception by implementing these aspects fully by signal processing. Examples for substitution schemes are predominantly noise reduction schemes, which generally aim at improving the signal-to-noise ratio (SNR). They might be based fully or in parts on a simulation (model) of the principles of auditory perception, or might follow a purely technical-based approach.

All approaches have to obey the limits for the processing delay of the hearing aid between output and input signal (algorithmic filter delay plus hardware delay (e.g., for signal buffering)). Synchronization of audio information and visual information (lip reading) is maintained up to a delay of about 70 ms (see also Chapter 37). However, in case the hearing loss is small in specific frequency regions, both the direct sound and the processed signal are audible, and disturbing comb filter effects may occur. This effect is especially pronounced when listening to one's own voice. Based on the empirical results, it is generally agreed that the delay of a hearing aid should not exceed 10 ms to avoid all potentially detrimental effects [1]. This limit is reached in most FFT-based systems, whereas filterbank-based systems show lower delays between 6 and 2 ms, decreasing with frequency. An important consequence of the delay limit is that models of higher cognitive processes, e.g., processes involving short-term memory, cannot be used in the signal path in principle.

2 • Factors of Hearing Impairment and Rehabilitation Strategies

Table 1 indicates several factors of hearing impairment, which have been identified based on physiological and psychoacoustical evidence (Chapter 10 and [2–4]) and lists generic rehabilitation strategies applied to compensate for their perceptual consequences. For further reading, details on specific algorithms are given in [5,6] and in the references listed in Table 1.

The major components of current hearing aids are an amplification stage and a compressive automatic gain control (AGC) stage (or dynamic range compression stage) to compensate for the frequency-specific increase in hearing threshold level and the reduced dynamic range between threshold level and uncomfortable loudness level ("recruitment"). Several rules for

TABLE 1 Factors of hearing impairment and rehabilitation strategies

Factor	Perceptual consequence	Rehabilitation strategy	Refs.
Attenuation component	Loss of sensitivity; increased threshold level	Increase audibility by	
		(a) frequency-specific amplification	[7,8]
		(b) frequency compression	[9,10]
Distortion component	(a) Loss of sensitivity; reduced dynamic range ("recruitment")	(a) Automatic gain control (AGC)	[11–13]
		(b) spectral enhancement	[14,15]
	(b) reduced frequency selectivity	(c) noise reduction (see "neural component")	
	(c) increased susceptibility to background noise		
Neural component	(a) increased susceptibility to background noise ("Cocktail-party effect")	(a) monaural noise reduction	[16,17]
		(b) directional microphones	[18]
	(b) impaired binaural capabilities	(c) beamformer	[19,20]
		(d) binaural noise reduction	[21,22]

individually fitting gain and compression parameters are used in practice [5]. Among the noise reduction schemes listed above, the first-order adaptive directional microphones [18] turned out to be the most successful strategies and are used in most current hearing aids for improved hearing in noise. Monaural noise reduction schemes [16] are also well established but are restricted to improving perceived noisiness of a noise condition, failing to improve speech intelligibility [23]. This limitation also holds for most published binaural noise reduction schemes [21].

In addition to the rehabilitation strategies, several technical limitations have to be addressed by DSP. In particular, feedback "howling" is a major problem because of the high acoustic gain and the close distance of microphone and speaker/receiver mounted in a single shell, giving rise to acoustic and mechanic feedback. DSP-based *feedback control* has been used to increase gain margins by about 10–15 dB in realistic conditions [24], but there is a demand for even more gain. As DSP increasingly offers specific processing schemes tailored to specific acoustic conditions (e.g., noise reduction for specific noise conditions like diffuse noise or stationary noise), an increasing demand for *automatic environment classification* [25] is observed.

Figure 1 shows a binaural hearing aid signal processing scheme, including the components listed above. Note that a fully binaural system which requires comparing signals from left and right ears is not commercially available to date because the

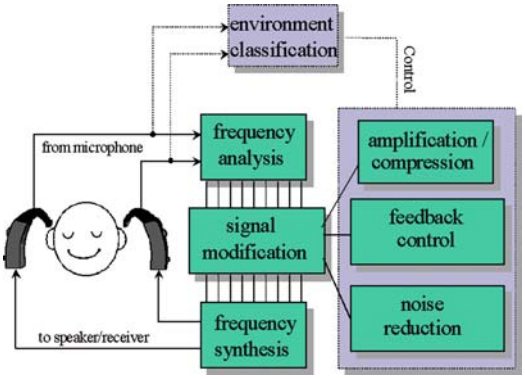


FIGURE 1 Block diagram of a generic binaural hearing aid signal processing scheme. Acoustic signals are picked up by microphones placed close to the ears, processed by a central processing device and played back at both ears. Processing includes frequency analysis and synthesis and several processing blocks (e.g., amplification and compression, feedback control, noise reduction schemes), which are controlled by automatic environment classification (e.g., activating noise reduction depending on estimated noise condition). Note that a fully binaural system as sketched here is currently not available in commercial hearing aids due to the lack of low-power wireless transmission.

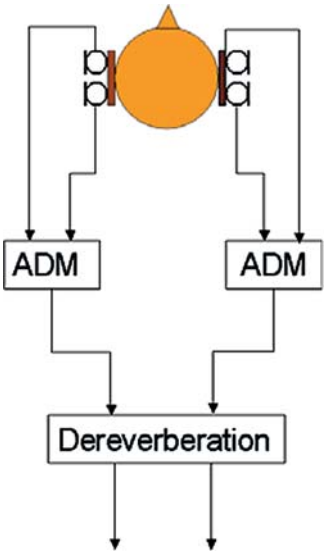


FIGURE 2 Generic binaural processing scheme combining two adaptive differential microphones (ADM), one at each ear, and a binaural dereverberation scheme, providing suppression of one directional noise source and of diffuse background noise and reverberation. The scheme provides a binaural output, i.e., the binaural cues are maintained and may be used by the residual binaural processing capabilities of the hearing impaired for localizing sound sources and exploiting binaural differences for noise reduction.

technology does not allow for signal transmission at the original data rate yet. Figure 2 sketches a promising combination of established processing schemes, i.e., the adaptive differential microphone [18] and a binaural dereverberation scheme [21], which possibly represents the current optimum that could be done in diffuse and reverberant noise conditions as soon as a binaural link is available.

3 • Limitations of Current Rehabilitation Strategies

Analyzing the performance of current hearing aids, further development should concentrate on two areas. First, AGC schemes use rather long adaptation time constants and a few frequency bands as compared to the normal hearing system, which provides a large number of overlapping bands and almost instantaneous amplitude compression [2]. Further research should be conducted to develop better simulations of the instantaneous spectro-temporal compressive processing implemented in the auditory system, which might be applied subsequently instead of the current AGC stages. Second, the benefit of noise reduction schemes, including published binaural schemes, is rather small and cannot compensate for degraded hearing in noise and reverberation associated with the most common types of hearing impairment. The normal hearing system, including the cognitive system, developed incredibly effective means of separating sound sources to ensure communication in difficult noise conditions. Some of the principles of the separation process have recently been experimentally revealed [26], but their simulation using computational auditory scene analysis (CASA) is still limited and not yet applicable to improve speech intelligibility in noise. Future research on noise reduction should therefore concentrate on exploiting CASA techniques. The properties of the hearing systems known so far suggest that the system exploits a priori knowledge on possible sound sources to identify sound objects and to extrapolate their temporal evolution in case part of the received acoustical information on these objects is ambiguous or even fully masked. Statistical filtering techniques (e.g., particle filters) incorporating dynamical models of possible sound sources [27] are therefore expected to play a major role in this line of research.

4 • Conclusions

Digital signal processing (DSP) is currently used to compensate for all aspects of hearing impairment and to overcome technical limitations of hearing aids. Increased user benefit of DSP has

been shown in many studies. However, many of the digital processing schemes used in up-to-date hearing aids are not more than generalized versions of the analog schemes used before, and some hearing deficiencies are not satisfactorily rehabilitated by DSP yet, especially when hearing in noise and when reverberation is involved. Basic research needs to be pursued in order to reveal fundamentally new processing schemes for an improved rehabilitation of the aurally handicapped in difficult communication conditions, thereby fully exploiting the possibilities of DSP. The increasing knowledge on the cognitive processes involved in auditory processing is expected to play a major role in this line of research, requiring the development of algorithms which model the higher auditory functions and demanding much higher processing power of future speech processing devices.

Acknowledgments This preparation was supported by the Center of Excellence on Hearing Aid Technology (BMBF 01EZ0212). The author thanks B. Kollmeier and other members of the Medical Physics group for their helpful comments on earlier versions of this manuscript.

References

1. Agnew J and Thornton JM (2000). Just noticeable and objectionable group delays in digital hearing aids. J. Am. Acad. Audiol. 11, 330–336.
2. Moore BCJ (1995) Perceptual consequences of cochlear damage. Oxford University Press, Oxford.
3. Kollmeier B (1999) On the four factors involved in sensorineural hearing loss. In: Dau T, Hohmann V and Kollmeier B (Eds): Psychophysics, physiology and models of hearing. World Scientific, Singapore 211–218.
4. Kollmeier B (Ed) (1996) Psychoacoustics, speech and hearing aids. World Scientific, Singapore.
5. Dillon H (2001) Hearing aids. Boomerang Press, Turramurra, Australia.
6. Kates JM (1998) Signal processing for Hearing Aids. In: Kahrs M and Brandenburg K (Eds): Applications of digital signal processing to audio and acoustics. Kluwer, Dordrecht.
7. Plomp R (1994) Noise, amplification, and compression: Considerations of three main issues in hearing aid design. Ear Hear. 15, 2–12.
8. Turner C and Henry B (2002) Benefits of amplification for speech recognition in background noise. J. Acoust. Soc. Am. 112, 1675–1680.

9. Turner CW and Hurtig RR (1999) Proportional frequency compression of speech for listeners with sensorineural hearing loss. J. Acoust. Soc. Am. 106, 877–886.
10. Sakamoto S, Goto K, Tateno M and Kaga K (2000) Frequency compression hearing aid for severe-to-profound hearing impairments. Auris Nasus Larynx. 27(4), 327–334.
11. Stone MA, Moore BCJ, Alcantara JI and Glasberg BR (1999) Comparison of different forms of compression using wearable digital hearing aids. J. Acoust. Soc. Am. 106, 3603–3619.
12. Verschuure J, Maas AJJ, Stikvoort E, de Jong RM, Goedegebure A and Dreschler WA (1996) Compression and its effect on the speech signal. Ear Hear. 17, 162–175.
13. Jenstad LM, Seewald RC, Cornelisse LE and Shantz J. Comparison of linear gain and wide dynamic range compression hearing aid circuits: Aided speech perception measures. Ear Hear. 20(2), 117–126.
14. Baer T and Moore BCJ (1994) Spectral enhancement to compensate for reduced frequency selectivity. J. Acoust. Soc. Am. 95, 2992.
15. Yang J, Luo FL and Nehorai A (2003) Spectral contrast enhancement: Algorithms and comparisons. Speech Commun. 39(1–2), 33–46.
16. Ephraim Y and Malah D (1985) Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. IEEE Trans. Acoust., Speech, Signal Process. ASSP 33(2), 443–445.
17. Tchorz J and Kollmeier B (2003) SNR estimation based on amplitude modulation analysis with applications to noise suppression. IEEE Trans. Speech Audio Process. 11(3), 184–192.
18. Elko GW and Pong A-TN (1995) A simple adaptive first-order differential microphone. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA.
19. Campbell DR and Shields PW (2003) Speech enhancement using sub-band adaptive Griffiths – Jim signal processing. Speech Commun. 39(1–2), 97–110.
20. Widrow B and Luo FL (2003) Microphone arrays for hearing aids: An overview. Speech Commun. 39(1–2): 139–146.
21. Wittkop T and Hohmann V (2003) Strategy-selective noise reduction for binaural digital hearing aids. Speech Commun. 39, 111–138.
22. Bodden, M (1993) Modeling human sound-source localization and the cocktail-party-effect. Acta Acustica 1, 43–56.
23. Marzinzik M and Kollmeier B (2003) Predicting the subjective quality of noise reduction algorithms for hearing aids. Acta Acustica/Acustica, 89, 521–529.
24. Greenberg JE, Zurek PM and Brantley M (2000). Evaluation of feedback-reduction algorithms for hearing aids. J. Acoust. Soc. Am. 108, 2366–2376.
25. Nordqvist P and Leijon A (2004) An efficient robust sound classification algorithm for hearing aids. J. Acoust. Soc. Am.115(6), 3033–3041.

26. Bregman AS (1990). Auditory scene analysis: The perceptual organization of sound. MIT Press, Cambridge, MA.
27. Nix J, Hohmann V (2007) Combined estimation of spectral envelopes and sound source direction of concurrent voices by multidimensional statistical filtering. IEEE Trans. Audio Speech and Lang. Proc. 15, 995–1008.

PART III
PSYCHOACOUSTICS

Hugo Fastl

Department of Human Machine Communication, Munich University of Technology, Munich, Germany

15 Methods for Psychoacoustics in Relation to Long-Term Sounds

Seiichiro Namba and Sonoko Kuwano 215

Introduction • Methods for Psychoacoustics • Information Processing along the Temporal Stream – The Method of Continuous Judgment by Category • Information Processing in Hearing – Summary

16 Masking and Critical Bands

Bernhard U. Seeber 229

Introduction • Description of Masking Effects • Critical Bands in Psychoacoustical Experiments • Approximations of Auditory Filter Bandwidth • Limits of the Critical Band Concept • Modeling Auditory Masking • Conclusions

17 Aspects of Modeling Pitch Perception

Ernst Terhardt 241

Introduction • Basic Issues • Pitch of Tonal Stimuli • Pitch of Random Stimuli

18 Calculation of Loudness for Normal and Hearing-Impaired Listeners

Josef Chalupper 251

Introduction • Model • Conclusions

19 Psychoacoustical Roughness

Peter Daniel 263

Introduction • Roughness in Dependence of Signal Parameters • Roughness Calculation

15

Methods for Psychoacoustics in Relation to Long-Term Sounds

Seiichiro Namba and Sonoko Kuwano

Osaka University, Osaka, Japan

1 • Introduction

It is well known that auditory signals are effectively transferred to the inner ear as the acoustic energy through the outer ear canal, the ear drum, and the ossicles in the middle ear. It has been made clear using Moessbauer method or using laser radiation that sharp tuning of frequencies is made by active control system on the basilar membrane in the inner ear. Also each cochlear nerve fiber has its characteristic frequency to which it selectively responds. This makes it possible to transfer the selected clues of the frequency to the auditory cortex.

The information of the sound intensity is transferred in terms of the rate of generation of action potentials of auditory nerves. Superior olivary nucleus and inferior colliculus are considered to contribute to the sound localization by detecting the temporal difference in the information transmitted through both ears [1,2].

It is being made clear that the physical properties of sounds such as frequency and intensity and their difference between two ears are processed around peripheral areas in hearing system. However, the physiological mechanisms are not evident how the quantitative evaluation of loudness, the identification of pitch or tonality, or the recognition of special localization or movement of sound sources in our surroundings are achieved. Moreover, it is not explained yet how and where the complex information such as speech, music, and environmental noises is recognized in the auditory system.

Recent progress in brain science in visual perception shows how effectively the information of shapes, colors, and movements is processed and discriminated separately and individually in each specific area in the brain. However, it is still unclear how the information processed separately is integrated into a whole pattern in the central cortex [3].

Since even in visual perception, which is thought to have made a big progress in brain science, there is a long way to solve all the details of the mechanism of perception and cognition from physiological viewpoint, it will take much time to find how all the information in hearing is processed physiologically. At present when there is a black box between physiological findings and information processing of complex phenomena, the methods of psychoacoustics are useful tools to estimate the mechanism of information processing. Even if the brain science makes a drastic development in future, psychoacoustical approach is inevitable in order to find the laws in the subjective world in hearing.

2 • Methods for Psychoacoustics

2.1 Methods for Psychophysics

Psychophysical studies of pitch and loudness of steady-state sounds have a history for more than 100 years. Various thresholds such as absolute threshold of sounds and differential limens of frequency and intensity and point of subjective equality (PSE) of sounds of different frequencies have been measured using conventional psychophysical methods [4]. Psychophysical methods have also been used for the measurement of more complex phenomena than thresholds, such as signal detection under the presence of masking sounds. Information that seems to be processed in the peripheral auditory system can be measured using conventional psychophysical methods, and the relation is being made clear between physical values of sounds and the subjective responses to them. A brief outline of the conventional

psychophysical methods is introduced taking an experiment to measure PSE of pitch or loudness as an example.

2.1.1 Method of Adjustment

This is the method where participants are required to control the frequency or the sound pressure level of the comparison stimulus so that both standard and comparison sounds are equally perceived. Sometimes the experimenter controls the stimulus according to the responses of participants. This method can easily be understood by participants, and usually it is fairly easy for participants to judge the stimuli with this method. On the other hand, it has a disadvantage that there is little reproducibility in this method.

2.1.2 Method of Limits

This is the method where the stimulus is presented by varying the stimulus parameter, such as frequency or sound pressure level, in a fixed step from low to high in ascending series or from high to low in descending series. Participants are required to compare the stimulus pair at each presentation and judge which sound is higher or louder in each pair. All the procedures are settled by the experimenter before starting the experiment and usually the experiment can be conducted in a short period. In experiments, participants can easily guess to which direction the stimulus is varied. This may possibly cause a bias in the judgment.

2.1.3 Method of Constant

This is the method where the stimulus pairs decided by a preliminary experiment are presented several times in random order and participants are required to judge each pair. This method has an advantage that it is free from various constant errors. At the same time, however, a disadvantage of this method is that it takes much time to obtain the data since many stimulus pairs must be judged.

2.1.4 Adaptive Procedure

In the above-mentioned psychophysical methods, the stimuli are presented in the same way as preliminarily decided. In the adaptive procedure, the stimuli to be presented are decided depending on the responses of the participants. The stimulus presentation is controlled by a computer. This procedure has a merit that useless trials can be skipped and it can shorten the time for experiments. On the other hand, experiments using adaptive procedures cannot be done without a computer for each participant and it is not easy to do experiments with a large group of participants together. There are several methods of adaptive procedure, such as “up and down method”, “up and down transformed method”, “parameter estimation by sequential testing”, “maximum likelihood sequential procedure”,

and "hybrid adaptive procedure". There is a special issue of the journal *Perception & Psychophysics* on adaptive procedure [5].

2.1.5 Signal Detection Theory

There are sensory process and decision process when people respond to stimuli. These two processes are in the inner world of people and not visible from the outer world. It is considered that when a stimulus is given, sensory information is transferred to the decision process and a decision is made in reference to expectation, prediction, attitude, and knowledge of participants. In psychophysical methods, the decision process is controlled as much as possible in order to find the relation between stimuli and sensation. In the signal detection theory, it becomes possible to estimate the criterion of the participant's judgments through analysis of experimentally obtainable hit and false alarm rates [6].

2.2 Limitation of Psychophysical Methods

Psychophysical methods are a useful tool, but it has a disadvantage that in principle it is applied only to sounds of short duration. A pair of a standard stimulus and a comparison stimulus is prepared in order to measure threshold or PSE using psychophysical methods. A pair of the sounds is presented successively to participants, and after listening to the pair of the sounds, they are required to compare the sounds and judge whether they are subjectively equal or not. In this case, if the duration of the sounds is long, the impression of the standard sound may become vague after the presentation of the comparison sound and it may become difficult to compare the two sounds. The measurement of threshold is aimed to measure the sensitivity of auditory sensation. However, at the moment when the judgments are made after the presentation of the two sounds, the standard sound belongs to memory and the latter part of the comparison sound belongs to perception. This means that two phenomena that belong to different worlds are compared. This problem remains even if the orders of presentation of the standard and comparison stimuli are reversed. Therefore, generally, the duration of sounds used in psychophysical measurement is limited within about 1 s.

Such problems do not occur when psychophysical methods are applied to the measurement of visual perception. Both standard stimulus and comparison stimulus can be presented simultaneously in a visual field and compared. There is no effect of time order or the duration of stimuli. If necessary, successive comparison is also possible in visual perception taking the afterimage into consideration. It is necessary to control constant errors such as the effect of spatial position where the stimuli are presented, time order error, etc. If the experiment is conducted under the well-controlled condition, psychophysical methods can be applied to visual stimuli with complex patterns.

In hearing, the information is conveyed along the temporal stream. Since the information is expressed in the temporal variation, sounds have complex temporal patterns such as speech, music, and environmental noises and it takes time to be transmitted. The duration of an event sometimes becomes several 10 s as can be found in the case of a phrase of speech, a melody of music, or an occurrence of event such as a flyover of an aircraft.

There is a limitation in the application of conventional psychophysical methods to sounds of long duration since it is related to the interaction between perception and memory. Scaling methods such as rating method and magnitude estimation can be applied to sounds of fairly long duration.

2.3 Scaling Method

There are several scaling methods such as rating method, the method of paid comparison, the method of rank order, and magnitude estimation [7]. Among them, absolute judgments of a stimulus without comparing two stimuli can be made using rating method and magnitude estimation. This suggests that it is possible to measure the impression of fairly long-term sounds using these methods. The outlines of these scaling methods are introduced.

2.3.1 Rating Method

The judgment using rating method is made using category scales. The scales are indicated in various forms, e.g., numerical scales from 1 to 7 and graphical indication using the length of a line. Since the judgment using rating scale is easy, this method can be used for the measurement of wide aspects of our daily life experience; they are, for example, the judgments of aesthetic aspects in experimental aesthetics, the impression in daily life using questionnaires in social surveys and in clinical test methods as well as the impression of visual and auditory stimuli. It is also possible to obtain the judgment of long-term experience in questionnaires [8].

2.3.2 Magnitude Estimation

Magnitude estimation is a method that Stevens [9] used for verifying the power law between physical properties of stimuli and attributes of sensations. In the absolute magnitude estimation, participants are required to judge the impression quantitatively correlating their impression to randomly presented stimuli with numerals. This method can be applied not only to the measurement of sensations but to stimuli whose physical properties are not defined such as the preference of wrist watches [10].

It is true that scaling methods can be used for measuring the impression of long-term stimuli and experiences. However, in

the case of long-term sounds with complex temporal variation, such as music, speech, and environmental noises, there is a problem that it is unclear whether the judgment is based on the overall impression of long-term sounds or whether some specific portions of the sounds have stronger effect on the overall judgment than the other portions.

In order to make clear the laws between factors involved in the stimuli and subjective responses to them, it is not easy to choose what kind of stimuli to be used: short-term sounds that include little information, but can be easily controlled, or long-term sounds that reflect our daily life experience and include much information, but may be difficult to obtain clear effect of each factor.

Scaling methods can also be applied to the judgments of comparison. The method of paired comparison and similarity judgment in multidimensional scaling can give us useful information. A slight difference among stimuli, especially, can be found and the sensory scale can be obtained by comparison. However, there is a limitation that they cannot be applied to stimuli of long duration.

3 • Information Processing along the Temporal Stream – The Method of Continuous Judgment by Category

As described above, sounds convey information along temporal stream. When the method of measurement of psychoacoustics is considered from the viewpoint of information processing, it is not desirable to divide the long-term event of sounds into short segments. This may destroy the frame of reference as a whole event. Conventional psychophysical methods are useful tools to find psychophysical laws between pitch and frequency or between loudness and intensity using short-term sounds. However, these laws show only a part of the information processing in peripheral regions of our auditory system.

It is possible to measure the impression of long-term temporally varying events such as a meaningful phrase of speech and music or a flyover of an airplane using conventional scaling methods. However, since components of these sounds are complex and varying with time, it is not easy to find a clear relationship between stimulus parameters and subjective responses.

The authors have developed a new method called “the method of continuous judgment by category” in order to obtain the

instantaneous impression of temporally varying sounds [11–15]. This method enables us to measure the subjective responses along the temporal stream without destroying the natural steam and meaning of the sounds.

The idea of this method comes from piano performance. Piano performance itself is a kind of response by a pianist, and it is an expression of the musical image in the mind of the pianist by the motion of fingers. This suggests that it would be possible to measure the impression of temporally varying sounds continuously by pressing keys on a keyboard just as is done in the piano performance. It is not seldom that some passages of piano music are played in a speed faster than the critical identification rates [16], i.e., the shortest time where the temporal order of sounds can be identified. Therefore, after an exercise for the smooth motions of fingers such as "Hannon Album" in piano training, it would be possible to measure the impression of sounds continuously as quick as the critical identification rate using a computer keyboard as a response box.

As described before, when only the overall impression of long-term sound is measured using a conventional category rating scale, it is not clear which portions of the long-term sound contribute to the overall impression. In the method of continuous judgment by category, instantaneous impression is measured asking participants to respond continuously by pressing a key. In addition to the continuous judgment, participants are usually required to judge the overall impression of the sound in a questionnaire conducted after the continuous judgment. The relation between the instantaneous impression and the overall impression gives us a clue to find how the long-term sound is processed and the overall impression of the sound is formed from each portion.

Overall impression and instantaneous impression are not independent of each other. The overall impression may be determined by giving a kind of subjective weight to the impression of each portion of the whole stimulus. The impression of each portion is also affected by the frame of reference formed by the overall context and past experience [17].

In the design of conventional psychological experiments, stimuli are presented in random order in order to avoid constant errors. If we assume that a long-term sound, such as music, is divided into short segments and the segments are randomly arranged, it can be easily understood how strange it is to judge the randomly rearranged sound stream continuously. It is important to make clear in which condition the method of continuous judgment is effectively used. That is, this method is used for

measuring the impression of the sounds that convey information along temporal steam in daily life such as speech, music, and environmental noises.

3.1 Introduction

When the instantaneous impression of loudness of a temporally varying sound is judged, for example, participants are required to judge the impression of the loudness at each moment using seven categories from very soft to very loud and press a corresponding key. They need not press the key if their impression does not change. The impression registered on the monitor will remain the same. When their impression of loudness changes, they have to press the appropriate key.

An example of the relation between instantaneous judgments and sound level in the case of traffic noises is shown in Figure 1 [18]. The instantaneous judgments were sampled every 100 ms and the judgments of eight participants were averaged taking the reaction time into account. High correlation can be seen between the instantaneous judgments and sound levels ($r = 0.969$). Another example of the analysis is shown in Figure 2. The instantaneous judgments for the sound level of every 1 dB step were averaged. The average of instantaneous judgments with the standard deviation is plotted against sound levels. Also high correlation can be seen ($r = 0.992$).

This method has many advantages as follows:

1. The judgment is easy for participants.
2. This method can be applied to a fairly long stimulus.
3. Since all the subjective responses are stored in a computer with the information of the timing they press the key, various analyses can be done after the experiment is over.

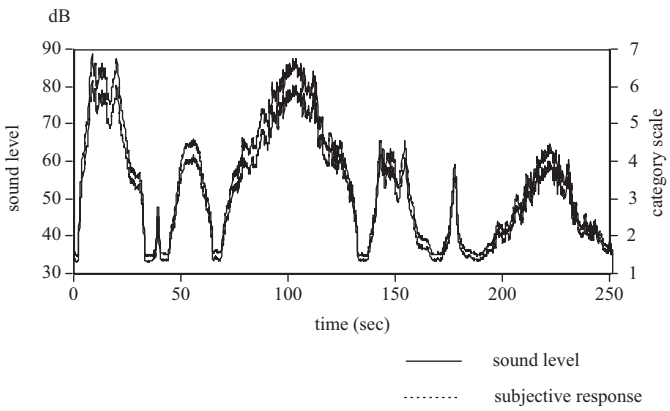


FIGURE 1 An example of the relation between sound level and instantaneous judgments sampled every 100 ms.

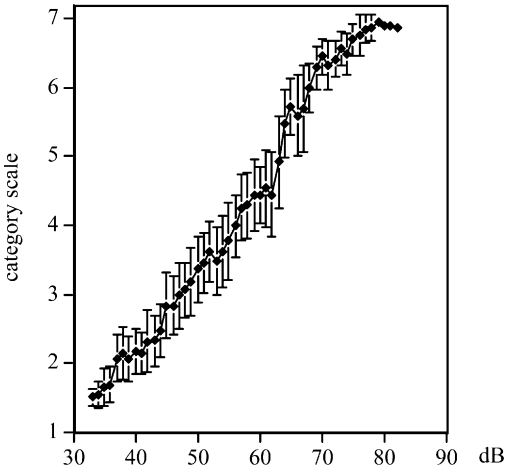


FIGURE 2 An example of the instantaneous judgments averaged for sound level of every 1 dB.

4. The relation between instantaneous judgments and physical values can give clues for understanding temporal characteristics of hearing.
5. The relation between overall impression and instantaneous impression can be compared and this makes it possible to investigate how the overall impression is determined by the impression of each portion.
6. In conventional psychophysical methods, participants have to pay attention to the stimulus to be judged. On the other hand, in experiments using this method, participants judge the impression along the temporal stream, i.e., in the same context, without paying attention to a specific portion of the long sound, and the impression of each portion of the sound can be obtained as in their daily life situations.
7. Multidimensional judgment is also possible using the method of continuous judgment by selected description [14,19].

3.2 Examples

3.2.1 Estimation of Psychological Present

In the continuous judgments, the judgment may possibly be affected by the preceding portions of the stimulus as well as the immediate portion. In order to find the duration of the preceding portions which may affect judgment at each moment, the sound energy was averaged for various durations and correlated with the instantaneous judgment [12]. From this it was found that the instantaneous judgment showed the highest correlation when the preceding sound energy was averaged for 2.5 s. This value may

reflect the psychological present [20]. The physical present is a point, but the psychological present is not and may have a certain duration during which subjective impressions are integrated and perceived as a unit. If the coefficient of correlation between the integrated physical values and the instantaneous judgments is taken as an index of the psychological present, the integration time which shows the highest correlation with the instantaneous judgment may represent the duration of the psychological present.

3.2.2 Relation Between Overall Judgment and Instantaneous Judgment

For the evaluation of long-term temporally varying sounds, it would be helpful to measure instantaneous impression and overall impression as well. The time interval between the presentation of sounds and the overall judgment should be chosen carefully. As time passes, the impression of the prominent portion becomes stronger and that of the less prominent portion becomes weaker [12]. It is natural that loud sounds or loud portions easily become prominent, but it was found that even soft sounds such as bird twittering can be prominent and contribute to the overall impression [21]. It is interesting and important to find the factors which contribute to the overall impression.

3.2.3 Habituation

With the lapse of time, subjective impression of sounds may possibly change. Scharf [22] has indicated that with low-level sounds and high-frequency sounds loudness adaptation occurs, and the impression of loudness decreases when the sounds are presented for a longer time.

Habituation to sounds is the phenomenon where, when a certain stimulus is repeatedly presented to the auditory organ, the response to the stimulus gradually diminishes and finally no response will occur. Habituation is different from adaptation or fatigue in the point that it is easy to overcome habituation when attention is directed to the stimulus. Habituation has a positive meaning to help people cope with the environment. It is not easy to measure habituation by conventional psychological methods where participants have to pay attention in order to judge the stimulus. In experiments using the method of continuous judgment by category, participants are not necessarily forced to pay attention to the specific stimuli all the time. This makes it possible to approach a measurement of habituation.

In an experiment [13], participants were engaged on a mental task and they were highly motivated to do the task. While occupied with the mental task, they were asked to judge the noisiness of sounds continuously, using the method of

continuous judgment by category. The sounds were of various kinds, such as road traffic noise, voices of street vendors, TV talk, and sounds from a washing machine and a vacuum cleaner. The duration of an experimental session was about one hour and a half. In this experiment when participants did not respond to the sounds for 30 s, they were given a signal on a monitor to respond. Participants did not always respond to the sounds in spite of the warning signal when they were devoted to the mental task. The length of time when they did not respond even if the warning signal was presented was used as an index of habituation. There was a high correlation between two trials in the habituation period for each participant. It was also found that the group of participants who showed long period of habitation said, in answer to a questionnaire, that they became habituated to noise fairly easily in daily life situations.

Another index of habituation is the number of responses made [23,24]. When participants pay attention to sounds, they respond to each change in their impression of sounds. On the other hand, when they are habituated to sounds, they are not conscious of the change and do not respond.

3.2.4 Trade-Off Effect in Subjective Responses among Stimulus Conditions

Taking the interior noise of a car, as an example, various sounds are audible in a car while driving, such as engine noise, air conditioner noise, winker signals, sounds from wipers, and road noise. The loudness of inside car noise while driving was judged using the method of continuous judgment by category [25]. The visual scenery while driving was also presented with the sound. This made the participants feel as if they were driving and could prevent participants from paying special attention to sounds. After the experiment, the averages of instantaneous judgments for each sound source were compared with background sounds. It was found that road noise when driving on a rough road is judged louder than the other portions in the same sound level conditions, which suggests that some countermeasures for noise due to road conditions are needed. It was also found in this experiment that air conditioner noise was judged louder than the other portions. This suggests that it is important to improve the air conditioner noise in order to make the cabin of a car comfortable in the case of the car used for the experiment.

3.3 Software

When the method of continuous judgment by category was developed in the 1970s, personal computers were not available. The responses of continuous judgments were recorded using a level recorder and the data at every 100 ms were measured

manually. Then the data were punched on IBM cards, and using a large-size computer, the coefficient of correlation between physical parameters of stimuli and instantaneous responses was calculated. It was very laborious work.

Some years later, the experiments were conducted using a personal computer. The software was made by the authors with MS-DOS. This made it possible to analyze data from various viewpoints.

Now the new software for the method of continuous judgment by category for computers with OS of Windows is available [26]. The data can be analyzed very quickly, and the results can be obtained immediately after the experiment is over. The examples shown in Figures 1 and 2 were the results of the analysis using this software. It is also possible to analyze each portion of the sound separately. This software has Japanese and English versions and is commercially available.

4 • Information Processing in Hearing – Summary

In conventional psychophysical methods, usually a sound of short duration is used and absolute threshold or differential limen of frequency or intensity is measured. These measurements provide us valuable information about the basic characteristics of hearing. The signal processing in hearing in the peripheral area is becoming clear combining the results of psychophysical experiments and physiological findings.

However, temporal information processing is not considered in conventional psychophysical experiments, and important aspects of hearing to convey information along temporal stream are neglected. It is important to make clear how the information is processed in hearing along temporal stream in order to understand the mechanism of hearing to temporally varying sounds such as music, speech, and environmental sounds.

No one may deny the importance of ecological validity in hearing research. There is no one who underestimates the importance of precise control of stimulus conditions in experiments either. Though both aspects are important, it is difficult to fulfill both at the same time.

In this chapter, the method of continuous judgment by category is introduced that makes it possible to control physical parameters of sounds continuously and measure quantitatively the instantaneous impression of temporally varying sounds along the temporal stream. The main findings using this method are introduced in relation to information processing in hearing; they are the estimation of psychological present (i.e., psychological time

window), the relation between overall impression and instantaneous impression, habituation to noise and trade-off effect among stimulus conditions. A newly developed software for the experiment using the method of continuous judgment by category is also introduced.

The information processing in hearing covers wide areas from short-term process to long-term experience by adjusting the frame of reference flexibly to the situation. It is expected that the information processing in hearing will be made clear in future investigations along the temporal stream.

References

1. K. Maki, K. Ito and M. Akagi, "Time–frequency response patterns of nerve firings in primary stage of auditory system", *The Journal of the Acoustical Society of Japan*, 59 (1), 52–58 (2003).
2. B. C. J. Moore, *Hearing*, 5th ed., Academic Press, 2003.
3. R. L. Gregory, *Eye and Brain: The Psychology of Seeing*, Oxford University Press, 1998.
4. G. A. Gescheider, *Psychophysics: Method, Theory, and Application*, 3rd ed., Lawrence Erlbaum Associates, Inc., 1997.
5. N. A. Macmillan (Ed.), "Psychometric functions and adaptive methods", special issue, *Perception & Psychophysics*, 63 (8), 1277–1455 (2001).
6. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*, Wiley, 1966.
7. S. Namba and S. Kuwano, *Method of Psychological Measurement for Hearing Research*, Corona-sha, 1998.
8. S. Namba, J. Igarashi, S. Kuwano, K. Kuno, M. Sasaki, H. Tachibana, A. Tamura and Y. Mishina, "Report of the Committee of the Social Survey on Noise Problems", *The Journal of the Acoustical Society of Japan (E)*, 17, 109–113 (1996).
9. S. S. Stevens, *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects*, Wiley, 1975.
10. T. Indow, "An example of motivation research applied to product design", *Chosa to Gijutsu*, 102, 445–460 (1961).
11. S. Namba and S. Kuwano, "The relation between overall noisiness and instantaneous judgment of noise and the effect of background noise level on noisiness", *The Journal of the Acoustical Society of Japan (E)*, 1, 99–106 (1980).
12. S. Kuwano and S. Namba, "Continuous judgment of level-fluctuating sounds and the relationship between overall loudness and instantaneous loudness", *Psychological Research*, 47, 27–37 (1985).
13. S. Namba and S. Kuwano, "Measurement of habituation to noise using the method of continuous judgment by category", *Journal of Sound and Vibration*, 127, 507–511 (1988).

14. S. Namba, S. Kuwano and M. Koyasu, "The measurement of temporal stream of hearing by continuous judgments – in the case of the evaluation of helicopter noise", *The Journal of the Acoustical Society of Japan (E)*, 14, 341–352 (1993).
15. S. Kuwano, "Continuous judgment of temporally fluctuating sounds", in H. Fastl, S. Kuwano and A. Schick (Eds.), *Recent Trends in Hearing Research*, BIS, 1996, pp. 193–214.
16. R. Teranishi, "Critical rate for identification and information capacity in hearing system", *The Journal of the Acoustical Society of Japan*, 33, 136–143 (1977).
17. H. Helson, *Adaptation Level Theory*, Harper & Row, 1964.
18. S. Namba and S. Kuwano, "The continuous judgment of traffic noises from mixed sources" (unpublished data in 2003).
19. S. Namba, S. Kuwano, T. Hato and M. Kato, "Assessment of musical performance by using the method of continuous judgment by selected description", *Music Perception*, 8, 251–276 (1991).
20. P. Fraisse, *Psychologie du Temps*, Press University, 1957 (Japanese translation by Y. Hara, revised by K. Sato).
21. S. Kuwano, S. Namba, T. Kato and J. Hellbrueck, "Memory of the loudness of sounds in relation to overall impression", *Acoustical Science and Technology*, 24 (4), 194–196 (2003).
22. B. Scharf, "Loudness adaptation", in J. B. Tobias and E. D. Schubert (Eds.), *Hearing Research and Theory*, Vol. 2, Academic Press, 1983, pp. 1–57.
23. S. Namba, S. Kuwano and A. Kinoshita, "Measurement of habituation to noise", *Proceedings of the Autumn Meeting of the Acoustical Society of Japan*, 759–760 (1994).
24. N. Hato, S. Kuwano and S. Namba, "A measurement method of car interior noise habituation", *Transactions of Technical Committee on Noise, The Acoustical Society of Japan*, N95-48, 1–4 (1995).
25. S. Kuwano, S. Namba and Y. Hayakawa, "Comparison of the loudness of inside car noises from various sound sources in the same context", *The Journal of the Acoustical Society of Japan (E)*, 18, 189–193 (1997).
26. S. Namba, S. Kuwano, H. Fastl, T. Kato, J. Kaku and K. Nomachi, "Estimation of reaction time in continuous judgment", *Proceedings of the International Congress on Acoustics*, 1093–1096 (2004).

16

Masking and Critical Bands

Bernhard U. Seeber

Auditory Perception Lab, Department of Psychology, University of California at Berkeley, 3210 Tolman Hall #1650, Berkeley, CA 94720-1650, USA, e-mail: bernhard@bseeber.de

1 • Introduction

Auditory masking is often observed in everyday life: While walking on a street a passing truck can disrupt our conversation. Likewise, one would not want a cell phone to ring in quiet parts of a classical concert but it might not even be heard in a loud rock-concert. Components of one sound interact with components of another sound similar in frequency and time and render them inaudible. This is referred to as masking. Partial masking can also occur: Components are not inaudible, but their loudness is reduced (c.f. Chapter 18). This chapter first introduces the frequent case of masking by one tone on another for simultaneous and successive presentation.

Many spectral masking effects can be described by comparing signal power, integrated over a certain frequency region, with and without the probe present. Detection of changes in energy in the bandpass-filtered signal forms the basis of the critical band concept. The critical band (CB) is defined as the rectangular bandwidth of the filter. Section 3 presents a collection of experiments which reveal the influence of critical band filters.

It turns out that CBs observed in psychophysical masking experiments grow in the same way as the physiologically measured distance on the basilar membrane or the psychophysical scales of just-audible pitch steps or of ratio pitch. Scales of psychophysically derived CBs are presented in Section 4 whereas Section 5 discusses the limits of the CB-concept. The last section gives a brief introduction to models of masking.

2 • Description of Masking Effects

2.1 Simultaneous Masking

In a simultaneous masking experiment the level of a probe signal is varied in the presence of a masking sound to find the threshold at which the probe is just audible. Three variations for presenting the results are common: (1) *tuning curves* describe the level of a masker necessary to mask a probe of fixed frequency and low level, (2) *masking growth functions* show the dependence of probe level on masker level for fixed probe and masker frequencies, and (3) *masking patterns* depict the level of probes of different frequencies that are just audible in the presence of a masker of fixed level and frequency.

Figure 1 presents *masking patterns* of a tonal masker for simultaneously presented probe tones. It is immediately evident that the level of the test tone (L_T) has to be raised to assure audibility as its frequency (f_T) approaches that of the masker (f_M), i.e., a tonal masker produces more masking in its spectral vicinity. Masking even occurs for very low masker levels (L_M) but the effect is spread widely across frequencies for higher levels.

The slopes of the masking pattern depend heavily on masker level. To illustrate this non-linearity Figure 1 also shows masking patterns mirrored at the masker frequency. For tonal maskers of a level $L_M \approx 40$ dB SPL the pattern is almost symmetrical. If the masker level is lower than 40 dB SPL the pattern tilts so that the low-frequency slope is shallower than the high-frequency slope. For masker levels above 40 dB SPL the opposite occurs: The low-frequency slope becomes somewhat steeper whereas the slope on the high-frequency side becomes far more shallow. Masking spreads out to frequencies far above the masker frequency, which is called the *upward spread of masking*. The excess in masking between the mirrored high- and the low-frequency side is labeled "A" in Figure 1. Additionally, the frequency of maximum masking can shift upward from the masker frequency.

The way masking (L_T) grows with increasing masker level (L_M) can also be captured from Figure 1. At $f_T = f_M = 1{,}000$ Hz the different masking patterns are about 10 dB apart – equal to the increase in masker level. More precisely, masking increases from

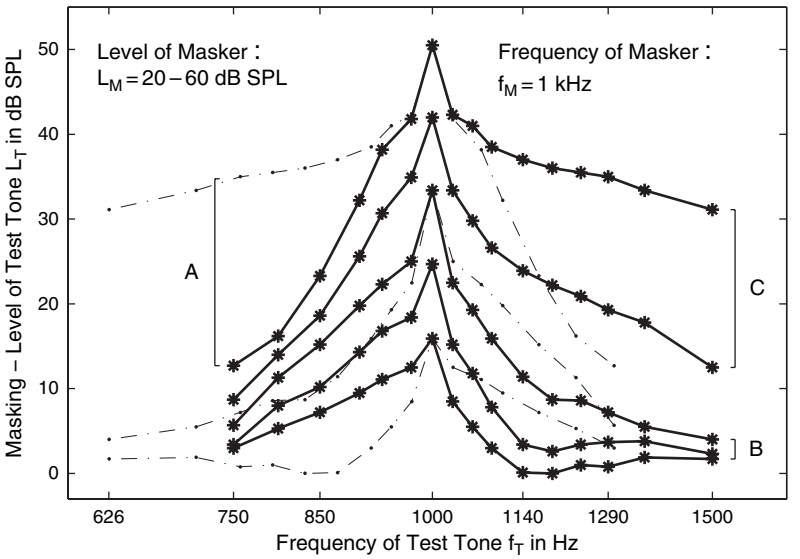


FIGURE 1 Tone-on-tone masking patterns for a simultaneous masker at $f_M = 1{,}000$ Hz with a level $L_M = 20$–60 dB SPL in 10 dB steps. The patterns for maskers at 20, 40 and 60 dB SPL are also plotted mirrored in frequency on the Bark scale (Section 4.1) to emphasize the non-linearity of the pattern (*dashed dotted line*). Data adopted from [28].

15.9 to 50.5 dB over a 40 dB change of L_M from 20 to 60 dB SPL. The small deviation from a linear increase is called the *"near miss" to Weber's law* and shows a ratio of 34.6 dB/40 dB ≈ 0.87.

On the high-frequency side masking shows a strong non-linear dependence on masker level. Unless the masker level exceeds a certain threshold no masking is present. This is labeled "B" in Figure 1 and evident for $f_T = 1{,}500$ Hz and $L_M \leq 40$ dB SPL. If the masker level exceeds this threshold masking occurs and it grows far stronger than linear with a ratio of 2–3 on average. In Figure 1, depicted with "C", masking increases from 12.5 to 31.1 dB SPL for a 10 dB change in masker level from 50 to 60 dB SPL. This equals a masking growth ratio of 1.9. Masking can even grow with a ratio of 6, i.e., for a 1 dB increase in masker level the probe must be raised by 6 dB.

The strong non-linear effect of the upward spread of masking can be exploited in audio compression systems. High compression ratios are possible for known playback levels as few strong components can produce masking over a wide frequency range. It should be mentioned that the measurement of tone-on-tone masking patterns is complicated by beats between

the tones. Difference tones originating in the non-linearity of the basilar membrane can be masked with noise which was done to obtain the data in Figure 1 [13,24,27,28].

2.2 Simultaneous Masking by a Narrow Band of Noise

A narrow band of noise appears to produce stronger masking and more regular masking patterns than a single tone. Pronounced differences occur at the low-frequency slope of the masking pattern of the noise masker which appears less steep and less dependent on level with a constant slope of about 27 dB/Bark. This can be attributed to additional masking produced by difference tones in the auditory system. The tip region is broadened by the bandwidth of the noise and the influence of temporal beats is reduced. The high-frequency side shows similar upward spread of masking for noise and tonal maskers. Several tones spaced on and between the corner frequencies of the noise are sufficient to reproduce the masking patterns of the noise [5,13,27].

2.3 Temporal Masking Effects

Forward masking or *postmasking* is measured with a brief probe following the masker. It shows the decay of masking at the end of the masker. Forward masking is a non-linear effect depending on the masker frequency, its duration and level, the time between masker and probe and the probe frequency. Additional nearby components can reduce forward masking (suppression).

Tuning curves show the masker level necessary to mask a probe of fixed frequency and level. They are thought to resemble the shape of auditory filters as the listener is likely to always use the same filter to detect a faint probe of fixed frequency. Figure 2 compares simultaneous and forward masking tuning curves. At first sight forward masking tuning curves form a more regular pattern due to the absence of beats and difference tones. Except for the very tip region of the tuning curve the lower and upper slopes are steeper in forward masking. This indicates a higher frequency selectivity in forward than in simultaneous masking which is attributed mostly to suppression.

Forward masking will last only a few milliseconds for very short masker durations, but pronounced forward masking occurs if the masker duration is longer than 200 ms. Forward masking decays little within 5 ms after the end of the masker and ceases after about 100–200 ms relatively independent of masker level. This requires a faster decay of masking for higher masker levels which is reflected in a less than linear growth of masking with increasing masker level at the masker frequency ($f_M = f_T$). While masking increases about linearly with masker level in simultaneous masking, the growth ratio becomes smaller with

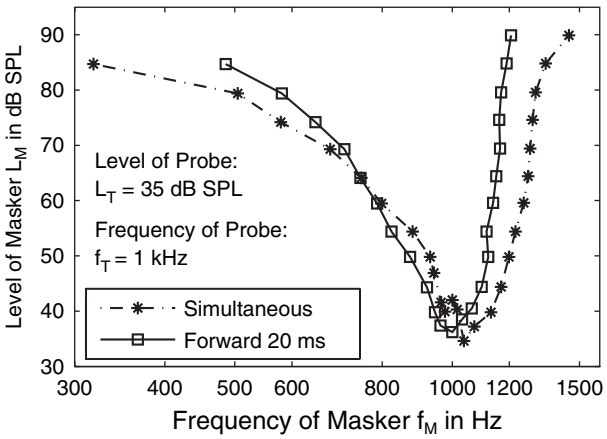


FIGURE 2 Psychophysical tuning curves for simultaneous and forward masking for a brief probe tone at $f_T = 1{,}000$ Hz with a level of $L_T = 35$ dB SPL. The tuning curves represent the level of a tonal masker required to keep the probe tone at the threshold of audibility. Data adopted from Figure 4 in [22].

increasing delay time in forward masking: 0.82, 0.5 and 0.29 for 0, 10, 30 ms delay. However, instead of showing the upward spread of masking for $f_T > f_M$, a linear growth of masking with increasing masker level is seen in forward masking [6,15].

In *premasking* or *backward masking* the probe precedes the masker. Although the masker appears later it can mask a preceding low-level probe within about 20 ms, but the effect is small and subject to high variance.

The tuning curves in Figure 2 show strong frequency selectivity. In the following section the idea of critical bands is developed based on the frequency selectivity seen in several psychoacoustical experiments.

3 • Critical Bands in Psychoacoustical Experiments

3.1 Band Widening

If the bandwidth of noise with constant spectral density is widened, the threshold of a tone masked by the noise will increase markedly less beyond a certain noise bandwidth, even though the overall level of the noise still increases. It seems that the noise power is spectrally integrated toward masking the tone up to this "critical" bandwidth and the energy outside this frequency band is disregarded. The band-widening experiment is the original measurement of the critical bandwidth [7]. Measurement difficulties arise from (1) the inherent modulation of the narrow-band noise masker which forms an additional detection cue besides

energy, and (2) from the limited contribution of energy from the slopes of the non-rectangular auditory filters (e.g., [10,13,17]).

In a variation of the band-widening experiment a complex harmonic tone is detected as a function of it's number of harmonics. Since every doubling of the number of harmonics increases the level of the complex by 3 dB the detection threshold decreases by 3 dB if components fall within a CB, but it remains fairly constant if components are processed outside the CB. As CB-filters are not of rectangular shape it can be difficult to estimate the corner frequency. The experiment can be used to measure CBs at different levels by playing *uniform masking noise* as a background masker. Uniform masking noise is filtered to produce the same masked threshold at all frequencies [27].

The observation of critical bands likely stemmed from another effect: If the bandwidth of white noise is increased but the overall level is kept constant, the loudness will remain constant up to the critical bandwidth, but increase thereafter ([27], see Chapter 18).

3.2 Masking in Spectral Notches

If the ear separates signals into critical bands then the threshold of a tone should be less affected if flanking energy can be processed separately in adjacent CBs. In the *two-tone masking* experiment the spectral separation of two flanking tones is varied around a narrow band of noise [27]. The threshold for detecting the noise is low as long as the flanking tones are outside the CB and it rises when the tones enter the CB of the noise. Difference tones complicate the measurement.

One way to reduce the influence of difference tones is to use flanking bands of noise as a masker and a tone as the probe. The method produces slightly smaller CBs than the above methods. The *notched-noise method* reduces off-frequency listening since it forces the listener to base tone-detection on energy in the defined region of the spectral notch [17].

The influence of CBs can also be seen in the sensitivity to phase differences between tones which becomes apparent in AM and FM detection experiments. The isolated audibility of partials in harmonic complex tones can also be explained by the CB-concept.

4 • Approximations of Auditory Filter Bandwidth

4.1 Critical Bands

Zwicker averaged values of critical bandwidth obtained from several different methods and published them in tabulated form [25]. The data can be seen in Figure 3 along with their numerical approximation after Zwicker and Terhardt [29]:

$$\Delta f_{\mathrm{G}}(f)\ [\mathrm{Hz}] = 25 + 75\left(1 + 1.4\,(f\ [\mathrm{kHz}])^2\right)^{0.69}. \qquad (1)$$

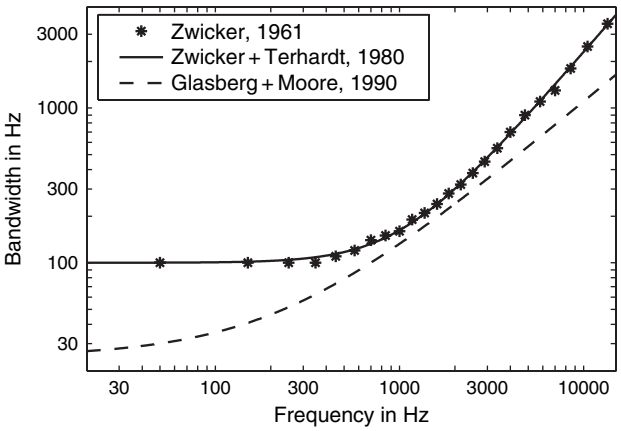


FIGURE 3 Bandwidth of psychophysically measured auditory filters: critical bandwidth according to the Bark-table published by Zwicker [25], "*", and according to the formula by Zwicker and Terhardt [29], "—", as well as the equivalent rectangular bandwidth (ERB) after a formula proposed by Glasberg and Moore [8], "- -".

The critical bands of Zwicker show a constant bandwidth of 100 Hz for $f < 500$ Hz and an approximately linear increase in bandwidth with $0.2f$ for $f > 500$ Hz. If 24 abutting CBs are lined up the *critical band rate scale* is formed which covers the audible frequency range. The unit is 1 Bark. An approximation of the critical band rate scale is given in [27,29]:

$$z(f)\ [\text{Bark}] = 13 \arctan(0.76 f\ [\text{kHz}]) + 3.5 \arctan\left(\frac{f\ [\text{kHz}]}{7.5}\right)^2. \tag{2}$$

4.2 Equivalent Rectangular Bandwidth

Glasberg and Moore [8] suggested a formula for the auditory filter bandwidth based entirely on notched-noise masking data. This bandwidth definition is commonly referred to as the *equivalent rectangular bandwidth (ERB)*. The following formulas give the ERB-bandwidth $B_{ERB}(f)$ as well as the derived number ("rate") scale $ERB_N(f)$ [8,13]:

$$B_{ERB}(f)\ [\text{Hz}] = 24.7\,(4.37 f + 1), \tag{3}$$

$$ERB_N(f)\ [\text{ERB}] = 21.4 \log_{10}(4.37 f + 1). \tag{4}$$

Software that implements all formulas can be found in [20]. When comparing Zwicker's critical bands and ERBs in Figure 3 it

becomes obvious that ERBs yield smaller values at all frequencies and that they decline for $f < 500$ Hz. With respect to auditory models CBs tend to be somewhat large at low frequencies whereas ERBs appear to be slightly small, e.g., in comparison with the filter bandwidth necessary to model notched-noise masking data [18]. The differences stem largely and somewhat from the use of different methods and somewhat from measuring at different levels. As auditory filters widen with increasing level some methods show a level dependence of the measured critical bandwidth.

5 • Limits of the Critical Band Concept

5.1 Temporal Effects and Influence of Non-linearities

Since the critical band concept is based on energy detection in spectral channels it can be applied only if perception is not dominated by temporal effects. Specifically, the following temporal effects can influence detection of masked components: (1) Beats produced by the interaction of the probe tone with masker components can be detected as temporal fluctuations. (2) The phase of spectral components strongly influences the temporal waveform of the signal. The auditory system is able to listen into temporal gaps of signals with high peak factors (dip-listening). The long-term spectrum is not sufficient to predict masking of those signals and a short-term spectral analysis has to be used. (3) Masking shows an overshoot for probes presented at the onset of the masker, i.e., the probe is harder to detect if it is switched on in synchrony with the masker. The *overshoot effect* is maximal for wide-band maskers of medium levels. The additional masking can amount to 10–15 dB. This huge effect might be explained by a spectral sharpening of the active basilar membrane filtering process in combination with central effects [13,26].

The peripheral filtering process takes place on the basilar membrane and is clearly visible in the auditory nerve [19]. It forms the basis for CBs and a first stage of auditory models. To accommodate for many temporal masking effects the static energy detectors monitoring the output of CB-filters can be extended to follow dynamic changes in energy.

Although not a limitation of the CB-concept per se it should be noted that modeling peripheral frequency selection by linear filters becomes a simplification if the non-linearity of the basilar membrane response is involved. The non-linearity can introduce distortion products which may be heard before the probe tone itself becomes audible in masking experiments. A combination of peripheral and more central non-linear effects is likely responsible for non-linear addition of masking of two maskers in

two-tone masking and suppression experiments. Finally, filter shape and bandwidth of peripheral filters are level dependent which somewhat influences effective critical bandwidth.

5.2 Comodulation Masking Release

If a masker is accompanied by flanking maskers in frequency bands away from the target band the detection threshold of a tone can decrease substantially if all maskers share the same low-frequency modulation. The effect of comodulation masking release (CMR) can be obtained over a wide frequency range and it can reach 1–6 dB for large and up to 16 dB for small spectral separations between the target and masker bands. The effect is maximal for low modulation frequencies and wide-band maskers [11]. CMR can be modeled either by comparing modulation across CBs or by assuming that the auditory system is able to listen into temporal gaps in the on-frequency band if the time instances for dip-listening are identified from flanking bands [2].

5.3 Profile Analysis

The detection of amplitude changes of a target component can improve in situations of stimulus uncertainty if the target is embedded in a background complex tone with multiple components [9,11]. It appears as if detection is facilitated by across-channel comparison of target level against background level. Profile analysis is linked to timbre perception because the effect degrades if the spectral extent of the background components is reduced or if the background deviates from a flat and dense spectral distribution.

5.4 Informational and Binaural Masking

Although this article is concerned with energetic masking it should be noted that several recent studies focused on the contribution of more central masking effects. Informational masking is commonly equated as masking that stems from attentional interference of sources or from stimulus uncertainty. It can be seen as masking that is not of peripheral, energetic origin [4]. Inter-subject variability is far higher in informational than energetic masking. Pronounced informational masking can be seen for speech sources (e.g., [1]).

When signals are played to both ears the masked threshold is dependent on the interaural phase difference between masker and probe, especially at low frequencies. If masker and probe are presented interaurally out of phase the threshold can be 15 dB lower than if they were presented in phase. The gain in detectability is called the *binaural masking level difference* [21].

6 • Modeling Auditory Masking

In most models of masking the detectability of signal components is derived from changes in the time- and frequency-dependent *excitation pattern*. Excitation patterns resemble a pattern close to masking patterns. They reflect the influence of the auditory periphery which analyzes the signal in frequency bands. Several functions for linear peripheral filters have been suggested. *Gammatone filters* consider the filter impulse response and are most widely used [16]. A simple, but powerful model to generate excitation patterns uses steep filters to compute the main excitation and applies slopes from normalized masking patterns to it. Frequency- and level-dependent non-linearities can be built into the lookup table [20,27].

Another approach builds on non-linear filterbanks which requires modeling the level-dependent compression and bandwidth in the filter center as well as the dependence of both tails on level (e.g., [12,18]). Temporal masking can be modeled by introducing compression and temporal integration (e.g., [3,14]). Inter-connectivity between channels for controlling channel-wise compression is necessary for simulating lateral suppression. Current modeling efforts aim to explain the non-linear behavior in simultaneous and forward masking. More general approaches will also model suppression effects, difference tones and the selection of spectral or temporal decision criteria.

7 • Conclusions

Since Helmholtz suggested in 1863 to view the cochlea as a series of resonators the filterbank approach to model spectral analysis became most widely accepted in auditory research [23]. The critical band concept adds the notion that masking can be predicted solely from energy comparisons at the output of these filters. The concept has proven successful in that numerous experimental results support this view. To predict results that cannot be explained directly by the CB concept, e.g., results on comodulation masking release or loudness, the critical band-based filterbank approach provides a starting point if it is followed by across-channel processes. Future models of masking have to be evaluated critically against the existing non-linear dynamic models to assess the contribution of other detection cues besides energy.

Acknowledgments I would like to thank Prof. H. Fastl, Dr. B. Edwards, and one anonymous reviewer for many helpful comments on previous versions of the manuscript.

References

1. D.S. Brungart, B.D. Simpson, C.J. Darwin, T.L. Arbogast, and G. Kidd. Across-ear interference from parametrically degraded synthetic speech signals in a dichotic cocktail-party listening task. *J. Acoust. Soc. Am.*, 118(5):292–304, 2005.
2. S. Buus. Release from masking caused by envelope fluctuations. *J. Acoust. Soc. Am.*, 78(6):1958–1965, 1985.
3. T. Dau, D. Püschel, and A. Kohlrausch. A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J. Acoust. Soc. Am.*, 99(6):3615–3622, 1996.
4. N.I. Durlach, C.R. Mason, G. Kidd, T.L. Arbogast, H.S. Colburn, and B.G. Shinn-Cunningham. Note on informational masking. *J. Acoust. Soc. Am.*, 113(6):2984–2987, 2003.
5. J.P. Egan and H.W. Hake. On the masking patterns of a simple auditory stimulus. *J. Acoust. Soc. Am.*, 22(5):622–630, 1950.
6. H. Fastl. Temporal masking effects: III. Pure tone masker. *Acustica*, 43:282–294, 1979.
7. H. Fletcher. Auditory patterns. *Rev. Mod. Phys.*, 12:47–65, 1940.
8. B.R. Glasberg and B.C.J. Moore. Derivation of auditory filter shapes from notched-noise data. *Hear. Res.*, 47(5):103–138, 1990.
9. D.M. Green. *Profile Analysis*. Oxford University Press, Oxford, 1988.
10. D.D. Greenwood. Auditory masking and the critical band. *J. Acoust. Soc. Am.*, 33(4):484–502, 1961.
11. J.W. Hall III, J.H. Grose, and L. Mendoza. Across-channel processes in masking. In B.C.J. Moore, editor, *Hearing*, Handbook of perception and cognition, pages 243–265. Academic Press, San Diego, 1995.
12. E.A. Lopez-Poveda and R. Meddis. A human nonlinear cochlear filterbank. *J. Acoust. Soc. Am.*, 110(6):3107–3118, 2001.
13. B.C.J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, San Diego, 5th edition, 2003.
14. A.J. Oxenham and B.C.J. Moore. Modeling the additivity of nonsimultaneous masking. *Hear. Res.*, 80:105–118, 1994.
15. A.J. Oxenham and C.J. Plack. Effects of masker frequency and duration in forward masking: further evidence for the influence of peripheral nonlinearity. *Hear. Res.*, 150:258–266, 2000.
16. R.D. Patterson, J. Holdsworth, I. Nimmo-Smith, and P. Rice. The auditory filterbank. SVOS final report, APU: 2341, 1988.
17. R.D. Patterson and B.C. Moore. Auditory filters and excitation patterns as representations of frequency resolution. In B.C. Moore, editor, *Frequency Selectivity in Hearing*, pages 123–177. Academic Press, London, 1986.
18. R.D. Patterson, M. Unoki, and T. Irino. Extending the domain of center frequencies for the compressive gammachirp auditory filter. *J. Acoust. Soc. Am.*, 114(3):1529–1542, 2003.
19. J.O. Pickles. The neurophysiological basis of frequency selectivity. In B.C. Moore, editor, *Frequency Selectivity in Hearing*, pages 51–121. Academic Press, London, 1986.

20. B. Seeber. Software on masking and critical bands. www.acoustics.bseeber.de.
21. S. van de Par and A. Kohlrausch. Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *J. Acoust. Soc. Am.*, 106(4):1940–1947, 1999.
22. L.L.M. Vogten. Low-level pure-tone masking: a comparison of "tuning curves" obtained with simultaneous and forward masking. *J. Acoust. Soc. Am.*, 63(5):1520–1527, 1978.
23. H. von Helmholtz. *Die Lehre von der Tonempfindung als physiologische Grundlage für die Theorie der Musik*. Vieweg Verlag, Braunschweig, 1863.
24. R.L. Wegel and C.E. Lane. The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Phys. Rev.*, 23(2):43–56, 1924.
25. E. Zwicker. Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *J. Acoust. Soc. Am.*, 33(3):248, 1961.
26. E. Zwicker. Temporal effects in simultaneous masking by white-noise bursts. *J. Acoust. Soc. Am.*, 37:653–663, 1965.
27. E. Zwicker and H. Fastl. *Psychoacoustics, Facts and Models*. Springer, Berlin Heidelberg New York, 2nd edition, 1999.
28. E. Zwicker and A. Jaroszewski. Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels. *J. Acoust. Soc. Am.*, 71(6):1508–1512, 1982.
29. E. Zwicker and E. Terhardt. Analytical expressions for critical band rate and critical bandwidth as a function of frequency. *J. Acoust. Soc. Am.*, 68(6):1523–1525, 1980.

17

Aspects of Modeling Pitch Perception

Ernst Terhardt

Technical University of Munich, Munich, Germany,
e-mail: teirhardt@ei.tum.de

1 • Introduction

There is hardly any topic of psychoacoustics the approaches to which have been as diverse and controversial as the topic of modeling pitch perception. In the present contribution a number of basic aspects of this problem are discussed, and a framework for its solution is outlined.

2 • Basic Issues

Regarding psychophysical processes such as the perception of pitch, one can make a sensible distinction between a *theory* and a *model*. A theory is a formal description of the stimulus-to-sensation relationships that have become known from observations and experiments. The theory of pitch is – or should be – confined to an algorithmic description of the relationships that exist between the "observable variables" involved, i.e., the parameters of the stimulus, and the perceived pitches. Inclusion of hypothetical "mechanisms" should be avoided.

A *model* is supposed to simulate the perceptual process; it relates the stimulus itself, i.e., the sound wave, to the perceived pitches. If the theory exists, a model can be obtained by adding

an algorithm that extracts from the sound wave the stimulus parameters which are defined and required by the theory. Any knowledge available about the perception of pitch – including the definition of the stimulus parameters – must be depicted by the theory, while extraction and measurement of the stimulus parameters as such is a purely physical problem whose solution does not require any knowledge about pitch perception.

An important practical consequence of this concept is that the theory – being self-contained – can be verified independently from the computationally expensive procedure of parameter extraction.

The theory of pitch perception can concisely be outlined by

$$P(t) = \Phi\{S_{\mathrm{L}}(t), S_{\mathrm{R}}(t), L\}, \tag{1}$$

where P denotes the set of perceived pitches; $S_{\mathrm{L}}, S_{\mathrm{R}}$ the sets of stimulus parameters for the left and right ears; and L a set of listener-specific parameters. $\Phi\{..\}$ denotes an algorithm (in terms of computer language: a *procedure* or *function*) that includes specifications about which of the stimulus- and listener parameters are relevant and how they affect the set of pitches. For simplicity, in the following considerations the listener-parameter set L is mostly left out.

The pitch set P in general consists of several or many pitches, each of which in turn is a set of attributes of which the most important ones are h (height), σ (strength/pronouncedness), and λ (lateralization):

$$\begin{aligned} P(t) &= \langle p_1(t), p_2(t), \ldots \rangle \\ &= \langle \langle h_1(t), \sigma_1(t), \lambda_1(t) \rangle, \langle h_2(t), \sigma_2(t), \lambda_2(t) \rangle, \ldots \rangle. \end{aligned} \tag{2}$$

The two stimulus-parameter sets S_{L}, S_{R} may either include a complete description of the respective stimuli or may be confined to those parameters which are sufficient to determine the pitch set. In any case there is an inherent interdependency between the stimulus parameters and the algorithm Φ. The latter constitutes the core of the theory and implicitly defines which kind of stimulus parameters it operates on. The theory is complete only if it includes that definition. There is no apparent criterion according to which the stimulus parameters must be of either a "temporal" or "spectral", or any other kind. In general, the parameters must be presumed to be time variant.

For the evaluation of the implications of modeling pitch perception it is helpful to use a kind of standard description of the stimulus. A description which, on the one hand, enables

specification of *any* type of stimulus and, on the other hand, is adapted to the fact that in pitch perception the sine wave plays a particular role, has the form

$$a(t) = \sum_{n=1}^{N} \hat{a}(t) \cos[2\pi f_n(t)t + \varphi_n]. \quad (3)$$

This is a *synthetic, parametric*, and universal description of the *audio signal a(t)*, i.e., the sound-pressure wave at each of the two ears.[1] As indicated, the sinusoidal components included in (3) are in general supposed to be time variant in frequency and amplitude. The time-variance of components ensures that *any* type of sound can be described in the form (3), i.e., with any number $N \geq 1$ of components.

By (3) the stimulus-parameter set S at each of the two ears is defined, i.e.,

$$S(t) = \langle \langle f_1(t), \hat{a}_1(t), \varphi_1 \rangle, \langle f_2(t), \hat{a}_2(t), \varphi_2 \rangle, \ldots, \langle f_N(t), \hat{a}_N(t), \varphi_N \rangle \rangle. \quad (4)$$

The actual number and values of the parameters may be different for the two ears, i.e., in the monaural and the dichotic conditions.

If only periodic sound waves would elicit pitch, and if there would occur only one pitch per periodic sound wave, specification of the stimulus-parameter set would be much simpler. For any singular periodic stimulus, S would not need to have the form (4) but would include only the type specification *periodic* and the length of the period or its reciprocal, i.e., oscillation frequency. The algorithm Φ would boil down to a mere formal frequency-to-pitch assignment. However, it is neither true that only periodic sound waves elicit pitch; nor is it true, for non-sinusoidal periodic sound waves, that there occurs only one pitch per periodic sound wave [25]. Therefore, any theory and/or model of pitch that rigorously depends on the stimulus' periodicity is inadequate.

The only type of periodic sound wave to which the above conditions apply – at least to a sufficient extent – is the singular sine wave. For singular sine waves the pitch theory is just as simple as was just hypothetically sketched for periodic signals in general: to a first approximation, there is a "one-to-one" relationship between a sine tone's oscillation frequency and pitch.

This provides for reliable and unambiguous measurement of the various pitches which in general are elicited by a complex stimulus: the test sound is presented to a listener in alternation

[1] This mathematical description must not be confused with a Fourier spectrum.

with a sine tone (the latter preferably presented diotically and with a definite SPL), and the sine tone's frequency is adjusted such that its pitch height matches the height of the test sound's particular pitch envisaged. The numerical value of any pitch can most conveniently be obtained by defining it to be identical to that of the match tone's frequency. Pitch height quantitatively expressed in this way is measured in *pitch units* (pu) [24–26]. The *strength* of any pitch can be determined by subjective assessment [9,29]. If required, also *lateralization* may be determined and included in the pitch set. In this way, the pitch set (2) of any complex sound can be obtained by fairly simple psychoacoustic experiments.

It has been proven helpful to make a distinction between *spectral pitch* and *virtual pitch* [22,25]. The former type of pitch is characterized by a "one-to-one" relationship to spectral singularities of any kind. Isolated sine tones are prominent examples of such singularity. However, spectral singularities may also occur in random and quasi-random types of stimuli. Any pitch which is not in a "one-to-one" relationship to such a singularity is termed virtual pitch. In general, virtual pitch is conceptualized as an entity that resides on the second or even a higher level of the perceptual hierarchy, while spectral pitch occurs on the first [22,25].

3 • Pitch of Tonal Stimuli

The tonal type of sound is characterized by the features that (a) the component frequencies included in (4) are to a large extent in a systematic relationship to one another (an objective criterion); and (b) a number of component frequencies elicit individual spectral pitches (an auditory criterion).

3.1 Spectral Pitch

Using (4), the formation of the two sets of spectral pitches $P_{s,L,R}$ elicited at the two ears can be outlined by

$$P_{s,L,R}(t) = \Phi_{s,L,R}\{S_{L,R}(t)\}, \tag{5}$$

where $S_{L,R}$ corresponds to (4). Equation (5) actually depicts two independent equations, i.e., one for each ear. There is compelling evidence that spectral pitch is constituted by each ear alone [2, 25], though its height may be slightly affected by a contralateral stimulus [23].

Each of the two algorithms $\Phi_{s,L,R}$ must yield a set of spectral pitches each of which is characterized by a numerical value for its height, another for its strength, and an appropriate characterization of lateralization (which is left out of consideration

here). To a first approximation, and by definition, the numerical value of pitch height can be set equal to that of the pertinent component's frequency, and the unit of height is pu [22,25]. As spectral-pitch height in general may also be affected by (a) SPL; (b) mutual interaction of adjacent components; (c) phase; and (d) interaural interaction, $\Phi_{s,L,R}$ should account for these effects. A preliminary algorithmic solution, which may serve as an example, was described in the virtual-pitch theory [24,26].

Pitch *strength* can be supposed to depend on (a) the pertinent component's amplitude (both absolute and relative to the threshold of hearing); (b) the degree to which the pertinent component can be assumed to stay unmasked by other adjacent components; (c) phase; (d) frequency-dependent prominence. (Components in the region around 700 Hz are most prominent, i.e., *dominant*, merely *because* they lie in that region [26].) As a result, each of the two spectral-pitch sets $P_{s,L,R}$ will include a number of pitches which is limited by the criterion that their strength must be $\sigma > 0$.

3.2 Virtual Pitch

One of the most striking features of virtural pitch is that it is not dependent on the presence of a sinusoidal component whose frequency corresponds to virtual-pitch height. However, the information about any virtual pitch surely is contained in the stimulus-parameter sets $S_{L,R}$.

Basically, several types of algorithm for the formation of virtual pitch are conceivable. In particular, it is tempting to suspect that virtual pitch may be obtained by an analysis of the temporal structure of the stimulus, i.e., essentially the structure created by those components which are not aurally "resolved" to such an extent that they can elicit individual spectral pitches [12,18]. However, already in the 1960s it was found that in fact virtual pitch essentially is dependent on components which evidently *are* resolved [22]. Moreover, it was found that the precise height of virtual pitches is in several ways correlated with the precise height of the spectral pitches which are elicited by the tonal stimulus [3,12,22,25,27]. These findings indicate that virtual pitch is not constituted by a mechanism which is independent of spectral pitch – irrespective of whether in the frequency- or time-domain – but is dependent on, and emerges from, the set of spectral pitches.

This conclusion is strongly supported by the observation that virtual pitch can even be elicited by a periodic stimulus whose set of harmonics is split, one subset being presented to the left ear, the other to the right [15]. While spectral pitch is monaural (Sect. 3.1), virtual pitch is binaural. Another compelling support comes from the observation that virtual pitches can be elicited

by a harmonic complex tone whose harmonics are presented in a rapid temporal succession instead of simultaneously [11]. As this stimulus is a sequence of non-simultaneous sinewaves, those virtual pitches are not at all encoded in the time structure of the stimulus. As the auditory system evidently is equipped with a short-term memory for pitch, formation of virtual pitch from a rapid succession of spectral pitches is well conceivable.

Taking into account that any kind of perception is hierarchical, the above findings support the conclusion that virtual pitch is a perceptual entity that resides on a higher level of the hierarchy of pitches than spectral pitch and that virtual pitch is by an appropriate algorithm derived from the set of spectral pitches [22].

A number of authors have pointed out that a faint virtual pitch may be elicited by a monaural or diotic harmonic complex tone which is high-pass filtered such that it includes only quite high-order harmonics (e.g., above the 20th) [5,16,18]. Though it may appear questionable if such high harmonics can elicit individual "sine tone pitches", one cannot be sure that such a high-pass filtered complex tone does not elicit any spectral pitch at all. It is even likely that there is a spectral pitch corresponding to the cut-off frequency – and *one* spectral pitch can be sufficient to elicit faint subharmonic virtual pitches [14]. Moreover, examples such as *overtone singing*, and melodies played on the *Jew's harp*, demonstrate that both spectral pitch and virtual pitch can be elicited by very-high-order harmonics of a periodic stimulus with a very low oscillation frequency. Whatsoever, the very faintness and singularity of the virtual pitch of "unresolved harmonics" render its potential theoretical implications overruled by the wealth of compelling evidence just outlined.

Hence, an adequate outline of the formation of virtual pitch can be denoted by the formula

$$P_{\mathrm{v}}(t) = \Phi_{\mathrm{v}}\{P_{\mathrm{s,L}}(t), P_{\mathrm{s,R}}(t)\}, \tag{6}$$

where $P_{\mathrm{s,L}}$ and $P_{\mathrm{s,R}}$ are obtained from (5).

A guideline for the development of the algorithm Φ_{v} is provided by the eminent *robustness* of the pitch elicited by complex stimuli. In real life, the auditory stimulus routinely suffers drastic linear distortion, such that even for systematic and steady types of sound the amplitude- and phase-distribution of components at the ear is variable and in several respects unpredictable. It is the spectral frequencies – and, therefore, the spectral pitches – which provide reliable clues to the characteristics of the sounds emitted by the sound sources. The evident perceptual constancy of the pitch percept suggests that, on a second or higher level of the auditory hierarchy, a *completion mechanism* must be involved, and it makes sense to assume that virtual pitches are the result of

such kind of pitch-pattern completion. The completion algorithm appears to operate on the presumption that in real life any set of spectral pitches emerges from the presence of one or several *harmonic* complex tones. In fact, harmonicity of components is a strong indicator for those components to be pertinent to one and the same periodically oscillating source. The principle of *subharmonic coincidence detection* applied in the virtual-pitch theory [24,26] may serve as an example for the way in which the above notions may be cast into the algorithm Φ_v.

3.3 The Tonal Percept

The ultimate pitch set P, which consists of both spectral and virtual pitches, can be assumed to emerge by evaluation and interpretation of the monaural pitch sets $P_{s,L}, P_{s,R}$, and the binaural set P_v, i.e., on a higher level of the perceptual hierarchy. For the description of the ultimate pitch set another algorithm Φ_p is required. This algorithm may also account for aspects such as formation of multiple sensory objects and contextual effects [12,18,25]. When listener-specific influences are also envisaged, one eventually obtains the formula

$$P(t) = \Phi_p\{P_{s,L}(t), P_{s,R}(t), P_v(t), L\}, \tag{7}$$

where $P_{s,L}$ and $P_{s,R}$ are obtained from (5), and P_v from (6).

3.4 Conclusion

The outline (7) of the pitch theory for tonal sounds is meant to account for those basic aspects and conditions which can be regarded as evident, while there remains plenty of space for specification of details and variants. Indeed, many details – in particular, where high-level pitch evaluation is concerned – remain to be specified. Several basic algorithms implied in (7) have been worked out and verified in the virtual-pitch theory [24–26]. Those algorithms may serve as examples.[2]

As the theory as such does not require processing of the audio signal, it can be tested and verified with moderate computational burden.

Extension of the theory into a *model* requires inclusion of an algorithm that extracts the component parameters (4) from the audio signal. A considerable number of such algorithms are available [4,10,13,17,21]. The model thus obtained provides a basis for applications such as pitch extraction from speech; automatic transcription of music; root-finding of musical chords.

[2]Software for testing the virtual-pitch theory may be downloaded from ftp://ftp.mmk.e-technik.tu-muenchen.de/pub/akom/ter/.

4 • Pitch of Random Stimuli

4.1 Pitch of Band-Limited Random Noise

When presented monaurally or diotically, narrow-band random noise typically elicits a pitch that essentially corresponds to its center frequency. When the bandwidth is increased and the spectral slopes are steep, the pitch splits into two pitches which approach the edges of the spectral contour [20,29]. It is conceivable that this behavior can be accounted for by a generalized form of the spectral-pitch algorithm Φ_s.

4.2 Pitch of Wideband Random Noise with Spectral Peaks

Equidistant peaks and gaps do occur in the power spectrum of wideband random noise when a truly random noise after a brief delay is superimposed to itself. For both monaural and diotic presentation, one or several pitches may be heard that are related to the reciprocal of the time delay [1,28,29]. It is conceivable that the phenomenon can be explained on the basis of a theory of the spectral pitch of narrow-band random noise (Section 4.1), and probably virtual pitches become involved in a way equivalent to that of tonal sounds (Section 3.2) [1].

4.3 Dichotic Pitch Phenomena

A number of dichotic pitch phenomena have been described which typically are elicited by dichotic presentation of two wideband random-noise stimuli, one of which was somehow manipulated [1,7,8]. Ordinarily, no pitch is heard when either of the stimuli is presented alone, i.e., monaurally. For example, the phase of the Fourier spectrum of a wideband random noise was within a narrow frequency interval continually shifted by 360°. When the noise thus manipulated is presented to one ear and the original noise to the other, a faint pitch can be heard that resembles that of a narrow-band noise situated in the frequency interval of phase transition [6,7]. As this type of pitch obviously is dependent on a kind of spectral singularity, the phenomenon demonstrates an exception from the rule that spectral pitch is basically monaural (cf. Section 3.1). To explain it, the algorithm for spectral-pitch formation needs to be generalized, and an account of binaural interaction must be included [1,7].

4.4 Pitch of Amplitude-Modulated Wideband Noise

When wideband random noise gets "pulsed", i.e., periodically switched off and on, a faint pitch corresponding to the modulation frequency may be heard for both monaural and diotic presentations [19,29]. The very faintness of the pitch excludes the possibility that periodicity of the temporal envelope as such can be the determining clue. The power spectrum of the modulated noise fails to reveal any clue. Hence, the faint pitch may be determined by the phase spectrum. As yet there does not appear to exist a theoretical description of that relationship [12].

4.5 Conclusion

For several of the pitches of random stimuli it is not readily apparent if and how they can be accounted for by a theory akin to the one for tonal sounds. Yet it is conceivable that by appropriate generalization of the algorithm for formation of spectral pitch a unified theory and model can be achieved.

References

1. F.A. Bilsen: J. Acoust. Soc. Am. **61**, 150 (1977)
2. G. van den Brink: Acustica **32**, 159 (1975)
3. G. van den Brink: Acustica **32**, 167 (1975)
4. J.C. Brown, H.S. Puckette: J. Acoust. Soc. Am. **94**, 662 (1993)
5. R.P. Carlyon, T.M. Shackleton: J. Acoust. Soc. Am. **95**, 3541 (1994)
6. E.M. Cramer, W.H. Huggins: J. Acoust. Soc. Am. **30**, 413 (1958)
7. J.F. Culling, A.Q. Summerfield, D.H. Marshall: J. Acoust. Soc. Am. **103**, 3509 (1998)
8. J.F. Culling, D.H. Marshall, A.Q. Summerfield: J. Acoust. Soc. Am. **103**, 3527 (1998)
9. H. Fastl, G. Stoll: Hear. Res. **1**, 293 (1979)
10. T. Funada: Signal Process. **13**, 15 (1987)
11. J.W. Hall, R.W. Peters: J. Acoust. Soc. Am. **69**, 509 (1981)
12. W.M. Hartmann: J. Acoust. Soc. Am. **100**, 3491 (1996)
13. W. Heinbach: Acustica **67**, 113 (1988)
14. T. Houtgast: J. Acoust. Soc. Am. **60**, 405 (1976)
15. A.J.M. Houtsma, J.L. Goldstein: J. Acoust. Soc. Am. **51**, 520 (1972)
16. A.J.M. Houtsma, J. Smurzynski: J. Acoust. Soc. Am. **87**, 304 (1990)
17. R.J. McAulay, T.F. Quatieri: IEEE Trans. ASSP **34**, 744 (1986)
18. B.C.J. Moore: *An Introduction to the Psychology of Hearing* (Academic, New York 1989)
19. I. Pollack: J. Acoust. Soc. Am. **45**, 237 (1969)
20. A. Rakowski: Pitch of filtered noise. In: *Proc. 6 Int. Congr. Acoust.* (Tokio 1968) pp. 105–108
21. S. Seneff: IEEE Trans. ASSP **26**, 358 (1978)
22. E. Terhardt: J. Acoust. Soc. Am. **55**, 1061 (1974)
23. E. Terhardt: Acustica **37**, 56 (1977)
24. E. Terhardt: Hear. Res. **1**, 155 (1979)
25. E. Terhardt: *Akustische Kommunikation* (Springer, Berlin Heidelberg New York 1998)
26. E. Terhardt, G. Stoll, M. Seewann: J. Acoust. Soc. Am. **71**, 679 (1982)
27. K. Walliser: Acustica **21**, 319 (1969)
28. W.A. Yost: J. Acoust. Soc. Am. **100**, 511 (1996)
29. E. Zwicker, H. Fastl: *Psychoacoustics*, 2nd edn (Springer, Berlin Heidelberg New York 1999)

18

Calculation of Loudness for Normal and Hearing-Impaired Listeners

Josef Chalupper

Siemens Audiological Engineering Group, Gebbertstrasse 125, Erlangen 91058, Germany, e-mail: Josef.Chalupper@siemens.com

1 • Introduction

Loudness is a hearing sensation which is almost ubiquitous in everyday life: traffic is *too loud*, announcements are *too soft*, children are *very loud* and a stereo can be adjusted to a *comfortable loudness*. Thus, it is not surprising that scientists started very early to carry out psychoacoustic experiments to investigate loudness. In 1933, Fletcher and Munson [1] measured "equal-loudness contours". These curves are obtained if a reference sound (e.g. a 1,000-Hz sinusoid) is fixed in level and the test sound is adjusted to give a loudness match. Some typical results are shown in Figure 1 [2]. Figure 1 shows equal-loudness contours for loudness levels from 3 to 100 phon.[1]

[1] "phon" is the unit of the loudness level and is defined as the level (dB SPL) of a 1-kHz sinusoid with the same loudness as the test sound.

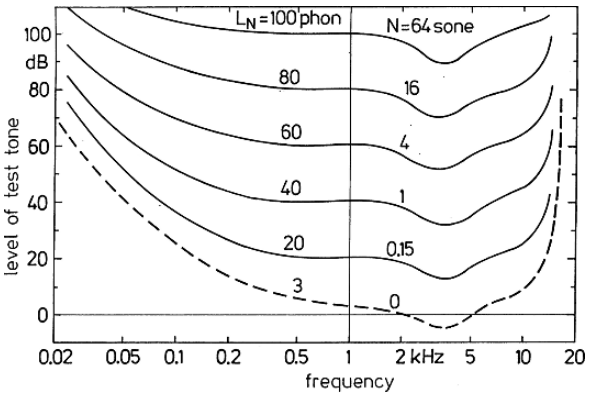


FIGURE 1 Equal-loudness contours for pure tones in a free sound field. Taken from [2].

Obviously, loudness of pure tones depends on frequency. Moreover, this dependency varies with level, as the dip between 2 and 5 kHz becomes deeper with increasing levels. The remarkable differences between the slopes of the equal-loudness contours below 1 kHz resulted in the definition of level-dependent weighting curves ("dB A", "dB C") which are often applied in technical acoustics [3]. In many other listening experiments (for comprehensive overviews see [2,4]), it was shown that loudness is affected by a large variety of signal parameters: frequency, bandwidth, duration, level and temporal structure. In addition, early experiments by Steinberg and Gardner [5] indicated that loudness perception also strongly depends on properties of the human listener. In particular, the rate of growth of loudness with increasing level is greater for listeners with cochlear hearing impairment than for normal listeners. This phenomenon is called "recruitment" (Figures 2 and 3).

In Figure 2, median loudness scaling data (circles) for 57 normal hearing subjects using the WHF method ("Würzburger Hörfeld", see [6]) are shown [7]. With this method, the loudness of 1/3-octave-band noises is measured by a combination of verbal categories and an underlying scale of "Categorical Units" (CU) ranging from 0 ("inaudible") to 50 ("too loud"). In general, the curves in Figure 2 are very similar across frequencies: starting with 0 CU for a sound level of 10 dB; values close to 50 CU are obtained for levels around 90 dB. Interestingly, the

---

"sone" is the unit of loudness. A 1-kHz sinusoid with a level of 40 dB SPL by definition elicits a loudness of 1 sone.

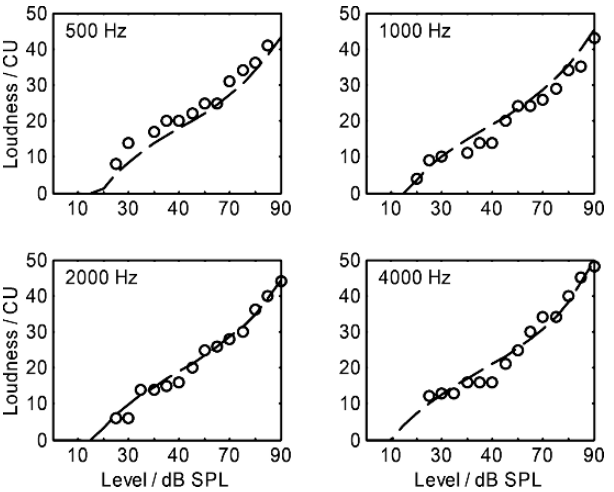


FIGURE 2 Loudness scaling by normal hearing subjects. Data measured by the WHF method (*circles*) [6] and predictions by the DLM (*broken curves*). From [7].

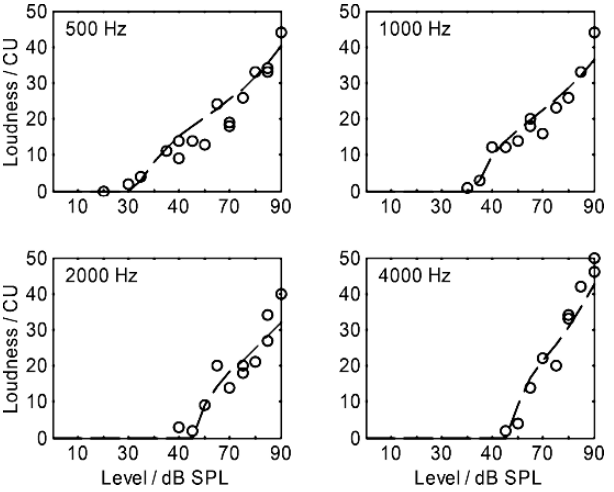


FIGURE 3 Same as Figure 2, but for a hearing-impaired subject with presbyacusis. From [7].

curves of the hearing-impaired subject (Figure 3) have nearly the same loudness at high levels, whereas the starting point of the loudness functions is shifted to higher levels due to the individual hearing loss. Thus, the loudness functions at 2 and 4 kHz, where hearing loss is about 50 dB HL, are nearly twice as steep as the

respective curves of the normal hearing listeners in Figure 2. If loudness functions are measured by assigning numbers to sounds according to their relative or absolute loudness ("magnitude estimation"), typically 10 dB increase in level is required to increase loudness by a factor of 2 [2].

As a consequence of this large number of complex dependencies of loudness on stimulus parameters, methodology and properties of the human hearing system, various computational models of loudness were developed to calculate the loudness of a given sound [1,8,9]. The loudness model of Zwicker [8,10] was originally designed for calculating the loudness of steady-state sounds for normal hearing listeners. In brief, first the excitation pattern is calculated from a spectral analysis based on properties of the human hearing system (as revealed mainly in masking), and then the excitation pattern is transformed into the "specific loudness pattern". The area under the specific loudness pattern indicates the overall loudness. Later on [11], this model was extended to account also for the loudness of time-varying sounds and was used as the basis for the calculation of other hearing sensations like roughness and sharpness [2]. In 1991, Zwicker's procedure for calculating of loudness became a normative standard [12]. New psychoacoustic data for loudness of nonstationary stimuli allowed for further refinement of the original dynamic model [2,13,14]. At present, it is widely used in many practical applications, in particular for sound quality engineering tasks [2]. In 2002, Zwicker's model was revised and extended to hearing-impaired listeners by Chalupper and Fastl [7]. The structure of this model (DLM – dynamic loudness model), specific aspects and some predictions for important test cases will be presented in more detail in the following section.

2 • Model

Figure 4 shows the basic structure of the dynamic loudness model.

The first stage is a filter, which takes into account the transmission through outer and middle ear. The frequency response of this filter depends on the acoustical conditions for the input signal. Currently, freefield conditions are assumed, but also diffuse field sounds can be used [2]. Moreover, if the loudness of in situ or KEMAR recordings shall be assessed, appropriate equalizing can be accomplished in this stage.

The next block is an aurally adequate spectro-temporal analysis consisting of a critical band filterbank. This filterbank was designed according to the Fourier-t transformation (FTT) [15].

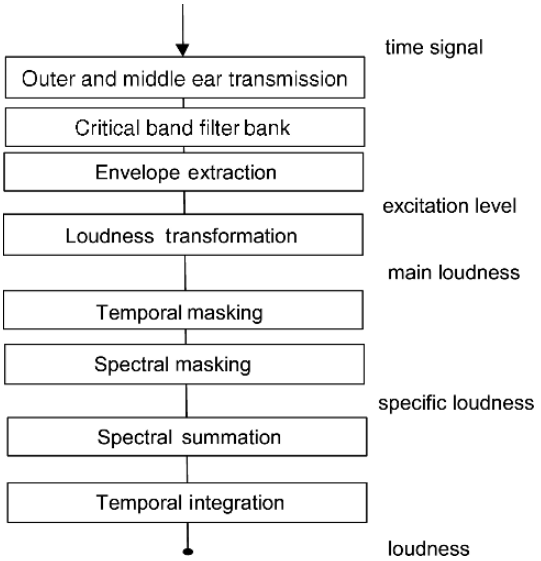


FIGURE 4 Dynamic loudness model (DLM) for normal hearing and hearing-impaired listeners.

Analysis frequencies are the centre frequencies of critical bands as described by Zwicker [16]. Essentially, the FTT analysis is identical to a Gammatone filterbank [17] with an equivalent rectangular bandwidth of 1 Bark [18].

The following block "envelope extraction" contains a window to extract the temporal envelopes of the outputs from the band-pass filters according to a proposal of Plack and Moore [19]. At the output of this block, the excitation level $L_{\mathrm{E}}(z,t)$ is available, which is related to excitation E by

$$L_E = 10 \cdot \log\left(\frac{E}{E_0}\right) \mathrm{dB}, \tag{1}$$

with E_0 the excitation corresponding to the reference value of sound intensity $I_0 = 10^{-12}\mathrm{W/m}^2$.

In contrast to other loudness models for hearing-impaired listeners [20–22], it is not required to adapt the spectro-temporal analysis to the reduced spectral and temporal resolution typically seen in cochlear hearing impairment [23,24]. This is possible because in the DLM, the effects of spectro-temporal resolution are splitted in a linear and a nonlinear part. Whereas the linear part ("critical band filterbank" and "envelope extraction") takes into account the maximum possible resolution of the human hearing system, the nonlinear part ("temporal masking" and

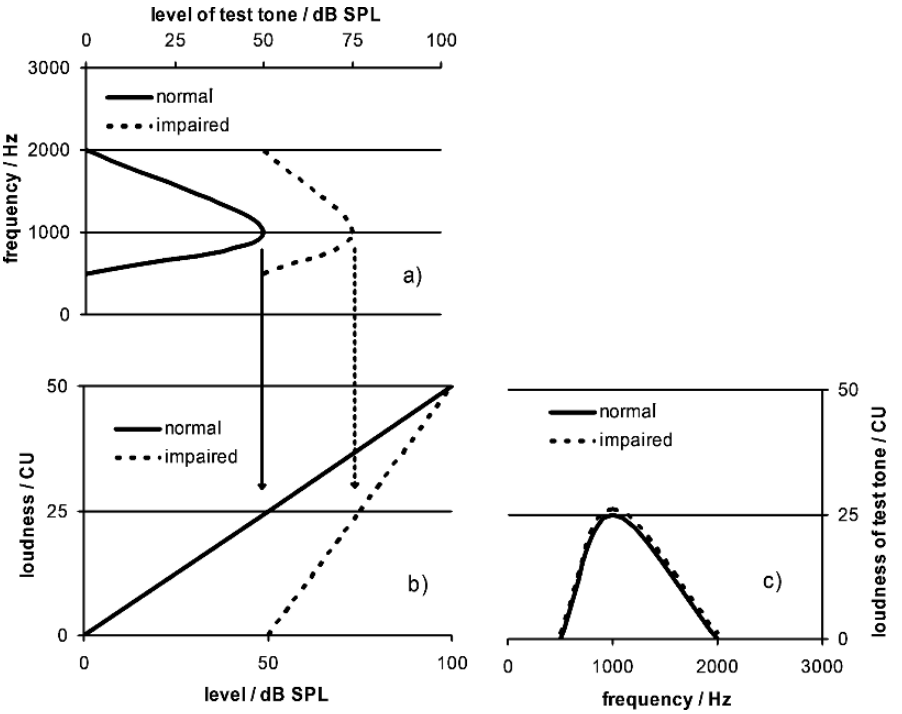


FIGURE 5 Slopes of masking patterns are shallower in hearing-impaired listeners (**a**). If the masking patterns are transformed to loudness by the individual loudness functions (**b**), the steepness of the slopes is the same for normal and hearing-impaired listeners (schematic sketch).

"spectral masking") is primarily designed to model dependencies on level and duration. Due to the position of these blocks after the loudness transformation, they also account for the effects of hearing loss on masking patterns, and thus on spectro-temporal resolution. This interesting behaviour can be explained best with the schematic sketch in Figure 5.

Figure 5a shows typical spectral masking patterns obtained with a narrow band noise. The levels of the masking noise have been adjusted to the same loudness for the normal and hearing-impaired listener. The slopes of the masking patterns of the normal listener are steeper by factor of 2 than the respective slopes of the hearing-impaired listener. If each data point of these masking patterns is transformed from level (in dB SPL) to loudness (in CU) by the respective loudness functions (Figure 5b), the slopes of the resulting patterns in the loudness domain have the same steepness for normal and hearing-impaired listeners. The same considerations also can be applied

to temporal masking patterns by simply changing the labelling of the x-axes in Figure 5a and c from "frequency" to "time". This means that reduced spectral and temporal resolution of hearing-impaired listeners is not required to be adapted explicitly if masking effects are modelled in terms of loudness. Since masking is taken into account after the loudness transformation in the DLM, the only adjustment required to adopt the model to an individual hearing-impaired listener is to modify the steepness of the loudness transformation.

The loudness transformation as proposed by Zwicker [8] transforms excitation into specific loudness:

$$N'(z) = N_0 \left(\frac{E_{\mathrm{THQN}}(z)}{s(z)E_0} \right)^{\alpha} \left[\left(1 - s(z) + s(z) \frac{E(z)}{E_{\mathrm{THQN}}(z)} \right)^{\alpha} - 1 \right] \quad \text{sone / Bark}, \tag{2}$$

where $E_{\mathrm{THQN}}(z)$ denotes the excitation at threshold in quiet, in case of normal hearing, and N_0 is a constant, which is chosen to achieve a loudness of 1 sone for a 1 kHz sinusoid with a level of 40 dB. The threshold factor s is the ratio between the intensity of the just-audible test tone and the intensity of the internal noise appearing within the critical band at the test tone's frequency. The exponent $\alpha = 0.23$ indicates that specific loudness is related to the fourth root of intensity. For details, the reader is referred to [10].

In order to account for individual differences in the loudness transformation for people with hearing impairment, a so-called two component approach [21,25] is used. In essence, the total hearing loss HTL is divided into a component which accounts for the attenuation $\mathrm{HTL_a}$ (seen in elevated thresholds) and another component realizing an expansion $\mathrm{HTL_e}$ (seen in the recruitment phenomenon). The following equations apply:

$$\mathrm{HTL} = \mathrm{HTL_a} + \mathrm{HTL_e}, \tag{3}$$

$$\mathrm{HTL_e} = k \times \mathrm{HTL} \quad (0 < k \leqq 1), \tag{4}$$

$$\mathrm{HTL_a} = (1 - k)\mathrm{HTL}. \tag{5}$$

In order to incorporate both components into the model, $E(z)$ in (2) is replaced by

$$E(z) = E_0 \cdot 10^{(L_{\mathrm{E}}(z) - \mathrm{HTL_a}/10)} \tag{6}$$

and $E_{\mathrm{THQ}}(z)$ by

$$E_{\mathrm{THQ}}(z) = E_{\mathrm{THQN}} \cdot 10^{(\mathrm{HTL_e}/10)}. \tag{7}$$

The factor k in (4) and (5) is determined by minimizing the difference between measured loudness scaling data and calculated loudness. This means that besides the hearing loss, loudness scaling data are required to adapt the loudness model to an individual hearing-impaired listener. If no loudness scaling data are available, these are either estimated from the uncomfortable level (UCL) or k is set to 0.8, which has been found empirically to be a good estimate [22]. In order to transform loudness N in sone into category units CU, the following formula was derived from empirical data [26]:

$$\mathrm{CU} = 0.6\ N/\mathrm{sone} + 11.8\ \log(16.5\ N/\mathrm{sone}). \tag{8}$$

Figure 6 shows the resulting specific loudness functions for $k = 0$ and $k = 1$ for a hearing loss of 50 dB in comparison to a normal hearing listener. If k is optimized individually, results of loudness scaling of narrow band noises are predicted with a high accuracy for both normal and hearing-impaired listeners (see Figure 2).

After the loudness transformation, the nonlinear aspects of temporal and spectral masking effects are modelled. At the output of these blocks, the specific loudness–time pattern is available with a resolution of 0.1 Bark. By integration across

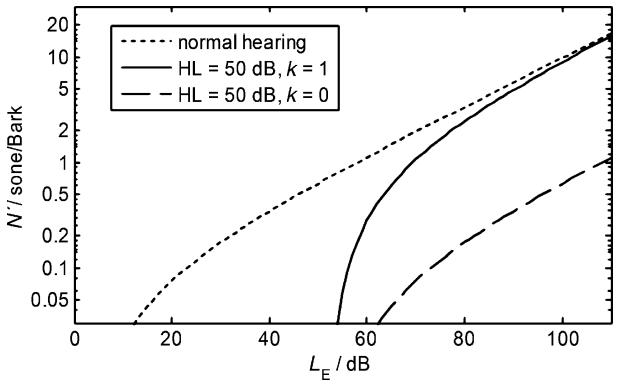


FIGURE 6 Specific loudness as a function of excitation level at 1 kHz. *Dotted*: normal hearing subject, *solid*: hearing-impaired subject with 50 dB hearing loss and $k = 1$. *Dashed*: hearing-impaired subject with 50 dB hearing loss and $k = 0$.

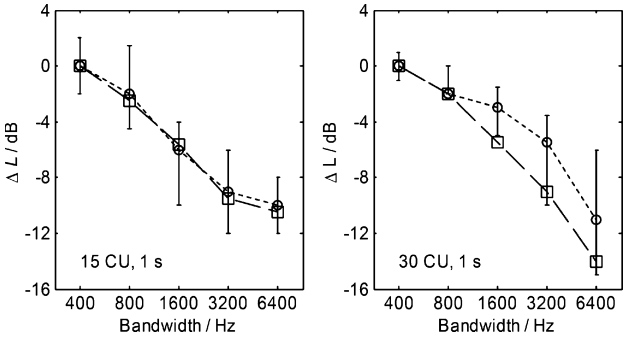


FIGURE 7 Level difference at same loudness between noises with various bandwidths and a narrow band noise of 400 Hz at 2 kHz. *Circles*: median and quartiles (*vertical lines*) of data from eight normal hearing subjects. *Squares*: prediction by the DLM. *Left panel*: loudness of the 400 Hz wide noise is 15 CU (45 dB). *Right panel*: 30 CU (70 dB). From [7].

frequency and time, the dependencies of loudness on bandwidth and duration are taken into account. If the level of narrow band noise is adjusted to match the loudness of a broad band noise, the resulting level is higher than the level of the broad band noise. This effect is called "loudness summation" and amounts to about 10–15 dB (depending on bandwidth and level) for normal hearing listeners (Figure 7). In subjects with sensorineural hearing loss, loudness summation is reduced (Figure 8). Reduced spectral resolution and recruitment in literature are often considered to be the reasons for reduced loudness summation [20,24]. As in Figures 7 and 8, the DLM is capable of predicting the amount of loudness summation both for normal and hearing-impaired listeners – without explicitly modifying spectral resolution (remember Figure 5).

In the last block of the DLM, the temporal integration of loudness is addressed. Experimental data have shown that loudness decreases for signal durations smaller than 100–200 ms. If the loudness of a 10 ms tone is matched to the loudness of a 100 ms tone, typically an increase in level of 10 dB is required for normal hearing listeners [2]. Similar to loudness summation, also the amount for temporal integration of loudness is reduced in hearing-impaired listeners [27]. Following the line of reasoning for reduced spectral resolution (Figure 5), in the DLM, the same low pass is sufficient to successfully model temporal integration for normal as well as hearing-impaired listeners [7].

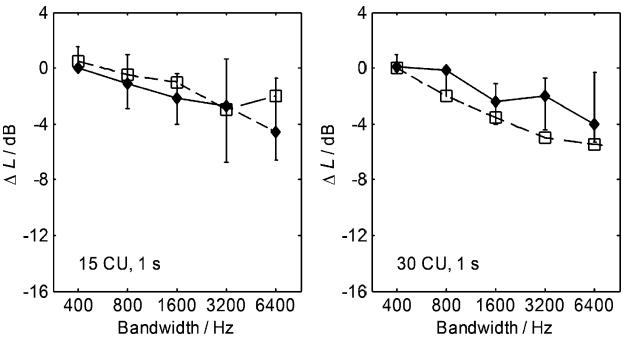


FIGURE 8 Same as Figure 7, but for subjects with cochlear impairment. *Filled symbols*: data from seven subjects with cochlear impairment (medians and quartiles). *Open squares*: predictions by the DLM. From [7].

3 • Conclusions

Loudness is a hearing sensation which depends on many properties of the assessed signal (stimulus) and the human hearing system (receiver). Since loudness has been of interest for science as well as industry (e.g. hearing aid manufacturers, sound quality labs), the most important effects are well understood today and computational models are available for calculating loudness of arbitrary sounds – not only for normal listeners but also for listeners with hearing impairment.

References

1. Fletcher, H., Munson, W.A., "Loudness, its definition, measurement and calculation", J. Acoust. Soc. Am., 5, 82–108, 1933.
2. Zwicker, E., Fastl, H., "Psychoacoustics: Facts and Models", 2nd ed., Berlin, Heidelberg, New York: Springer, 1999.
3. Beranek, L., "Acoustical measurements", J. Acoust. Soc. Am., 1988.
4. Moore, B.C.J., "An Introduction to the Psychology of Hearing", 4th ed., London: Academic Press, 1997.
5. Steinberg, J., Gardner, M., "Dependence of hearing impairment on sound intensity", J. Acoust. Soc. Am., 9, 11–23, 1937.
6. Heller, O., "Orientated category scaling of loudness and speech-audiometric validation", In: Contributions to Psychological Acoustics, Vol. 5, A. Schick, Ed., Bibliotheks- und Informationssystem der Univ. Oldenburg, Oldenburg, 135–159, 1991.
7. Chalupper, J., Fastl, H., "Dynamic loudness model (DLM) for normal and hearing-impaired listeners", Acustica – Acta Acustica, 88, 378–386, 2002.

8. Zwicker, E., "Über Psychologische und Methodische Grundlagen der Lautheit", Acustica, 8, 237–258, 1958.
9. Zwicker, E., Scharf, B., "A model of loudness summation", Psychol. Rev., 72, 3–26, 1965.
10. Zwicker, E., "Ein Verfahren zur Berechnung der Lautstärke", Acustica, 10, 304–308, 1960.
11. Zwicker, E., "Procedure for calculating loudness of temporally variable sounds", J. Acoust. Soc. Am., 62, 675–682, 1977.
12. DIN45631, "Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum, Verfahren nach E. Zwicker", DIN 45631, 1991.
13. Zwicker, E., "Dependence of post-masking on masker duration and its relation to temporal effects in loudness", J. Acoust. Soc. Am., 75, 219–223, 1984.
14. Widmann, U., Lippold, R., Fastl, H., "Ein Computerprogramm zur Simulation der Nachverdeckung für Anwendungen in Akustischen Meßsystemen", In: Fortschritte der Akustik, DAGA'98, Dt. Gesell. für Akustik e. V., Oldenburg, 96–97, 1998.
15. Terhardt, E., "Fourier transformation of time signals: conceptual revision", Acustica, 57, 242–256, 1985.
16. Zwicker, E., "Subdivision of the audible frequency range into critical bands (Frequenzgruppen)", J. Acoust. Soc. Am., 33, 248, 1961.
17. Patterson, R.D., Holdsworth, J., Nimmo-Smith, I., Rice, P., "SVOS final report: the auditory filterbank", MRC Applied Psychology Unit, Cambridge, APU report 2341, 1986.
18. Chalupper, J., Fastl, H., "Simulation of hearing impairment based on the Fourier Time Transformation", Proceedings ICASSP 2000, Istanbul, Turkey, 857–860, 2000.
19. Plack, C.J., Moore, B.C.J., "Temporal window shape as a function of frequency and level", J. Acoust. Soc. Am., 87, 2178–2187, 1990.
20. Florentine, M., Zwicker, E., "A model of loudness summation applied to noise-induced hearing loss", Hear. Res., 1, 121–132, 1979.
21. Moore, B.C.J., Glasberg, B.R., "A model of loudness perception applied to cochlear hearing loss." Aud. Neurosci., 3, 289–311, 1997.
22. Moore, B.C.J., Glasberg, B.R., Vickers, D.A., "Further evaluation of a model of loudness perception applied to cochlear hearing." J. Acoust. Soc. Am., 106, 898–907, 1997.
23. Tyler, R., "Frequency resolution in hearing-impaired listeners", In: Frequency Selectivity in Hearing, B.C.J. Moore, Ed., Academic Press, London, 309–371, 1986.
24. Moore, B.C.J., "Perceptual Consequences of Cochlear Damage", London: Oxford University Press, 1995.
25. Launer, S., Hohmann, V., Kollmeier, B., "Modeling loudness growth and loudness summation in hearing-impaired listeners" In: Modeling Sensorineural Hearing Loss, W. Jesteadt, Lawrence Erlbaum Mahwah, NJ: 175–185, 1997.

26. Baumann, U., Stemplinger, I., Arnold, B., Schorn, K., "Kategoriale Lautstärkeskalie-rung in der klinischen Anwendung", In: Fortschritte der Akustik, DAGA'96, Dt. Ge-sell. für Akustik e. V., Oldenburg, 128–129, 1996.
27. Buus, S., Florentine, M., Poulsen, T., "Temporal integration of loudness in listeners with hearing losses of primarily cochlear origin", J. Acoust. Soc. Am., 105, 3464–3480, 1999.

19

Psychoacoustical Roughness

Peter Daniel

Brüel & Kjær, GmbH, Bremen, Germany

1 • Introduction

Roughness is a fundamental hearing sensation for fast amplitude modulations. Fast means in this sense that modulation is not any longer perceived as loudness fluctuation but instead as a new quality of the sound: the sound becomes rough. The transition between the sensation of loudness fluctuation and roughness occurs at about 12 Hz for sinusoidal amplitude-modulated sounds.[1]

The sensation of roughness is similar to the sensation of dissonance caused by the beating of tones. It seems to be produced at early stages of the hearing process. If the two beating tones are presented separately to both ears, roughness disappears [9]. As basic listening sensation, roughness was extensively investigated by numerous researchers back from the nineteenth century. Detailed description can be found, e.g., already in Hermann von Helmholtz fundamental book *Die Lehre von den Tonempfindungen* dating from 1863 [15].

[1] The perception of loudness fluctuation – called fluctuation strength – is maximal at a modulation frequency of about 4 Hz, correlating well with the average rate of syllables in speech.

In general, roughness correlates highly with the perceived unpleasantness and annoyance of sounds [8,4,16]. Thus, it is an important parameter for the assessment of the perceived quality of sounds, mainly of technical sounds. Deviations in gears cause, e.g., increased friction with dominating rough components.

2 • Roughness in Dependence of Signal Parameters

The dependency of roughness on signal parameters is very complex and was thoroughly investigated in jury tests [2,6,7, 10–14]. A comprehensive survey and summary of these studies is given in [4,5,17].

2.1 Quantitative Definition of Psychoacoustical Roughness

In order to quantify roughness perception, *the roughness produced by a 1-kHz tone sinusoidal modulated with a modulation frequency of 70 Hz and a degree of modulation of 1 is defined to be one asper.*

By this definition pychoacoustical roughness is linked to the character of this reference sound. In the following the term roughness is used in this sense!

Its dependencies on the different signal parameters can be summarized as follows.

Threshold of roughness. The lower limit of roughness perception equals 0.07 asper for a 1-kHz tone modulated with 70 Hz.

Just noticeable difference. In order to perceive roughness differences, the degree of modulation of a sinusoidal amplitude-modulated sound has to be changed by about 10 % resulting in a just noticeable roughness difference of about 17 %.

Degree of modulation. Main factor influencing roughness perception is the degree of modulation m. For sinusoidal amplitude-modulated sounds perceived roughness R is on average proportional to $m^{1.6}$. (Figure 1)

Modulation frequency and carrier frequency. Roughness shows a bandpass characteristic in dependence on modulation frequency and carrier frequency. The bandpass ranges from about 12 to 300 Hz, with its maximum depending on the carrier frequency. It varies from a modulation frequency of about 30 Hz for a carrier frequency of 125 to 70 Hz for carrier frequencies of 1 kHz and higher. (Figure 2)

Sound pressure level. Sound pressure level has a minor but significant influence on roughness. Roughness increases by about 25 % when the SPL is increased by 10 dB. This implies on the

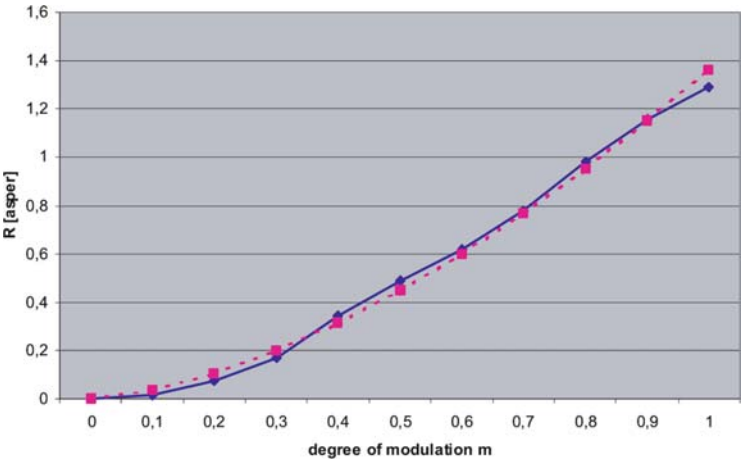


FIGURE 1 Roughness in dependence of the degree of modulation for a 1-kHz tone sinusoidal modulated at 70 Hz (SPL = 70 dB): jury test data (*dotted line*) from [17] approximated in very good agreement by the power function $m^{1.6}$, *solid line* model data.

other side that a reduction of the loudness of a rough component is often not solving sufficiently a sound quality problem.

Specific roughness. Investigations on roughness summation by Aures, Terhardt and Vogel [2,11,13] showed that total roughness sensation is composed of specific roughness, a concept which is similar to the composition of total loudness. The specific roughnesses of the critical bands simply add up to overall roughness if the contributing modulated parts of the sound do not interfere with the basilar membrane. (Figure 3)

Roughness of noise. Modulated broadband noise attains considerably higher roughness values than amplitude-modulated tones. The maximal roughness of amplitude-modulated broadband noise is about 4 asper. Even if noise is unmodulated, a considerable roughness sensation may occur for narrow band noises due to random envelope fluctuations (Figure 5). Unmodulated broadband noise however is not perceived as rough. As most environmental and technical sounds contain noise, it is very important to model the roughness of modulated and unmodulated noise correctly.

3 • Roughness Calculation

Two main approaches can be distinguished for the calculation of roughness: (1) calculating roughness out of an effective degree of modulation [2,5] and (2) calculating roughness out

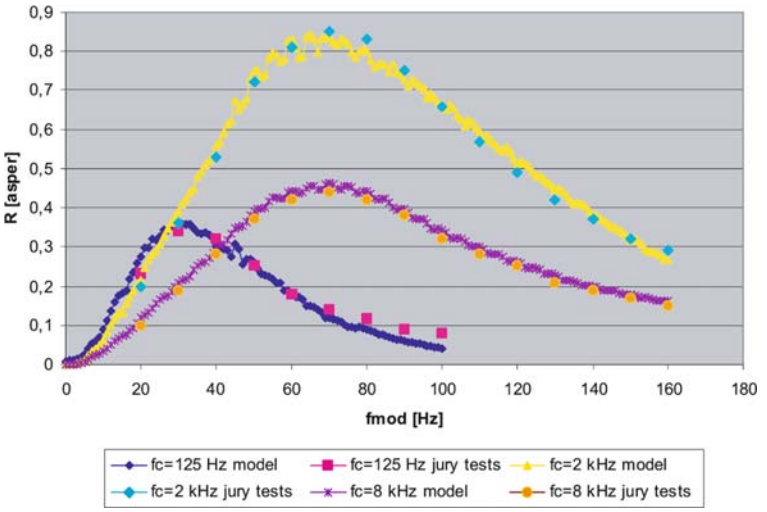


FIGURE 2 Roughness of 100 % amplitude-modulated tones at different carrier frequencies in dependence of modulation frequency: model data (*solid lines*) and data from jury tests from [17] (bigger symbols, see legend).

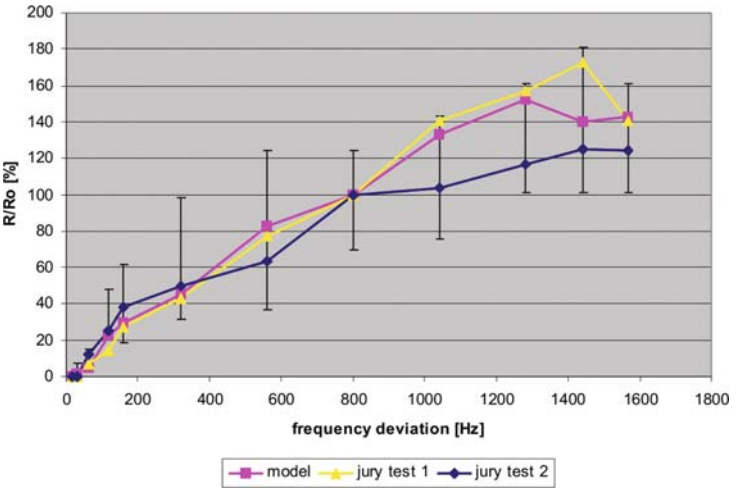


FIGURE 3 Relative roughness of frequency-modulated 1.6 kHz tones ($f_{\mathrm{mod}} = 70$ Hz, $L = 60$ dB) in dependence of frequency deviation. Two different standards were used in the jury tests from [7] (*black* and *white curve*). Quartiles are plotted for better visibility only for the jury test with weaker standard (an amplitude modulated 1.6 kHz tone). Data from model (*gray curve*) are within the jury test data.

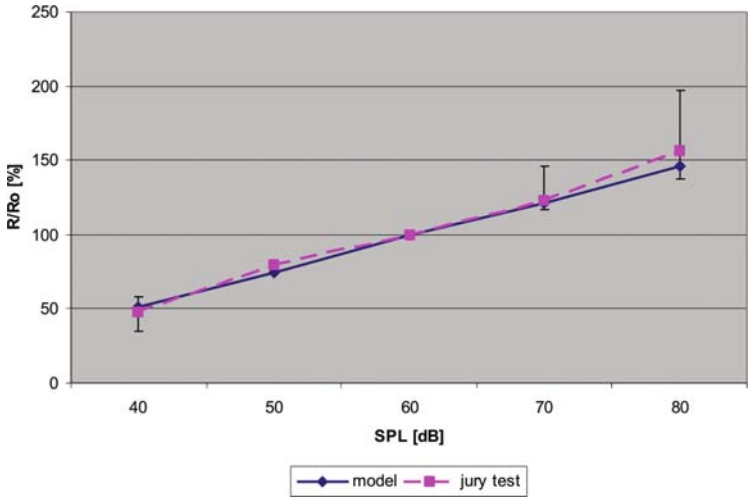


FIGURE 4 Relative roughness of a 1.6-kHz FM tone in dependence of level: *solid line* data from model; *dotted line* data from jury test of [7].

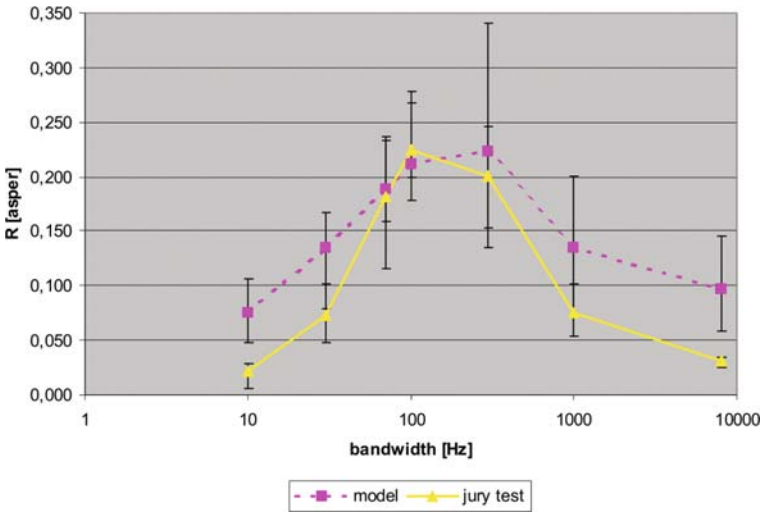


FIGURE 5 Roughness of unmodulated bandpass noise in dependence of bandwidth. Model data (*solid line* with quartiles of 21 calculated values) and jury test data from [1] (*dotted line* with quartiles).

of the product of modulation frequency and excitation level differences [6,16,17]. There are several different implementations available on the market including combinations of the two approaches. Their results differ often considerably for the same sound, and especially for low roughness values, the precision is often not sufficient to decide whether a sound contains rough components or not. Due to overestimation of specific roughness of uncorrelated noise events in the critical bands by various roughness models, calculated roughness of technical sounds often exceeds, e.g., the roughness of 0.1 asper by far, although they are not perceived as rough by experienced test persons. On a 5-point equally distant category scale ranging from not to very rough, category 3 (medium roughness) ranged, e.g., already at 0.1 asper, whereas a value of 0.05 asper near the threshold of roughness is ranked to be slightly rough [4].

3.1 Model Description

The goal of the described model is to provide calculated roughness values with a precision within the range of just noticeable difference of psychoacoustical roughness. As reference values for psychoacoustical roughness, the data of [17] are used. As a basis for this approach, the model of Aures [2] was selected in [4,5] due its flexibility. Several modifications of the model of Aures have been necessary in order to reduce deviations between the calculated roughness and data from jury tests below the just noticeable difference of 17 %. The main changes are doubling of the analysis window from 100 ms in the model of Aures to 200 ms and a doubling of the analysis channels to model psychoacoustical roughness of amplitude-modulated tones in dependence of modulation and carrier frequency correctly. Although the correlation of the envelopes of neighboring critical bands is already taken into account in the model of Aures, a stronger reduction of roughness of uncorrelated narrow band noises had to be introduced in addition.

The structure of the model is shown in Figure 6 and it is summarized in the following (a more detailed description is given in [4,5]).

The roughness is calculated from time frames of 200 ms weighted by a Blackman window. After a discrete Fourier transformation, the spectrum is multiplied by the factor $a_0(f)$ (representing the transmission between free field and the peripheral hearing system) [17]. Then each frequency component is transformed into a triangular excitation pattern on the bark scale, modeling its excitation at the different places on the basilar membrane. The excitation level of the component is set to the level of the Fourier component at the place on the bark scale which corresponds to the frequency of the component.

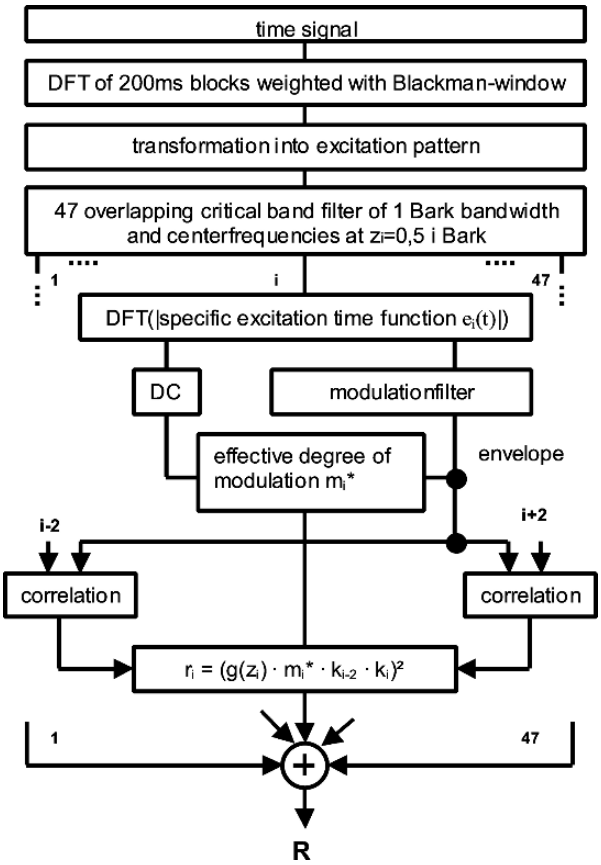


FIGURE 6 Structure of the roughness model.

The excitation decreases with a constant slope toward lower frequencies and with a level dependent slope toward higher frequencies. The slopes are chosen in accordance with the proposal of Terhardt [3]:

$S_1 = -27\,\text{dB}$ /Bark
for the lower slopes and
$S_2 = [-24 - 0.23\text{kHz}/f + 0.2L/\text{dB}]$ dB/Bark
for the level-dependent upper slopes, with f and L denoting frequency and level of the spectral component concerned.

The changes in levels and slopes between different frames result in a time-dependent filtering characteristic.

After the transformation into an excitation pattern, specific excitations are calculated within 47 overlapping 1-Bark-wide channels with equally spaced center frequencies z_i at

$$z_i = 0.5 \cdot i \text{ [Bark]}, \quad i = 1, 2, \ldots, 47.$$

A specific excitation spectrum in a channel i is obtained by linearly superimposing the excitation contributions in an interval $[z_i - 0.5,\ z_i + 0.5]$. The contribution of a spectral component is determined as follows: If the Bark value $z(f)$ of the frequency f of a component is higher than the upper limit of the interval $z(f) > z_i + 0.5$ Bark, the excitation of its lower slope S_1 at $z_i + 0.5$ Bark is taken as the contribution of the component to the specific excitation spectrum of the interval i (for the transformations from the frequency domain into the Bark domain $f \rightarrow z(f)$ the table in [17] is used with linear interpolations). If $z(f) < z_i - 0.5$ Bark, the contribution is the excitation of its upper slope S_2 at $z_i - 0.5$ Bark, and if the component f, i.e., $z(f)$ falls in the interval $z_i - 0.5$ Bark $\leq z(f) \leq z_i + 0.5$ Bark, its unchanged level L fully contributes to the specific excitation spectrum of the channel i. Excitation contributions which are less than the threshold in quiet L_{TQ} minus the logarithmic transmission factor a_0 in a given channel z_i are omitted (see also [17]). This way of calculating specific excitation spectra in each of the 47 bark channels allows the incorporation of the essential level-dependent nonlinear behavior of the excitation pattern into the model and is therefore superior to a constant linear filterbank as input stage. It is equivalent to the use of trapezoidal critical-band filters with level-dependent lower slopes.

The original phases of the frame spectrum are used in a subsequent inverse discrete Fourier transform of each of the 47 specific excitation spectra, yielding 200-ms-long specific excitation time functions $e_i(t)$.

The fluctuations of the envelope are contained in the low-frequency part of the spectrum $F(|e_i(t)|)$ of the absolute value of the specific excitations. This spectrum is appropriately weighted in order to model the bandpass characteristic of the roughness on modulation frequency. The weighting functions $H_i(f_{\text{mod}})$ have been optimized for that purpose.

The shapes of the weighting functions $H_i(f_{\text{mod}})$ of the channels i are described in detail as they have a decisive influence on the final relation between the total roughness R and the modulation frequency f_{mod} for AM tones. The spectral weighting functions $H_4(f_{\text{mod}})$ at $z_4 = 2$ Bark, $H_{16}(f_{\text{mod}})$ at $z_{16} = 8$ Bark and $H_{42}(f_{\text{mod}})$ at $z_{42} = 21$ Bark are drawn in Figure 7. They are basic in the sense that the weighting functions in the other channels can be derived from them in the following way:

for $i = 1, 2, .., 4$ ($z_i = 0.5, 1, .., 2$ Bark)
$H_1 = H_2 = H_3 = H_4$

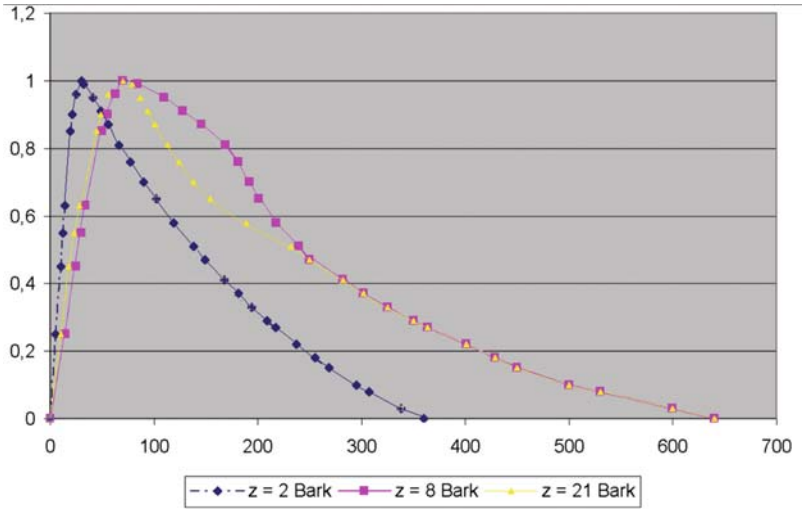


FIGURE 7 Weighting functions $H_i(f_{\text{mod}})$ of the channels $i = 2, 16, 42$.

for $i = 6, 8, \ldots, 16$ ($z_i = 3, 4, .., 8$ Bark)
$H_{i-1} = H_i$ and H_i is linearly interpolated between H_4 and H_{16}
for $i = 17, 18, 19, 20$ ($z_i = 7.5, 8, \ldots, 10$ Bark) $H_i = H_{16}$
for $i = 22, 24, \ldots, 42$ ($z_i = 11, 12, \ldots, 21$)
$H_{i-1} = H_i$ and H_i is linearly interpolated between H_{20} and H_{42}
for $i = 43, 44, \ldots, 47$ ($z_i = 21.5, 22, .., 23.5$ Bark) $H_i = H_{42}$

The time functions of the bandpass-filtered envelopes $h_{\text{BP},i}(t)$ are calculated via inverse Fourier transform.

A generalized degree of modulation m_i^* is calculated by

$$m_i^* = \frac{\tilde{h}_{\text{BP}i}}{\text{DC}_i} \quad \text{if} \quad \frac{\tilde{h}_{BPi}}{\text{DC}_i} \leq 1$$
$$\text{else} \quad m_i^* = 1.$$

with $\tilde{h}_{\text{BP}i}$ equal to the rms value of the filtered envelope $h_{\text{BP}}(t)$ and its DC value before the modulation filtering:

$$DC = \overline{|e_i(t)|}.$$

The roughness dependency on the frequency of the carrier is introduced into the model by multiplying the calculated m_i^* by a weighting function $g(z_i)$ with factors from 0.6 to 1.1 according to the dependence of roughness on the carrier frequency of amplitude-modulated tones.

A simple addition of the squares of $g(z_i) \cdot m_i^*$ would overestimate the total roughness if the modulations in the different

channels are uncorrelated as in the case of noise events. Thus, the weighted generalized degree of modulation $g(z_i) \cdot m_i^*$ of each band is multiplied by the crosscorrelation coefficients k_{i-2} between the envelopes of the channels k_{i-2} and i and k_i calculated from channels i and $i+2$ with $\Delta z = 1$ Bark in both cases.[2] The squared product of the crosscorrelation factors instead of their average as used by Aures [1] has proved to be more effective in reducing roughness of unmodulated noise. The specific roughness is given by

$$r_i = (g(z_i) \cdot m_i^* \cdot k_{i-2} \cdot k_i)^2.$$

The use of the exponent 2 for defining the specific roughness does not necessarily mean that the overall relation between the degree of modulation and the entire calculated roughness will also be of power 2 as can be seen in Figure 1. The total roughness is given by

$$R = 0.25 \sum_{i=1}^{47} r_i \text{ [asper]}.$$

The calibration factor 0.25 is chosen such that the model produces a roughness of 1 asper for a 1-kHz tone at 60 dB SPL that is 100 % amplitude modulated with a modulation frequency of 70 Hz.

3.2 Discussion

The model is fitted to the roughness of amplitude-modulated tones in dependence on modulation frequency and carrier frequency. Results for the degree of modulation and modulation frequency are shown for typical carrier frequencies in Figures 1–3 together with jury test data from [17]. A more detailed analysis can be found in [4,5]. The differences between results from jury tests and the model are in general within the just noticeable difference of 17 %.

The roughness of unmodulated bandpass noise is also predicted fairly well by the model. Typical examples are shown in Figure 4 for bandpass noises centered at 4 kHz with varying bandwidth. Model data (solid line) with quartiles for 21 calculated values are plotted together with data from jury tests from [1]. Data for other center frequencies fit even better to the

[2] In the regions of $z = 0.5$, 1 Bark and $z = 23$, 23.5 Bark only the possible crosscorrelation factors k_i and k_{i-2}, respectively, are used. A compensation for the reduced correlation is given by the enlarged threshold of hearing in these regions.

data from jury tests [5]. For a bandwidth of 10 Hz the random fluctuation is about 6 Hz and thus calculated roughness near 0 asper. The jury test data are also near or below the threshold of roughness of 0.07 asper.

The model was verified by several tests, especially by comparing calculated roughness of frequency-modulated sounds with data from jury tests. Figure 7 shows the relative roughness of frequency-modulated 1.6 kHz tones ($f_{mod} = 70$ Hz, $L = 60$ dB) in dependence of frequency deviation. Two different standards were used in the jury tests from [7] (black and white curves). Quartiles are plotted for better visibility only for the jury test with weaker standard (an amplitude-modulated 1.6 kHz tone). Data from model (gray curve) are within the jury test data. Figure 6 shows the dependence of roughness of frequency-modulated tones on sound pressure level. The data from jury tests from [7] are very well predicted by the model without being fitted to this parameter.

References

1. Aures, W.: Berechnungsverfabren für den Wohcklang Selicdiger Schallsignale, ein Beitrog zurigehōrsezogenen Schall analyse. PhD thesis, Tu München (1984).
2. Aures, W.: Ein Berechnungsverfahren der Rauhigkeit. Acustica 59, 130–141 (1985).
3. Terhardt, E.: Calculating virtual pitch. Hearing Research 1, 155–182 (1979).
4. Daniel, P.: Berechnung und kategoriale Beurteilung der Rauhigkeit und Unangenehmheit von synthetischen und technischen Schallen. Fortschritt-Berichte VDI Reihe 10 Nr. 465. VDI Verlag, Düsseldorf (1997).
5. Daniel, P. and Weber, R.: Psychoacoustical roughness: implementation of an optimized model. Acustica – Acta Acustica 81, 1–12 (1995).
6. Fastl, H.: Roughness and temporal masking patterns of sinusoidal amplitude modulated broadband noise. Psychophysics and Physiology of hearing, Academic Press, London, 403–414 (1977).
7. Kemp, S.: Roughness of frequency-modulated tones. Acustica 50, 126–133 (1992).
8. Mellert, V. and Weber, R.: Gehōrbezogene Verfahren zur Lārmsenrteilung. Walcher, K.P. and Schick, A., editors, Bedentungslehre des Schalls, Bern, Verlag Peter lang, 183 (1984).
9. Roederer, J.: Physikalische und psychoakustische Grundlagen der Musik. Springer, Berlin, Heidelberg (1977).
10. Terhardt, E.: Über akustische Rauhigkeit und Schwankungsstärke. Acustica 20, 215–224 (1968).
11. Terhardt, E.: On the perception of periodic sound fluctuation (roughness). Acustica 30, 201–213 (1974).

12. Über ein Äguivalenzgesetz für Intervalle akustischer Empfindungsgrōβen. Kybernetik 5, 127–133 (1968).
13. Vogel, A.: Roughness and its relation to the time-pattern of psychoacoustical excitation, 241. Springer, Berlin, Heidelberg, New York (1974).
14. Vogel, A.: Über den Zusammenhang zwischen Rauhigkeit und Modulationsgrad. Acustica 32, 300–306 (1975).
15. von Helmholtz, H.: die Lehre von den Tonempfindungen. Wissenschaftliche Buchgesellschaft, Darmstadt, 7. unveränderte Auflage (1. Auflage Braunschweig 1863), (1968).
16. Widmann, U.: Ein Modell der Psychoakustischen Lästigkeit von Schallen und seine Anwendung in der Praxis der Lärmbeurteilung. PhD thesis, TU München (1992).
17. Zwicker, E. and Fastl, H.: Psychoacoustics. Springer, Berlin, Heidelberg (1990).

PART IV
MUSICAL ACOUSTICS

James W. Beauchamp

University of Illinois at Urbana-Champaign, Urbana, IL, USA

20 Automatic Music Transcription

Anssi Klapuri and Tuomas Virtanen 277

Introduction • Musical Sounds and F0 Estimation • Auditory Model-Based Multiple-F0 Analysis • Sound Separation Using Sinusoidal Modeling • Statistical Inference Within Parametric Signal Models • Unsupervised Learning Techniques • Summary and Conclusions

21 Music Structure Analysis from Acoustic Signals

Roger B. Dannenberg and Masataka Goto 305

Introduction • Features and Similarity Measures • Segmentation • The Similarity Matrix • Finding Repeating Sequences • Music Summary • Evaluation • Summary and Conclusions

22 Computer Music Synthesis and Composition

Lydia Ayers 333

Introduction • Music with Live Performers • Form, Structure and Time • Timbre and Texture • Tuning Systems • Musique Concrète and Collage • Environments • Conclusion

23 Singing Voice Analysis, Synthesis, and Modeling

Youngmoo E. Kim 359

Introduction • Principles of Singing Voice Modeling • Models of the Glottal Source • Models of the Vocal Tract • Synthesis of the Singing Voice • Conclusion

24 Instrument Modeling and Synthesis

Andrew B. Horner and James W. Beauchamp 375

Introduction • Spectral Analysis • Wavetable Synthesis • Wavetable Matching • Frequency Modulation (FM) Synthesis Models • Frequency Modulation Matching • Conclusions

25 Digital Waveguide Architectures for Virtual Musical Instruments

Julius O. Smith 399

Introduction • Vibrating Strings • Wind Instruments • Recommended Reading

26 Modeling of Musical Instruments

Rolf Bader and Uwe Hansen 419

Introduction • Finite-Element Formulation • Finite-Difference Formulation • Turbulent $k-\epsilon$ Model of Flute-Like Instruments • Multiphonics in Blown Instruments in Terms of a Split-Phase Pressure Description • Time-Dependent Model of a Classical Guitar Body • Guitar Back-Plate Timbre Brightening by Stress Bending • Church Bell Clapper Time-Dependent Striking Force

20

Automatic Music Transcription

Anssi Klapuri and Tuomas Virtanen

Institute of Signal Processing, Tampere University of Technology, Tampere, Finland

1 • Introduction

Written musical notation describes music in a symbolic form that is suitable for performing a piece using the available musical instruments. Traditionally, musical notation indicates the pitch, target instrument, timing, and duration of each sound to be played. The aim of music transcription either by humans or by a machine is to infer these musical parameters, given only the acoustic recording of a performance. In terms of data representations, this can be seen as transforming an audio signal into a MIDI[1] file. Signals of particular interest here are *polyphonic* music signals where several sounds are playing simultaneously.

Automatic recovery of the musical notation of an audio signal allows modifying, rearranging, and processing music at a high abstraction level and then resynthesizing it again. Structured audio coding is another important application: For example, a MIDI-like representation is extremely compact yet retains the characteristics of a piece of music to an important degree. Other uses of music transcription comprise information retrieval,

[1] Musical instrument digital interface (MIDI) is a standard for exchanging performance data and parameters between electronic musical devices.

musicological analysis of improvised and ethnic music, and interactive music systems which generate an accompaniment to the singing or playing of a soloist.

Attempts to automatically transcribe polyphonic music have been reported over a timescale of about 30 years, starting from the work of Moorer in 1970s [50].

State-of-the-art transcription systems are still clearly behind human musicians in accuracy and flexibility, but considerable progress has been made during the last 10 years, and it is the purpose of this chapter to survey selected literature representing these advances. The main emphasis will be on signal processing methods for resolving mixtures of pitched musical instrument sounds. Methods will be discussed both for estimating the fundamental frequencies (F0s) of concurrent sounds and for separating component sounds from a mixture signal. Other subtopics of music transcription, in particular beat tracking, percussion transcription, musical instrument classification, and music structure analysis, will not be discussed here, but an overview can be found in [37].

This chapter is organized as follows. Section 2 will introduce musical sounds and the basic principles of F0 estimation. Section 3 will discuss multiple-F0 estimators that follow principles of human hearing and pitch perception. Section 4 will describe sound separation methods that are based on the sinusoidal representation of music signals. Section 5 will discuss statistical inferences from parametric signal models. Section 6 describes unsupervised learning methods which make no assumptions about the nature of the sound sources and are thus suitable for percussion transcription, too. Concluding remarks are made in Section 7.

2 • Musical Sounds and F0 Estimation

In the majority of Western music, melody and harmony are communicated by *harmonic* sounds. These are sounds that are nearly periodic in the time domain and show a regular spacing between the significant spectral components in the frequency domain (Figure 1). An ideal harmonic sound consists of frequency components at integer multiples of its F0. In the case of plucked and struck string instruments, however, the partial frequencies are not in exact integral ratios but obey the formula $f_j = jF\sqrt{1+\beta(j^2-1)}$, where F is the fundamental frequency, j is the partial index, and β is an inharmonicity factor due to the stiffness of real strings [20]. Despite this imperfection, the general structure of the spectrum is similar to that in Figure 1.

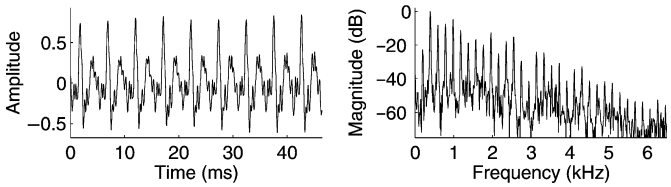


FIGURE 1 A harmonic sound in the time and frequency domains. The example represents a violin sound with fundamental frequency 196 Hz and fundamental period 5.1 ms.

Some of the methods discussed later in this chapter assume harmonic sounds as input. This limitation is not very severe in Western music where harmonic sounds are produced by most instrument families, namely by bowed and plucked string instruments, brass and reed instruments, flutes, pipe organs, and the human voice. Outside this category are mallet percussion instruments (e.g., vibraphone, xylophone, and marimba) whose waveforms are nearly periodic but whose spectra are definitely inharmonic.

A large number of different methods are available for monophonic F0 estimation. For speech signals, comparative evaluations of these can be found in [56,26,13]. Here, only the main ideas of different algorithms are introduced so as to provide a background for the analysis of polyphonic music signals described in detail in subsequent sections.

Algorithms that measure the periodicity of a time-domain signal have been among the most frequently used F0 estimators [64,14]. As noted in [13], quite accurate F0 estimation can be achieved simply by an appropriate normalization of the autocorrelation function (ACF) $r(\tau)$. Then, the F0 can be computed as the inverse of the lag τ that corresponds to the maximum of $r(\tau)$ within a predefined lag range. The ACF of a signal $x(n)$ within a K-length analysis frame is given by

$$r(\tau) = \frac{1}{K} \sum_{k=0}^{K-1-\tau} x(k)x(k+\tau). \quad (1)$$

Figure 2 illustrates the ACF for the violin waveform shown in Figure 1.

An indirect way of measuring time-domain periodicity is to match a harmonic pattern to the signal in the frequency domain. According to the Fourier theorem, a periodic signal with period τ can be represented with a series of sinusoidal components at frequencies j/τ, where j is a positive integer.

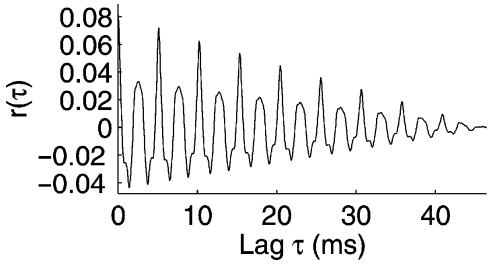


FIGURE 2 Autocorrelation function calculated within a 46 ms time frame for the violin waveform of Figure 1.

As an example, Doval and Rodet [16] performed maximum-likelihood spectral pattern matching to find the F0 which best explained the observed frequency partials. Other strategies for spectral pattern matching have been proposed by Brown [6] and Maher and Beauchamp [44]. The time-domain ACF calculation is closely related to these methods since, according to the Wiener–Khintchine theorem [25, p. 334], the ACF of a signal $x(n)$ is the inverse Fourier transform (IDFT) of its power spectrum $|X(k)|^2$. Writing out the IDFT we get

$$r(\tau) = \frac{1}{K} \sum_{k=0}^{K-1} \cos\left(\frac{2\pi\tau k}{K}\right) |X(k)|^2, \tag{2}$$

where K is the size of the analysis frame. It is easy to see that Eq. (2) emphasizes frequency components at harmonic spectral locations (integer multiples of $k = K/\tau$) because $r(\tau)$ is maximized when the cosine maxima line up with the frequencies of the harmonics (Figure 4).

Another class of F0 estimators measures the periodicity of the Fourier spectrum of a sound [40]. These methods are based on the observation that harmonic sounds have a quasi-periodic magnitude spectrum, the period of which is the F0. In its simplest form, the autocorrelation function $\tilde{r}(m)$ of a K-length magnitude spectrum is calculated as

$$\tilde{r}(m) = \frac{2}{K} \sum_{k=0}^{K/2-m-1} |X(k)||X(k+m)|. \tag{3}$$

The above formula bases F0 calculations on a fundamentally different type of information than the time-domain ACF: Here, any two frequency components with a certain spectral interval support the corresponding F0. An interesting difference between this method and the time-domain ACF is that measuring the periodicity of the time-domain signal is prone to F0 halving since the signal is periodic at twice the fundamental period

too, whereas the methods that measure the periodicity of the magnitude spectrum are prone to F0 doubling since the spectrum is periodic at twice the F0 rate, too. There are ways to combine these complementary approaches to improve the overall result, as will be discussed in the next section.

3 • Auditory Model-Based Multiple-F0 Analysis

The human auditory system is very efficient at analyzing sound mixtures. It is therefore reasonable to learn from its function as much as possible, especially since the peripheral parts of hearing are relatively well known, and precise auditory models exist which are able to approximate the signal in the auditory nerve [49]. This enables the computation of a data representation similar to that used by the central auditory system.

3.1 Model of the Auditory Periphery

Computational models of the peripheral auditory system comprise two main parts which can be summarized as follows:

(1) An acoustic input signal is passed through a bank of bandpass filters (aka *channels*) which represents the frequency selectivity of the inner ear. Typically, about 100 filters are used with center frequencies f_c uniformly distributed on a critical-band scale,

$$f_c = 229(10^{\xi/21.4} - 1), \tag{4}$$

where ξ is the critical band number. Usually *gammatone* filters are used, with the bandwidths b_c of the filters obeying $b_c = 0.108 f_c + 24.7\,\text{Hz}$.

(2) The signal at each band is processed to simulate the transform characteristics of *hair cells* that produce neural impulses to the auditory nerve [27]. In signal processing terms, this involves compression, half-wave rectification, and lowpass filtering.

These seemingly very simple steps account for some important properties in pitch perception. In particular, the half-wave rectification operation in subbands allows a synthesis of the time-domain and frequency-domain periodicity analysis mechanisms discussed in Section 2. The half-wave rectification (HWR) is defined as

$$HWR(x) = \max(x, 0) = \frac{1}{2}(|x| + x). \tag{5}$$

Figure 3 illustrates the HWR operation for a subband signal which consists of five overtones of a harmonic sound. What is

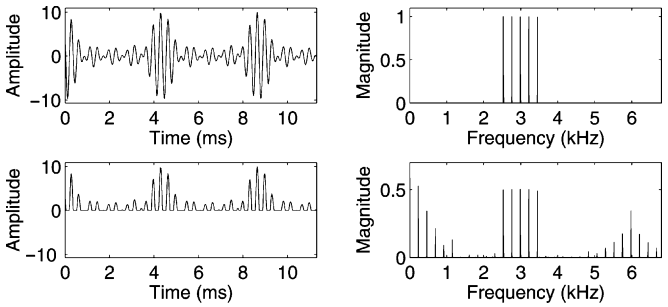


FIGURE 3 *Upper* panels show a signal consisting of the overtone partials 11–15 of a sound with F0 230 Hz in the time and the frequency domains. Lower panels illustrate the signal after half-wave rectification.

important is that the rectification generates spectral components that correspond to the frequency intervals between the input partials. The components below 1 kHz in the figure represent the amplitude envelope of the input signal. Any signal that consists of more than one frequency component exhibits periodic fluctuations, beating, in its time-domain amplitude envelope. That is, the partials alternatingly amplify and cancel out each other, depending on their phase. The rate of the beating caused by each pair of two frequency components depends on their frequency difference and, for a harmonic sound, the frequency interval corresponding to the F0 dominates.

After the rectification, periodicity analysis of the subband signals using, for instance, the ACF, would make use of both time and frequency domain periodicity by matching a harmonic pattern to both the input partials and the beating partials, which leads to more reliable F0 analysis [39]. Balance between the two types of information can be determined by applying a lowpass filter which partly suppresses the original passband at higher subbands but leaves the spectrum of the amplitude envelope intact.

3.2 Pitch Perception Models

The above-described two steps produce a simulation of the signal that travels in the different fibers ("channels") of the auditory nerve. Pitch perception models try to explain how this signal is further processed to derive a stationary percept of pitch. The processing mechanisms in the brain are not accurately known, but the prevailing view is that periodicity analysis of some form takes place for the signals within each auditory channel, and the results are then combined across channels [11]. Meddis and Hewitt [48] implemented these two steps as follows:

3) ACF estimates $r_c(\tau)$ are computed within channels.
4) The ACFs are summed across channels to obtain a summary autocorrelation function $s(\tau) = \sum_c r_c(\tau)$. The maximum of $s(\tau)$ within a predefined lag range is used to predict the perceived pitch.

Together with the above-mentioned peripheral processing stages, this became known as the "unitary model" of pitch perception, and the authors showed that it can reproduce many important phenomena in pitch perception, such as missing fundamental, repetition pitch, and pitch shift of equally spaced inharmonic components [48].

Some music transcription systems have applied an auditory model to compute an intermediate data representation that is then used by a higher level inference procedure. Martin [46] proposed a system for transcribing piano performances of four-voice Bach chorales. His system used the log–lag correlogram model of Ellis [17] (similar to the unitary model) as a front end to an inference architecture, where knowledge about sound production was integrated with rules governing tonal music. More recently, Marolt [45] used adaptive oscillators and neural networks to detect notes at the output of the peripheral hearing model described above.

3.3 Extensions of the Unitary Pitch Model

The pitch perception model outlined above is not sufficient as such for accurate multiple-F0 estimation. It suffers from certain shortcomings, and the following described methods can be seen as attempts to alleviate these problems. Most importantly, the model accounts only for a single simultaneous pitch. Several pitches in a mixture signal cannot be detected simply by picking several local maxima in the summary ACF. De Cheveigné and Kawahara [12] addressed this problem by proposing a system where pitch estimation was followed by the cancellation of the detected sound in order to reveal the other sounds in the mixture. The cancellation was performed either by subband selection or by performing within-band cancellation filtering, and the estimation step was then iteratively repeated for the residual signal.

Tolonen and Karjalainen [65] addressed the computational complexity of the unitary model by first prewhitening the input signal and then dividing it into only two subbands, below and above 1 kHz. A "generalized ACF" was then computed for the lower channel signal and for the amplitude envelope of the higher channel signal, and the two ACFs were summed. Despite the drastic reduction in computation compared to the original unitary model, many important characteristics of the model were preserved. Extension to multiple-F0 estimation was

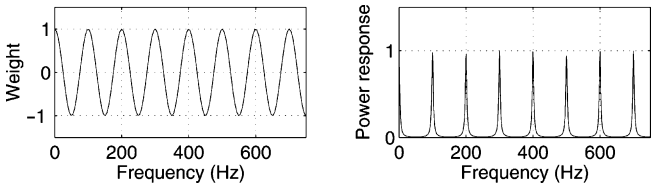


FIGURE 4 The *left* panel shows the weights $\cos(2\pi\tau k/K)$ of ACF calculation in (2), when the lag τ is 10 ms (corresponding to 100 Hz). The *right* panel shows the power response of a comb filter with a feedback lag of 10 ms.

achieved by cancelling subharmonics in the summary ACF, by clipping the summary ACF to positive values, timescaling it to twice its length, and by subtracting the result from the original clipped summary ACF. This cancellation operation was repeated for scaling factors up to about five. From the resulting enhanced summary autocorrelation function, all F0s were picked without iterative estimation and cancellation. The method is relatively easy to implement and produces good results when the component F0s are clearly below 1 kHz.

The generally weak robustness of the unitary model in polyphonic music was addressed by Klapuri [39]. He used a conventional peripheral hearing model (Section 3.1) but replaced the ACF by a periodicity analysis mechanism where a bank of comb filters was simulated in the frequency domain. Figure 4 illustrates the power response of a comb filter for a time lag candidate τ and, for comparison, the corresponding weights $\cos(2\pi\tau k/K)$ of the ACF calculation in (2). Giving less emphasis to the spectrum between the harmonic partials of an F0 candidate alleviates the interference of concurrent sounds. Estimation of multiple F0s was achieved by canceling each detected sound from the mixture and repeating the estimation for the residual, similarly to [12]. Quite accurate multiple-F0 estimation was achieved in the pitch range 60–2100 Hz.

4 • Sound Separation Using Sinusoidal Modeling

Sound separation refers to the task of estimating the signal produced by an individual sound source from a mixture. Multiple-F0 estimation and sound separation are closely related: An algorithm which achieves sound separation facilitates multiple-F0 estimation, and vice versa. Even when the F0s of musical tones have already been found, additional information obtained using sound separation is useful for classifying the tones to sources. For example, similar F0s and spectra in adjacent

frames can be concatenated to form notes, and notes with different F0s can be grouped to obtain the whole passage of an instrument.

An efficient initial decomposition for the sounds produced by musical instruments is afforded by the sinusoids plus noise model, which represents the signal as a sum of deterministic and stochastic parts, or, as a sum of a set of sinusoids plus a noise residual [4,58]. Sinusoidal components represent the vibration modes of sound sources, and the residual contains the energy produced by the excitation mechanisms and other components which are not a result of these modes. The deterministic part of the model, which is called the sinusoidal model, has been used widely in audio signal processing, for example in speech coding by McAulay and Quatieri [47]. In musical signal processing, it became known by the work of Smith and Serra [61].

The usefulness of the sinusoidal model in sound source separation and automatic music transcription stems from its physical plausibility. Since all sinusoids correspond to the vibrating modes of the sources in a mixture, it is plausible to map the sinusoids into the individual sound sources and to estimate higher level information such as notes played by each source. Sound source separation algorithms which apply the sinusoidal model can be roughly divided into three categories: (1) methods which first estimate sinusoids and then group them into sound sources, (2) methods which first estimate F0s of the sources and then estimate sinusoids using partial frequencies predicted by the F0s, and (3) methods, which jointly estimate the number of sources, their F0s, and parameters of the sinusoids. The general signal model is discussed in Section 4.1, and the three separation approaches are discussed in Sections 4.2, 4.3, and 5, respectively.

4.1 Signal Model

The parameters of natural sound sources are usually slowly varying; therefore, the signal is analyzed in short segments or frames. In general, the parameters are assumed to be fixed during each frame, although continuous time-varying parameters are used in some systems. The sinusoidal model for one frame of a music signal can be written as

$$x(n) = \sum_{j=1}^{J} a_j \cos(2\pi f_j n + \theta_j) + e(n), \tag{6}$$

where the J sinusoids represent the harmonic partials of all sources, n is the time index, a_j, f_j, and θ_j are the amplitude,

frequency, and phase of the jth sinusoid, respectively, and $e(n)$ is the residual.

Basic algorithms for the estimation of sinusoids from music signals have been reviewed by Rodet [57] and by Serra [58], and a theorethical framework for the parameter estimation is discussed by Kay [35]. Usually the estimation is done in the frequency domain. The frequencies of sinusoids can be estimated by picking the most prominent peaks from the magnitude spectrum. Many useful practical details for estimating the peaks are given in [3]. Also, matching pursuit algorithms have been used, which use a dictionary of time-domain elements to decompose the signal. For harmonic sounds, the use of harmonic atoms [24] provides a decomposition which is a good basis for the analysis of music signals.

4.2 Grouping Sinusoids to Sources

Psychoacoustically motivated methods have been among the most widely used in sound source separation. The cognitive ability of humans to perceive and recognize individual sound sources in mixture signals is called *auditory scene analysis* [5]. Computational models of this function typically consist of two main stages where an incoming signal is first decomposed into its elementary time–frequency components and these are then organized to their respective sound sources.

Bregman pointed out a number of measurable acoustic "cues" which promote the grouping of time–frequency components to a common sound source as perceived by human listeners [5]. Among these are proximity in time–frequency, harmonic frequency relationships, synchronous changes in the frequency or amplitude of the components, and spatial proximity (i.e., the same direction of arrival). In the computational modeling of auditory scene analysis, the most widely used cues are the proximity in time and harmonic frequency relationships; usually the signals are analyzed in short frames, and they are assumed to be harmonic or close to harmonic.

Kashino and Tanaka [34] implemented a subset of Bregman's cues for the purpose of sound source separation in music signals. Using the sinusoidal model as a decomposition, the authors viewed the source separation task as two grouping problems where sinusoids were first clustered to sound events which were then grouped to particular sound sources (instruments). Harmonic mistuning and onset asynchrony were used as the cues to initialize new sound events, and sinusoids were then grouped to these. Grouping the sound events to their respective sources was achieved by using timbre models and an "old-plus-new" heuristic. The latter means that a complex sound is interpreted as a combination of

"old sounds" as much as possible, and the remainder is perceived as a "new sound". Evaluation results were shown for polyphonies up to three simultaneous sounds and for several different instruments. In a subsequent paper, the authors implemented the system in the framework of a Bayesian probability network and integrated musical knowledge to it [33].

The two-way mismatch procedure proposed by Maher and Beauchamp implements grouping based on harmonic frequency relationships [44]. Their method can be used to find the most likely F0s of harmonic sounds within one frame, given the frequency and amplitude estimates of prominent sinusoids. The fundamental frequencies are chosen so that the mismatch error between the estimated sinusoids and the partials predicted from trial F0 values is minimized. The name two-way mismatch stems from a procedure in which each estimated sinusoid is matched to the closest predicted partial, and each predicted partial is matched to the closest estimated sinusoid, and the total mismatch is measured as the average discrepancy in all the matches. The exact procedure and mismatch function is explained in [44]. The method requires a search for possible F0s within a designated range, and the number of sources has to be set manually. The algorithm is relatively straightforward to implement and has been used for the separation of duet signals [43] and to analyze the melody line in a polyphonic recording [53].

Sterian [62] implemented perceptual grouping rules as a set of likelihood functions, each of which evaluated the likelihood of the observed sinusoids given a hypothesized grouping of the sinusoids to note candidates. Distinct likelihood functions were defined to take into account onset and offset timing, harmonicity, low partial support, partial gap, and partial density (see the reference for the definitions of the latter concepts). The product of all the likelihood functions was used as a criterion for optimal grouping.

Godsmark and Brown [21] applied an auditory model to estimate dominant time–frequency components at different bands. These were mapped to sound sources by performing grouping according to onset and offset synchrony, temporal and frequency proximity, harmonicity, and common frequency movement.

4.3 Sound Separation Given the Fundamental Frequencies

When the number of concurrent sounds is high, the robustness of the grouping methods presented in the previous section decreases mainly because overlapping partials (partials whose frequencies are very close or the same) are difficult to resolve at the grouping stage. Estimation of the partials can be done more robustly if some higher level information is available, such as the F0s of the

sources, which can usually be estimated more accurately than individual sinusoids. This section deals with the estimation of sinusoids assuming that the F0s are estimated in advance using some other method.

Usually sinusoidal parameters are estimated by minimizing the reconstruction error between the sinusoidal model and the observed signal. Since the sinusoidal model is nonlinear with respect to frequencies, in general, a global solution for optimal parameters cannot be guaranteed. However, estimated F0s can be used to generate rough predictions of the partial frequencies, and iterative procedures which alternately estimate amplitudes and phases while frequencies are fixed and then update the frequencies, repeating until convergence is achieved, can be used. Most authors have used the energy of the residual to measure the reconstruction error, which leads to least-squares estimation. A general procedure, using a nonlinear least-squares method has been described in [63].

For fixed frequencies, amplitudes and phases can be estimated using the standard least-squares approach [36] which usually produces good results, even when the frequencies of the partials are close to each other. However, when the frequencies of two or more partials are equal, the exact amplitudes and phases of these partials cannot be resolved based on spectral peaks. This phenomenon is very common in musical signals, since F0s are often in harmonic relationships. However, an estimate of the amplitudes can be obtained by interpolating from adjacent frames or partials. The shape of a typical instrument spectrum is slowly varying with respect to time and frequency, so that, in general, interpolation of amplitudes produces tolerable results. The phases are perceptually less important, so they can be easily interpolated to produce smooth transitions between frames.

Quatieri and Danisewicz [55] used adjacent frame interpolation for separation of speech voices. The method is not as useful for musical signals since F0s are often in harmonic relationship for relatively long durations. Maher [43] suggested using interpolation from adjacent partials. However, a problem arises if several adjacent partials overlap. The method proposed by Virtanen and Klapuri [71] performs interpolation implicitly, since it forces smooth amplitude spectra of sources by means of linear overtone series models.

Using the estimated amplitudes and phases, frequencies can be updated, for example, by using an algorithm proposed by Depalle and Hélie [15]. Each partial's frequency can be separately

estimated, or frequency ratios can be fixed, so that only the fundamental frequency is updated, as done by Quatieri and Danisewicz [55] and by Virtanen and Klapuri [70].

If an F0 estimate fails, the signal $x(n)$ will contain components for which no sinusoids have been assigned in the model given by Eq. (6). The additional components may affect the estimate, since both the target and the interfering sounds have a harmonic spectral structure. The effect of harmonic interference can be decreased, for example, by postprocessing the estimated parameters using a perceptually motivated spectral smoothing method proposed by Klapuri [38].

5 • Statistical Inference Within Parametric Signal Models

By making certain assumptions about the component sounds in a music signal, the multiple-F0 estimation problem can be viewed as estimating the parameters of a chosen signal model. In this section, statistical methods to this end are discussed.

5.1 Frequency-Domain Models

Let us first consider a model for the short-term power spectrum of a signal. Here it is useful to apply a logarithmic frequency scale and to measure frequency in cent units. One cent is 1/100 of a semitone in equal temperament and there are 1,200 cents in an octave. The relationship between cents and frequencies in Hertz is given by

$$z = 1200 \log_2(f_{\mathrm{Hz}}/440) + 6900, \tag{7}$$

where the reference frequency 440 Hz is fixed to 6,900 cents.

To enable statistical modeling, the power spectrum $\Psi(z)$ of a signal is considered to be a distribution of small units of energy along the frequency axis. In terms of an underlying random variable ζ, a realization $\zeta = z$ is thought to assign a unit of spectral energy to cents-frequency z. Consider the following parametric model for the spectral energy distribution proposed by Goto [22]:

$$p(\zeta = z|\theta) = \sum_{i=1}^{I} \alpha_i \sum_{j=1}^{J_i} a_{i,j} G(z; \omega_i + 1200 \log_2(j), \sigma), \tag{8}$$

where $p(\zeta = z|\theta)$ represents the probability of observing energy at cents-frequency z given the model parameters θ. The function $\mathrm{G}(z; z_0, \sigma)$ is a Gaussian distribution with mean z_0 and standard deviation σ. The above formula assumes a mixture of I harmonic sounds where ω_i is the fundamental frequency of sound i in cents,

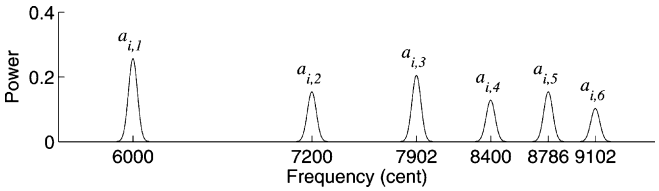


FIGURE 5 Illustration of the parametric model for spectral energy distribution in Eq. (8). Here ω_i is 6,000 cents (260 Hz).

α_i is the gain of sound i, and $a_{i,j}$ are the powers of the partials j of sound i. The shorthand notation θ represents all the parameters, $\theta = \{\omega_i, \alpha_i, a_{i,j}, I, J_i\}$.

Figure 5 illustrates the above model. Intuitively, each of the component sounds is modeled as a mixture of Gaussian distributions centered at integer multiples of F0. The variance around each partial models spreading of spectral energy due to time-domain windowing, F0 variations, and the inharmonicity phenomenon mentioned in Section 2. A nice feature of the model is that, due to the logarithmic frequency scale applied, the actual variance is larger for the higher order partials although σ is common for all.

Goto [23] extended the above model by introducing one additional dimension, so that multiple adaptive tone models were estimated for each F0. In practice, the weights became two-dimensional ($\alpha_{i,m}$) and the adaptive tone models three-dimensional ($a_{i,m,j}$), where m denotes the model index. This was done in order to allow various kinds of harmonic structures to appear at each F0 and to enable the use of several alternative tone model priors (spectral templates) side by side.

The problem to be solved, then, is to infer the model parameters θ given an observed short-time spectrum and possibly some prior information about the parameter distributions. Goto derived a computationally feasible expectation-maximization (EM) algorithm which iteratively updates the tone models and their weights, leading to maximum *a posteriori* parameter estimates. He avoided the explicit estimation of the F0 parameters in Eq. (8) by distributing a large number of F0s uniformly on the log-frequency scale and by estimating only the weights for each F0. The algorithm requires only a few iterations to converge and is suitable for real-time implementation. Goto used the method to detect the melody and the bass lines in real-world CD recordings. Temporal continuity of the estimates was considered by framewise tracking of the F0 weights within a multiple-agent architecture. Although the overall method in [23]

is quite complex, the core EM algorithm is straightforward to implement and particularly suitable for detecting one (predominant) F0 in polyphonic signals.

Kameoka et al. [32] applied the signal model (8) and performed maximum-likelihood estimation of the F0s ω_i, weights α_i, and tone model shapes $a_{i,j}$ using another EM algorithm. In addition, they used an information theoretical criterion to estimate the number of concurrent sounds I. They proposed a two-stage estimation procedure where F0 values were first constrained within a one-octave range and equal amplitudes were used for all harmonics; then the true F0s were found among the harmonics and subharmonics of each candidate and the spectral shape of the sounds was estimated. Promising results were reported for polyphonic music transcription and for pitch estimation of two simultaneous speakers.

5.2 Time-Domain Models

In time-domain models, the main difference compared to frequency-domain models is that the phases of partials have to be taken into account. Walmsley [72] performed parameter estimation in a Bayesian framework using a reformulated sinusoidal model for the time-domain signal:

$$y(t) = \sum_{i=1}^{I} \gamma_i \sum_{j=1}^{J_i} \left[a_{i,j} \cos(j\omega_i t) + b_{i,j} \sin(j\omega_i t) \right] + e(t), \qquad (9)$$

where γ_i is a binary indicator variable switching the sound i on or off, ω_i is the F0 of sound i, and $a_{i,j}$, $b_{i,j}$ together encode the amplitude and phase of individual partials. The term $e(t)$ is a residual noise component. Similarly to Eq. (8), the above model assumes a mixture of harmonic signals, but a significant difference is that here the phases of the partials are also taken into account and estimated.

Another novelty in Walmsley's method was that the parameters were estimated jointly across a number of adjacent frames to increase robustness against transient events. A joint posterior distribution for all the parameters (given an observed signal segment) was defined which took into account the modeling error $e(t)$, prior distributions of the parameters, and the dependencies of the parameters on longer term hyperparameters that modeled frequency variation over time. Optimal parameter estimates were produced by locating regions of high probability in the joint posterior distribution. Whereas this method proved intractable for analytical optimization, Walmsley was able to use Markov chain Monte Carlo (MCMC) methods to generate samples from the posterior. A transition kernel was proposed which consisted

of heuristic rules for the fast exploration of the parameter space. For example, sounds were switched on and off, F0 values were switched between their harmonics and subharmonics, and the residual was analyzed to detect additional notes.

Davy and Godsill [10] extended Walmsley's model to accommodate time-varying amplitudes, nonideal harmonicity, and nonwhite residual noise. They also improved the MCMC algorithm and reconsidered the prior structure. The resulting system was reported to work robustly for polyphonies up to three simultaneous sounds.

Another interesting time-domain model has been proposed by Cemgil et al. [9] who placed emphasis on explicitly modeling the sound generation process. They modeled musical sounds as a sum of harmonically related and damped "noisy" sinusoids drawn from a stochastic process. The damping factors were tied to a global factor which was assumed to be trained in advance. The sound generators were controlled by a collection of binary indicator variables $\gamma_{i,t}$, a "piano roll", which represented the activity of different notes i as a function of time t. A matrix $[\mathbf{\Gamma}]_{i,t} = \gamma_{i,t}$ was used to denote the entire piano roll.

The transcription problem was then viewed as the task of finding a piano-roll $\mathbf{\Gamma}^*$ which maximizes the posterior probability $p(\mathbf{\Gamma}|\mathbf{y})$ given a time-domain signal $\mathbf{y}$. Since this optimization task was analytically intractable, Cemgil et al. developed inference procedures for certain special cases. Chords were identified by a greedy algorithm which started from an initial configuration and iteratively added or removed single notes until the probability of the mixture converged. Piano-roll inference was achieved by analyzing the signal in short time-windows and by assuming that only one note may switch on or off within a window.

6 • Unsupervised Learning Techniques

Recently, information theoretical methods such as independent component analysis (ICA) have been successfully used to solve blind source separation problems in several application areas [30]. ICA is based on the assumption of statistical independence of sources, and in certain conditions, it enables the blind estimation of sources from mixed observations. The term blind means that there is no prior knowledge of the sources. Basic ICA cannot be directly used to separate one-channel time-domain signals, but this becomes possible with a suitable signal representation, as described in Section 6.1. In addition to ICA, other unsupervised learning algorithms such as nonnegative matrix factorization (NMF) [41] and sparse coding have been used in

the analysis and separation of monaural (single-channel) music signals. Unlike ICA, NMF and sparse coding do not aim at statistical independence but just at representing the observed data efficiently. In the case of music signals, this usually results in the separation of sources, at least to some degree. The basic principles of ICA, NMF, and sparse coding are described in Sections 6.2, 6.3, and 6.4, respectively.

6.1 Signal Model for Monaural Signals

Basic ICA and many other blind source separation algorithms require that the number of sensors is greater than or equal to the number of sources. In multichannel sound separation, this means that there should be at least as many microphones as there are sources. However, automatic music transcription usually aims at finding notes in monaural or stereo signals, for which basic ICA methods may not be adequate. The most common approach to overcome this limitation is to represent the input signal in the time-frequency domain, for example by using the short-time Fourier transform (STFT). Basic generative models of ICA, NMF, and sparse coding are linear: For each analysis frame m, the short-time spectrum vector $\mathbf{x}_m$ can be expressed as a weighted sum of basis spectra $\mathbf{b}_j$. The model is not necessarily noise free. Thus, with a residual term $\mathbf{e}_m$, it can be written as

$$\mathbf{x}_m = \sum_{j=1}^{J} g_{j,m}\mathbf{b}_j + \mathbf{e}_m, \tag{10}$$

where J is the number of basis spectra and $g_{j,m}$ is the gain of the jth basis spectrum for the mth frame. For M frames, the model can be written in a matrix form as

$$\mathbf{X} = \mathbf{BG} + \mathbf{E}, \tag{11}$$

where $\mathbf{X} = [\mathbf{x}_1, \ldots \mathbf{x}_M]$ is the spectrogram matrix, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_J]$ is the basis matrix, $[\mathbf{G}]_{j,m}$ is the gain matrix, and $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_M]$ is the residual matrix. Usually $\mathbf{X}$ is given as the input and $\mathbf{B}$, $\mathbf{G}$, and $\mathbf{E}$ are unknown. Each $\mathbf{b}_j$ corresponds to a short-time spectrum, which has a time-varying gain $g_{j,m}$. The term *component* is used to refer to one basis spectrum and its time-varying gain. Each sound source is usually modeled as a sum of one or more components. Multiple components per source are used since each pitch value of a harmonic source corresponds to a different spectrum. Also, even one note of an instrument may not have a stationary spectrum over time, so that multiple components may be needed. However, the model is flexible in the sense that it is suitable for representing both harmonic and percussive sounds. A simple

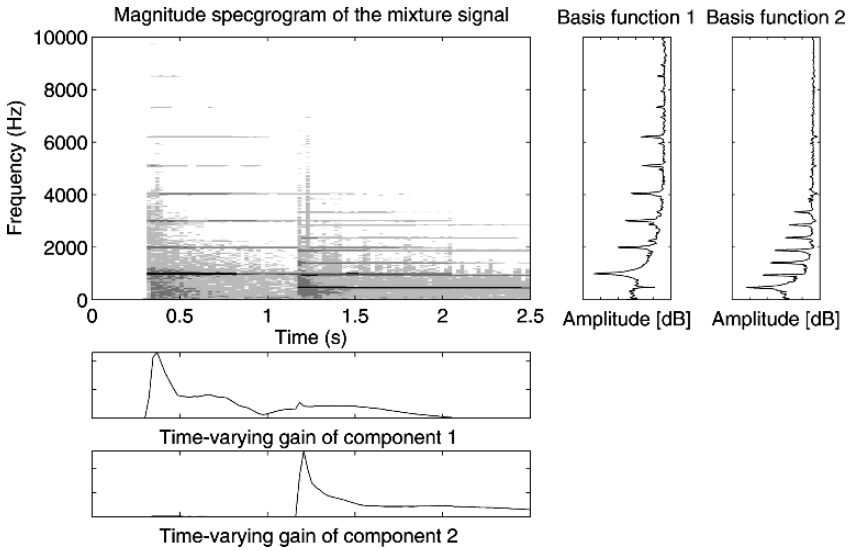


FIGURE 6 A mixture signal consisting of two piano notes (B5 and A#4, starting at times $t = 0.3$ and $t = 1.2$, respectively) and two separated components. Both components have a fixed spectrum with a time-varying gain.

two-note example is illustrated in Figure 6 where the components were estimated using the nonnegative matrix factorization algorithm described in [60] and [41].

The phase spectra of natural sound sources are very unpredictable. Therefore, the phases are often discarded and the estimation is done using the magnitude or power spectra. Linear superposition of time-domain signals does not imply linear superposition of the magnitude or power spectra. However, a linear superposition of incoherent time-domain signals can be approximated as a linear superposition of short-time power spectra.

In most systems, a discrete Fourier transform of fixed window size is used, but, in general, estimation algorithms allow the use of any time-frequency representation. For example, short-time signal processing has been used without an explicit frequency transform by Abdallah [1] and Jang and Lee [31]. It turns out that their systems learned basis functions from time-domain music and speech signals which were very similar to those used in wavelet and STFT analysis.

There are several alternative criteria for the estimation of the unknown parameters, the basic principles of which are discussed in the next three sections. In many cases, the number of basis functions, J, is unknown, and principal component analysis (PCA)

can be used to estimate it. However, in the case of monaural sound source separation, the problem of estimating the number of sources J has not received much attention yet.

Once the short-time spectrum of the input signal is separated into components, the components are further analyzed to obtain musically important information. This includes, for example, the onset and offset times and F0s of each component. Ideally, a component is active when its gain $g_{j,m}$ is nonzero, but in practice, activity detection has to be made using a threshold which is larger than zero. There are some alternative possibilities for the estimation of the F0. For example, prominant peaks can be located from the spectrum and the two-way mismatch procedure [44] can be used, or the fundamental period can be estimated from an autocorrelation estimate. Sometimes a component may represent more than one pitch. This happens especially when notes occur consistently simultaneously, as in the case of chords.

6.2 Independent Subspace Analysis

In monaural sound source separation, the term independent subspace analysis (ISA) has been used to denote the ICA of the spectrogram of a time-domain signal [8]. In ISA, either the time-varying gains or the spectra of the components are assumed to be statistically independent. There are several criteria and algorithms for obtaining the statistical independency, but they all have a common mathematical background [42]. The criteria include, for example, nongaussianity and negentropy, which are usually measured using high-order cumulants such as kurtosis [30]. As a preprocessing for ICA, the observation matrix is usually whitened and the dimensions are reduced by PCA. The core of ICA is the estimation of an unmixing matrix. Independent components are obtained by multiplying the whitened observation matrix by the estimate of the unmixing matrix. There are several ICA algorithms, some of which are freely available, for example, FastICA [29] and JADE [7].

ISA has been used in several automatic music transcription and sound source separation systems, for example, by Casey and Westner in general audio source separation [8] and by FitzGerald in percussion transcription [18].

6.3 Nonnegative Matrix Factorization

If short-time power spectra are used as observations, each component has a fixed power spectrum with a time-varying gain. It is reasonable to restrict these to be entry-wise nonnegative, so that the components are purely additive and the spectra do not have negative values. A problem with ISA is that the standard ICA algorithms do not allow non-negativity restrictions, and

in practice, the algorithms produce also negative values. This problem has been addressed, e.g., by Plumbley and Oja [54], whose nonnegative PCA algorithm solves the problem with some limitations.

Nonnegative matrix factorization (NMF) has been successfully used in unsupervised learning with the non-negativity restrictions [41]. When $\mathbf{X} \approx \mathbf{BG}$, the non-negativity of $\mathbf{B}$ and $\mathbf{G}$ alone seems to be a sufficient condition for the blind estimation of sources in many cases. In the analysis of music signals, NMF has been used, for example, by Smaragdis and Brown [60].

Lee and Seung proposed two cost functions and estimation algorithms for obtaining $\mathbf{X} \approx \mathbf{BG}$ [41]. The cost functions are the square of the Euclidean distance, or, the Frobenius norm of the error between the observation matrix $\mathbf{X}$ and the model $\mathbf{BG}$, given by

$$||\mathbf{X} - \mathbf{BG}||_F^2 = \sum_{k,m} \left([\mathbf{X}]_{k,m} - [\mathbf{BG}]_{k,m}\right)^2 \tag{12}$$

and divergence D, defined as

$$D(\mathbf{X}||\mathbf{BG}) = \sum_{k,m} d\left([\mathbf{X}]_{k,m}, [\mathbf{BG}]_{k,m}\right), \tag{13}$$

where the function d is given by

$$d(p,q) = p \log\left(\frac{p}{q}\right) - p + q. \tag{14}$$

Both cost functions are lower bounded by zero, which is obtained only when $\mathbf{X} = \mathbf{BG}$. The estimation algorithms presented in [41] initialize $\mathbf{B}$ and $\mathbf{G}$ with random values, and then update them iteratively, so that the value of the cost funtion is nonincreasing at each update.

Since the NMF aims only at representing the observed spectrogram with nonnegative components, it does not guarantee the separation of sources. Especially when the sources are always present simultaneously, the algorithm tends to represent them with a single component. However, it has turned out that the factorization of a magnitude spectrogram by minimizing the divergence given by Eq. (13) produces better separation results than ISA in most cases [69].

6.4 Sparse Coding

A technique called sparse coding has been successfully used to model the early stages of vision [51]. The term sparse refers to a signal model in which the data are represented in terms of a small number of active elements chosen out of a larger set.

The sparseness restriction is usually placed for the gains $[\mathbf{G}]_{j,m}$ in Eq. (11). Sparseness of $\mathbf{G}$ means that the probability of an element of $\mathbf{G}$ being zero is high, so that only a few components are active at a time, and each component is active only in a small number of frames. For musical sources, this is usually a valid assumption, since each component typically corresponds to a single pitch value or to a percussive source.

A sparse representation is obtained by minimizing a cost function which is the sum of a reconstruction error term and a term which incurs a penalty on the nonzero elements of $\mathbf{G}$. An example of such a cost function c is given by

$$c(\mathbf{B}, \mathbf{G}) = ||\mathbf{X} - \mathbf{BG}||_F^2 + \lambda \sum_{j,m} f([\mathbf{G}]_{j,m}). \tag{15}$$

The function f is used to penalize nonzero entries of $\mathbf{G}$, and the scalar $\lambda \geq 0$ is used to balance the reconstruction error cost and the sparseness cost. For example, Olshausen and Field [51] used $f(x) = \log(1 + x^2)$, and Hoyer [28] used $f(x) = |x|$. This approach requires that the scale of either $\mathbf{B}$ or $\mathbf{G}$ is fixed, for example to unity variance.

As in NMF, the parameters are usually solved using iterative algorithms. Hoyer [28] proposed a nonnegative sparse coding algorithm by combining NMF and sparse coding. For musical signal analysis, sparse coding has been used for example by Abdallah and Plumbley [1,2] to analyze pitched sounds and by Virtanen [67] to transcribe drums from synthesized MIDI signals.

It is not clear whether the explicit assumption of sparseness ($\lambda > 0$) really increases the quality of the separation. In [69] the sparseness criterion did not improve the average SNR of the separated sources, although it may be useful for musical transcription.

6.5 Discussion

Manual music transcription requires a lot of prior knowledge and training. It is not known whether automatic transcription of music is possible without prior information of the sources, for example the knowledge that they are harmonic. In some simple cases, it is possible to estimate the components without prior information, but this may not always be the case. Some attempts to utilize prior information have been made. Usually they are based on supervised learning since it is difficult to constrain harmonic basis functions in the model given by Eq. (10). Vincent and Rodet proposed a polyphonic transcription system based on ISA and hidden Markov models which were trained using monophonic

material [66]. FitzGerald [18] used spectral templates of instruments as an initialization of his prior subspace analysis based drum transcription system. The time-varying gains obtained for the components were further processed using ICA. All these systems use a limited instrument set. In general, all the possible instruments cannot be trained in advance but some kind of model adaptation is needed.

The linear model (10) is not well suited for separating singing voice signals, since different phonemes have different spectra. Also, other instruments with a strongly varying spectral shape are problematic. Researchers have tried to overcome the limitations of the model by using a more complex model which includes a two-dimensional time–frequency basis function for each component instead of a static spectrum. Initial experiments with this kind of approach have been presented by Smaragdis [59] and Virtanen [68]. The model can also be used to represent time-varying fundamental frequencies [19].

The dynamic range of music signals is wide, and low-intensity observations may be perceptually important. The power spectrum domain is problematic in the sense that it causes separation algorithms to concentrate on high-intensity observations, thus failing to separate low-energy sources. This has been addressed by Vincent and Rodet [66], who used a specific algorithm in the log-power spectral domain, and by Virtanen [68], who used perceptually motivated weights to mimic human loudness perception.

Automatic music transcription and sound source separation using unsupervised learning techniques is currently an active research topic, and none of the proposed methods is clearly better than others. The existing algorithms are most successful in cases where instrument sets and polyphony are limited. For example, FitzGerald [18] reported good results for an algorithm extended from ISA, and Paulus and Virtanen [52] successfully used NMF for the transcription of drum patterns consisting of bass, snare, and hi-hat drums.

7 • Summary and Conclusions

As the preceding sections have shown, music transcription can be performed using very different kinds of methods and assumptions. The underlying assumptions made for unsupervised learning methods are completely different than those made for signal-model-based statistical inference methods, yet both approaches can yield meaningful results. At the present time, none of the described main approaches stands out as clearly

the most promising. Instead, one of the decisive factors determining the popularity of different methods is their conceptual simplicity vs. performance. A problem with many transcription systems is that they are very complex entities. However, there are methods in all the main categories that are quite straightforward to implement and lead to good analysis results. Among these are, for example, the auditorily oriented method of Tolonen and Karjalainen [65], the sinusoidal-modeling method of Maher and Beauchamp [44], the expectation-maximization algorithm of Goto [23] for detecting the most predominant F0 in polyphonic signals, and the non-negative matrix factorization algorithm of Smaragdis and Brown [60].

Research is being carried out to combine the advantages of different approaches, and in many cases, this has produced the most successful results. Also, it should be remembered that this chapter has focused primarily on acoustic signal analysis methods without addressing the use of musicological information or the larger scale structure of music signals.

References

1. S. A. Abdallah. *Towards Music Perception by Redundancy Reduction and Unsupervised Learning in Probabilistic Models*. PhD thesis, Dept. Electronic Eng., King's College London, 2002.
2. S. A. Abdallah and M. D. Plumbley. An independent component analysis approach to automatic music transcription. *Audio Eng. Soc. 114th Convention*, Preprint No. 5754, Amsterdam, Netherlands, March 2003.
3. M. Abe and J. O. Smith. Design criteria for simple sinusoidal parameter estimation based on quadratic interpolation of FFT magnitude peaks. *Audio Eng. Soc. 117th Convention*, Preprint No. 6256, San Francisco, CA, 2004.
4. X. Amatriain, J. Bonada, A. Loscos, and X. Serra. Spectral processing. In U. Zölzer, editor, *DAFX – Digital Audio Effects*. Wiley New York, 2002.
5. A. S. Bregman. *Auditory Scene Analysis*. MIT Press, Cambridge, MA, 1990.
6. J. C. Brown. Musical fundamental frequency tracking using a pattern recognition method. *J. Acoust. Soc. Am.*, 92(3):1394–1402, 1992.
7. J.-F. Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11(1), 157–192, 1999.
8. M. A. Casey and A. Westner. Separation of mixed audio sources by independent subspace analysis. *Proc. 2000 Int. Computer Music Conf.*, Berlin, Germany, 2000, pp. 154–161.
9. A. T. Cemgil, B. Kappen, and D. Barber. A generative model for music transcription. *IEEE Trans on Speech and Audio Processing*, 14(2):679–694, 2006.

10. M. Davy and S. Godsill. Bayesian harmonic models for musical signal analysis. *Proc. Seventh Valencia Int. Meeting on Bayesian Statistics 7*, pp. 105–124, Tenerife, Spain, June 2002.
11. A. de Cheveigné. Pitch perception models. In C. J. Plack, A. J. Oxenham, R. R. Fay, and A. N. Popper, editors, *Pitch*. Springer, New York, 2005, pp. 169–233.
12. A. de Cheveigné and H. Kawahara. Multiple period estimation and pitch perception model. *Speech Communication*, 27:175–185, 1999.
13. A. de Cheveigné and H. Kawahara. Comparative evaluation of F0 estimation algorithms. *Proc. 7th European Conf. Speech Communication and Technology*, pp. 2451–2454, Aalborg, Denmark, 2001.
14. A. de Cheveigné and H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *J. Acoust Soc. of Am.*, 111(4): 1917–1930, 2002.
15. Ph. Depalle and T. Hélie. Extraction of spectral peak parameters using a short-time fourier transform modeling and no sidelobe windows. *Proc. 1997 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 19–22, New Palz, NY, 1997.
16. B. Doval and X. Rodet. Estimation of fundamental frequency of musical sound signals. *Proc. 1991 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3657–3660, Toronto, Canada, 1991.
17. D. P. W. Ellis. *Prediction-Driven Computational Auditory Scene Analysis*. PhD thesis, Massachusetts Institute of Technology, 1996.
18. D. FitzGerald. *Automatic Drum Transcription and Source Separation*. PhD thesis, Dublin Institute of Technology, 2004.
19. D. FitzGerald, M. Cranitch, and E. Coyle. Generalised prior subspace analysis for polyphonic pitch transcription. *Proc. Int. Conf. Digital Audio Effects*, pp. 77–81, Madrid, Spain, 2005.
20. N. H. Fletcher and T. D. Rossing. *The Physics of Musical Instruments*. Springer, Berlin, Germany, 2nd edition, 1998.
21. D. Godsmark and G. J. Brown. A blackboard architecture for computational auditory scene analysis. *Speech Communication*, 27(3): 351–366, 1999.
22. M. Goto. A robust predominant-F0 estimation method for real-time detection of melody and bass lines in cd recordings. *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 757–760, Istanbul, Turkey, June 2000.
23. M. Goto. A real-time music scene description system: predominant-F0 estimation for detecting melody and bass lines in real-world audio signals. *Speech Communication*, 43(4):311–329, 2004.
24. R. Gribonval and E. Bacry. Harmonic decomposition of audio signals with matching pursuit. *IEEE Trans. Signal Processing*, 51(1):101–111, 2003.
25. W. M. Hartmann. *Signals, Sound, and Sensation*. Springer, New York, 1998.
26. W. J. Hess. Pitch and voicing determination. In S. Furui and M. M. Sondhi, editors, *Advances in Speech Signal Processing*, pp. 3–48. Marcel Dekker, New York, 1991.

27. M. J. Hewitt and R. Meddis. An evaluation of eight computer models of mammalian inner hair-cell function. *J. Acoust Soc. Am.*, 90(2):904–917, 1991.
28. P. Hoyer. Non-negative sparse coding. In *IEEE Workshop on Networks for Signal Processing XII*, Martigny, Switzerland, 2002.
29. A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Networks*, 10(3):626–634, 1999.
30. A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley, New York, 2001.
31. G.-J. Jang and T.-W. Lee. A maximum likelihood approach to single channel source separation. *J. Machine Learning Research*, 23: 1365–1392, 2003.
32. H. Kameoka, T. Nishimoto, and S. Sagayama. Separation of harmonic structures based on tied gaussian mixture model and information criterion for concurrent sounds. In *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, (ICASSP 2004), Montreal, 2004.
33. K. Kashino, K. Nakadai, T. Kinoshita, and H. Tanaka. Organisation of hierarchical perceptual sounds: music scene analysis with autonomous processing modules and a quantitative information integration mechanism. In *Int. Joint Conf. Artificial Intelligence*, pp. 158–164, Montreal, 1995.
34. K. Kashino and H. Tanaka. A sound source separation system with the ability of automatic tone modeling. *Proc. 1993 Int. Computer Music Conf.*, pp. 248–255, Hong Kong, China, 1993.
35. M. Kay. *Modern Spectral Estimation*. Prentice Hall, Englewood Cliffs, NJ, 1988.
36. S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, Englewood Cliffs, NJ 1993.
37. A. Klapuri and M. Davy, editors. *Signal Processing Methods for Music Transcription*. Springer, New York, 2006.
38. A. P. Klapuri. Multiple fundamental frequency estimation based on harmonicity and spectral smoothness. *IEEE Trans. Speech and Audio Processing*, 11(6):804–815, 2003.
39. A. Klapuri. Multipitch analysis of polyphonic music and speech signals using an auditory model. *IEEE Trans. Audio, Speech and Language Processing*, 16(2):255–266, 2008.
40. M. Lahat, R. Niederjohn, and D. A. Krubsack. Spectral autocorrelation method for measurement of the fundamental frequency of noise-corrupted speech. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 6:741–750, June 1987.
41. D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Neural Information Processing Systems*, pp. 556–562, Denver, CO, 2001.
42. T.-W. Lee, M. Girolami, A. J. Bell, and T. J. Sejnowski. A unifying information-theoretic framework for independent component analysis. *Computers and Mathematics with Applications*, 31(11): 1–21, 2000.

43. R. C. Maher. Evaluation of a method for separating digitized duet signals. *J. Audio Eng. Soc.*, 38(12):956–979, 1990.
44. R. C. Maher and J. W. Beauchamp. Fundamental frequency estimation of musical signals using a two-way mismatch procedure. *J. Acoust. Soc. Am.*, 95(4):2254–2263, 1994.
45. M. Marolt. SONIC: transcription of polyphonic piano music with neural networks. *Proc. MOSART Workshop on Current Research Directions in Computer Music*, pp. 217–224, Barcelona, Spain, November 2001.
46. K. D. Martin. Automatic transcription of simple polyphonic music: robust front end processing. Technical Report 399, MIT Media Laboratory, Perceptual Computing Section, 1996.
47. R. J. McAulay and T. F. Quatieri. Speech analysis/synthesis based on a sinusoidal representation. *IEEE Trans. on Speech and Audio Processing*, 34(4):744–754, 1986.
48. R. Meddis and M. J. Hewitt. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: pitch identification. *J. Acousti. Soc. Am.*, 89(6):2866–2882, 1991.
49. B. C. J. Moore, editor. *Hearing – Handbook of Perception and Cognition*. Academic Press, San Diego, CA, 2nd edition, 1995.
50. J. A. Moorer. *On the Segmentation and Analysis of Continuous Musical Sound by Digital Computer*. PhD thesis, Department of Music, Stanford University, 1975. Distributed as Department of Music report No. STAN-M-3.
51. B. A. Olshausen and D. F. Field. Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Research*, 37:3311–3325, 1997.
52. J. Paulus and T. Virtanen. Drum transcription with non-negative spectrogram factorisation. *European Signal Processing Conf.*, Antalya, Turkey, 2005.
53. G. Peterschmitt, E. Gómez, and P. Herrera. Pitch-based solo location. *Proc. MOSART Workshop on Current Research Directions in Computer Music*, Barcelona, Spain, 2001.
54. M. D. Plumbley and E. Oja. A "non-negative PCA" algorithm for independent component analysis. *IEEE Trans. on Neural Networks*, 15(1):66–67, 2004.
55. T. F. Quatieri and R. G. Danisewicz. An approach to co-channel talker interference suppression using a sinusoidal model for speech. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38(1):56–69, 1990.
56. L. R. Rabiner, M. J. Cheng, A. E. Rosenberg, and C. A. McGonegal. A comparative performance study of several pitch detection algorithms. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 24(5):399–418, 1976.
57. X. Rodet. Musical sound signal analysis/synthesis: Sinusoidal+ residual and elementary waveform models. *Proc. IEEE Time-Frequency and Time-Scale Workshop*, Coventry, UK, 1997.
58. X. Serra. Musical sound modeling with sinusoids plus noise. In C. Roads, S. Pope, A. Picialli, and G. De Poli, editors, *Musical Signal Processing*. pp. 91–122, Swets & Zeitlinger, 1997.

59. P. Smaragdis. Discovering auditory objects through non-negativity constraints. *Proc. ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing*, Jeju, Korea, 2004.
60. P. Smaragdis and J. C. Brown. Non-negative matrix factorization for polyphonic music transcription. *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 177–180, New Palz, NY, 2003.
61. J. O. Smith and X. Serra. Parshl: an analysis/synthesis program for Proc. non-harmonic sounds based on a sinusoidal representation. *Int. Computer Music Conf.*, pp. 290–297, Urbana, IL, 1987.
62. A. D. Sterian. *Model-Based Segmentation of Time–Frequency Images for Musical transcription*. PhD thesis, University of Michigan, 1999.
63. P. Stoica and R. L. Moses. *Introduction to Spectral Analysis*. Prentice Hall, Englwood Cliffs, NJ, 1997.
64. D. Talkin. A robust algorithm for pitch tracking. In W. B. Kleijn and K. K. Paliwal, editors, *Speech Coding and Synthesis*, pp. 495–517. Elsevier, Amsterdam, 1995.
65. T. Tolonen and M. Karjalainen. A computationally efficient multipitch analysis model. *IEEE Trans. on Speech and Audio Processing*, 8(6):708–716, 2000.
66. E. Vincent and X. Rodet. Music transcription with ISA and HMM. *Proc. 5th Int. Symp. on Independent Component Analysis and Blind Signal Separation*, pp. 1119–1204, London, U.K., 2004.
67. T. Virtanen. Sound source separation using sparse coding with temporal continuity objective. *Proc. 2003 Int. Computer Music Conference*, pp. 231–234, Singapore, 2003.
68. T. Virtanen. Separation of sound sources by convolutive sparse coding. *Proc. ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing*, Jeju, Korea, 2004.
69. T. Virtanen, Monaural Sound Source Separation by Non-Negative Matrix Factorization with Temporal Continuity and Sparseness Criteria. *IEEE Trans. Audio, Speech, and Language Processing*, 15(3):1066–1074, 2007
70. T. Virtanen and A. Klapuri. Separation of harmonic sounds using multipitch analysis and iterative parameter estimation. *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 83–86, New Paltz, NY, 2001.
71. T. Virtanen and A. Klapuri. Separation of harmonic sounds using linear models for the overtone series. *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 1747–1760, Orlando, FL, 2002.
72. P. J. Walmsley. *Signal Separation of Musical Instruments. Simulation-Based Methods for Musical Signal Decomposition and Transcription*. PhD thesis, Department of Engineering, University of Cambridge, September 2000.

21

Music Structure Analysis from Acoustic Signals

Roger B. Dannenberg[1] *and Masataka Goto*[2]

[1] *Carhegie Mellon University, Pittsburgh, PA, USA*
[2] *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan*

1 • Introduction

Music is full of structure, including sections, sequences of distinct musical textures, and the repetition of phrases or entire sections. The analysis of music audio relies upon feature vectors that convey information about music texture or pitch content. Texture generally refers to the average spectral shape and statistical fluctuation, often reflecting the set of sounding instruments, e.g., strings, vocal, or drums. Pitch content reflects melody and harmony, which is often independent of texture. Structure is found in several ways. Segment boundaries can be detected by observing marked changes in locally averaged texture. Similar sections of music can be detected by clustering segments with similar average textures. The repetition of a sequence of music often marks a logical segment. Repeated phrases and hierarchical structures can be discovered by finding similar sequences of feature vectors within a piece of music. Structure analysis can be used to construct music summaries and to assist music browsing.

Probably everyone would agree that music has structure, but most of the interesting musical information that we perceive lies

hidden below the complex surface of the audio signal. From this signal, human listeners perceive vocal and instrumental lines, orchestration, rhythm, harmony, bass lines, and other features. Unfortunately, music audio signals have resisted our attempts to extract this kind of information. Researchers are making progress, but so far, computers have not come near to human levels of performance in detecting notes, processing rhythms, or identifying instruments in a typical (polyphonic) music audio texture.

On a longer time scale, listeners can hear structure including the chorus and verse in songs, sections in other types of music, repetition, and other patterns. One might think that without the reliable detection and identification of short-term features such as notes and their sources, it would be impossible to deduce any information whatsoever about even higher levels of abstraction. Surprisingly, it *is* possible to automatically detect a great deal of information concerning music structure. For example, it is possible to label the structure of a song as AABA, meaning that opening material (the "A" part) is repeated once, then contrasting material (the "B" part) is played, and then the opening material is played again at the end. This structural description may be deduced from low-level audio signals. Consequently, a computer might locate the "chorus" of a song without having any representation of the melody or rhythm that characterizes the chorus.

Underlying almost all work in this area is the concept that structure is induced by the repetition of similar material. This is in contrast to, say, speech recognition, where there is a common understanding of words, their structure, and their meaning. A string of unique words can be understood using prior knowledge of the language. Music, however, has no language or dictionary (although there are certainly known forms and conventions). In general, structure can only arise in music through repetition or systematic transformations of some kind.

Repetition implies there is some notion of similarity. Similarity can exist between two points in time (or at least two very short time intervals), similarity can exist between two sequences over longer time intervals, and similarity can exist between the longer-term statistical behaviors of acoustical features. Different approaches to similarity will be described.

Similarity can be used to segment music: contiguous regions of similar music can be grouped together into segments. Segments can then be grouped into clusters. The segmentation of a musical work and the grouping of these segments into clusters is a form of analysis or "explanation" of the music.

2 • Features and Similarity Measures

A variety of approaches are used to measure similarity, but it should be clear that a direct comparison of the waveform data or individual samples will not be useful. Large differences in waveforms can be imperceptible, so we need to derive features of waveform data that are more perceptually meaningful and compare these features with an appropriate measure of similarity.

2.1 Feature Vectors for Spectrum, Texture, and Pitch

Different features emphasize different aspects of the music. For example, Mel-frequency cepstral coefficients (MFCCs) seem to work well when the general shape of the spectrum but not necessarily pitch information is important. MFCCs generally capture overall "texture" or timbral information (what instruments are playing in what general pitch range), but some pitch information is captured, and results depend upon the number of coefficients used as well as the underlying musical signal.

When pitch is important, e.g., when searching for similar harmonic sequences, the chromagram is effective. The chromagram is based on the idea that tones separated by octaves have the same perceived value of *chroma* (Shepard, 1964). Just as we can describe the *chroma* aspect of pitch, the short-term frequency spectrum can be restructured into the *chroma spectrum* by combining energy at different octaves into just one octave. The *chroma vector* is a discretized version of the *chroma spectrum* where energy is summed into 12 log-spaced divisions of the octave corresponding to pitch classes (C, C#, D, …B). By analogy to the spectrogram, the *discrete chromagram* is a sequence of chroma vectors.

It should be noted that there are several variations of the chromagram. The computation typically begins with a short-term Fourier transform (STFT) which is used to compute the magnitude spectrum. There are different ways to "project" this onto the 12-element chroma vector. Each STFT bin can be mapped directly to the most appropriate chroma vector element (Bartsch and Wakefield, 2001), or the STFT bin data can be interpolated or windowed to divide the bin value among two neighboring vector elements (Goto, 2003a). Log magnitude values can be used to emphasize the presence of low-energy harmonics. Values can also be averaged, summed, or the vector can be computed to conserve the total energy. The chromagram can also be computed by using the wavelet transform.

Regardless of the exact details, the primary attraction of the chroma vector is that, by ignoring octaves, the vector is relatively

insensitive to overall spectral energy distribution and thus to timbral variations. However, since fundamental frequencies and lower harmonics of tones feature prominently in the calculation of the chroma vector, it is quite sensitive to pitch class content, making it ideal for the detection of similar harmonic sequences in music.

While MFCCs and chroma vectors can be calculated from a single short-term Fourier transform, features can also be obtained from longer sequences of spectral frames. Tzanetakis and Cook (1999) use means and variances of a variety of features in a 1-s window. The features include the spectral centroid, spectral rolloff, spectral flux, and RMS energy.

Peeters et al. (2002) describe "dynamic" features, which model the variation of the short-term spectrum over windows of about 1 s. In this approach, the audio signal is passed through a bank of Mel filters. The time-varying magnitudes of these filter outputs are each analyzed by a short-term Fourier transform. The resulting set of features, the Fourier coefficients from each Mel filter output, is large, so a supervised learning scheme is used to find features that maximize the mutual information between feature values and hand-labeled music structures.

2.2 Measures of Similarity

Given a feature vector such as the MFCC or chroma vector, some measure of similarity is needed. One possibility is to compute the (dis)similarity using the Euclidean distance between feature vectors. Euclidean distance will be dependent upon feature magnitude, which is often a measure of the overall music signal energy. To avoid giving more weight to the louder moments of music, feature vectors can be normalized, for example, to a mean of zero and a standard deviation of one or to a maximum element of one.

Alternatively, similarity can be measured using the scalar (dot) product of the feature vectors. This measure will be larger when feature vectors have a similar direction. As with Euclidean distance, the scalar product will also vary as a function of the overall magnitude of the feature vectors. If the dot product is normalized by the feature vector magnitudes, the result is equal to the cosine of the angle between the vectors. If the feature vectors are first normalized to have a mean of zero, the cosine angle is equivalent to the correlation, another measure that has been used with success.

Lu et al. (2004) use a constant-Q transform (CQT), and found that CQT outperforms chroma and MFCC features using a cosine distance measure. They also introduce a "structure-based" distance measure that takes into account the harmonic structure of spectra

to emphasize pitch similarity over timbral similarity, resulting in additional improvement in a music structure analysis task.

Similarity can be calculated between individual feature vectors, as suggested above, but can also be computed over a window of feature vectors. The measure suggested by Foote (1999) is vector correlation:

$$S_w(i,j) = \frac{1}{w} \sum_{k=0}^{w-1} (V_{i+k} \bullet V_{j+k}), \quad (1)$$

where w is the window size. This measure is appropriate when feature vectors vary with time, forming significant temporal patterns. In some of the work that will be described below, the detection of temporal patterns is viewed as a processing step that takes place after the determination of similarity.

2.3 Evaluation of Features and Similarity Measures

Linear prediction coefficients (LPC) offer another low-dimensional approximation to spectral shape, and other encodings such as moments (centroid, standard deviation, skewness, etc.) are possible. Aucouturier and Sandler (2001) compare various approaches and representations. Their ultimate goal is to segment music according to texture, which they define as the combination of instruments that are playing together. This requires sensitivity to the general spectral shape, and insensitivity to the spectral details that vary according to pitch. They conclude that a vector of about 10 MFCCs is superior to LPC and discrete cepstrum coefficients (Galas and Rodet, 1990).

On the other hand, Hu et al. (2003) compare features for detecting similarity between acoustic and synthesized realizations of a single work of music. In this case, the goal is to ignore timbral differences between acoustic and synthetic instruments, but to achieve fine discrimination of pitches and harmonies. They conclude that the chroma vector is superior to pitch histograms and MFCCs.

3 • Segmentation

One approach to discovering structure in music is to locate segments of similar musical material and the boundaries between them. Segmentation does not rely on classification or the discovery of higher order structure in music. However, one can envision using segmentation as a starting point for a number of more complicated tasks, including music summarization, music analysis, music search, and genre classification. Segmentation

can also assist in audio browsing, a task that can be enhanced through some sort of visual summary of music and audio segments.

3.1 Segmentation Using Texture Change

Tzanetakis and Cook (1999) perform segmentation as follows: Feature vectors V_i are computed as described above. A feature time differential, Δ_i, is defined as the Mahalanobis distance:

$$\Delta_i = (V_i - V_{i-1})^{\mathrm{T}} \Sigma^{-1} (V_i - V_{i-1}), \tag{2}$$

where Σ is an estimate of the feature covariance matrix, calculated from the training data, and i is the frame number (time). This measure is related to the Euclidean distance but takes into account the variance and correlations among features. Next, the first order differences of the distance, $\Delta_i - \Delta_{i-1}$, are computed. A large difference indicates a sudden transition. Peaks are picked, beginning with the maximum. After a peak is selected, the peak and its neighborhood are zeroed to avoid picking another peak within the same neighborhood. Assuming the total number of segments is given a priori, the neighborhood is 20% of the average segment size. Additional peaks are selected and zeroed until the desired number of peaks (segment boundaries) has been obtained.

3.2 Segmentation by Clustering

Logan and Chu (2000) describe a clustering technique for discovering music structure. The goal is to label each frame of audio so that frames within similar sections of music will have the same labels. For example, all frames within all occurrences of the chorus should have the same label. This can be accomplished using bottom-up clustering to merge clusters that are similar. Initially, the feature vectors are divided into fixed-length contiguous segments and each segment receives a different label. The following clustering step is iterated:

> Calculate the mean μ and covariance Σ of the feature vectors within each cluster. Compute a modified Kullback–Leibler (KL) distance between each pair of clusters, as described below. Find the pair of clusters with the minimum KL2 distance, and if this distance is below a threshold, combine the clusters. Repeat this step until no distance is below the threshold.

The KL2 distance between two Gaussian distributions A and B is given by

$$\mathrm{KL2}(A, B) = \mathrm{KL}(A; B) + \mathrm{KL}(B; A) \tag{3}$$

$$= \frac{\Sigma_A}{\Sigma_B} + \frac{\Sigma_B}{\Sigma_A} + (\mu_A - \mu_B) \cdot \left(\frac{1}{\Sigma_A} + \frac{1}{\Sigma_B} \right). \tag{4}$$

3.3 Segmentation and Hidden Markov Models

Another approach to segmentation uses a hidden Markov model (HMM). In this approach, segments of music correspond to discrete states Q and segment transitions correspond to state changes. Time advances in discrete steps corresponding to feature vectors and transitions from one state to the next are modeled by a probability distribution that depends only on the current state. This forms a Markov model that generates a sequence of states. Note that states are "hidden" because only feature vectors are observable. Another probability distribution, $p(V_i|q_i)$, models the generation of feature vector V_i from state q_i. The left side of Figure 1 illustrates a four-state ergodic Markov model, where arrows represent state transition probabilities. The right side of the figure illustrates the observation generation process, where arrows denote conditional probabilities between variables.

The HMM has advantages for segmentation. In general, feature vectors do not indicate the current state (segment class) unambiguously, so when a single feature vector is observed, one cannot assume that it was generated by particular state. However, some features are more likely to occur in one state than another, so one can observe the *trend* of feature vectors, ignoring the unlikely outliers and guessing the state that is most consistent with the observations. If transitions are very unlikely, one may have to assume many outliers occur. On the other hand, if transitions are common and segments are short, one can change states rapidly to account for different feature vectors. The HMM formalism can determine the segmentation (the hidden state sequence) with the maximum likelihood given a set of transition probabilities and observations, thus the model can formalize the trade-offs between minimizing transitions and matching features to states. Furthermore, HMM transition probabilities can be estimated from unlabeled training data, eliminating the need to guess transition probabilities manually.

Aucouturier and Sandler (2001) model the observation probability distribution $P(V_i|q_j)$ as a mixture of Gaussian distributions over the feature space:

$$P(V_i|q_j) = \sum_{m=1}^{M} c_{j,m} \cdot \mathbf{N}(V_i, \mu_{j,m}, \Gamma_{j,m}), \tag{5}$$

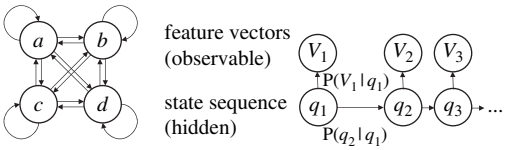


FIGURE 1 Hidden Markov model with four hidden states a, b, c, and d. As shown, feature vectors depend only upon the current state, which depends only upon the previous state.

where **N** is a Gaussian probability density function with mean $\mu_{i,m}$, covariance matrix $\Gamma_{j,m}$, and $c_{j,m}$ is a mixture coefficient. Here, i indexes time and j indexes state. They train the HMM using the Baum–Welch algorithm using the sequence vectors from the single song chosen for analysis. The Viterbi algorithm is then used to find the sequence of hidden states with the maximum likelihood, given the observed feature vectors.

One potential drawback of this approach is that the HMM will segment the signal according to fine-grain changes in spectral content rather than long-term elements of musical form. For example, in one of Aucoturier and Sandler's test cases (see Figure 2), the HMM segmentation appears to isolate individual words of a singer rather than divide the song according to verses and instrumental interludes (Aucouturier et al., 2005). In other words, the segments can be quite short when there are rapid changes in the music. Although this might be the desired result, it seems likely that one could detect longer-term, higher-level music structure by averaging features over a longer time span or applying further processing to the state sequence obtained from an HMM.

Peeters et al. (2002) approach the problem of clustering with a two-pass algorithm. Imagine a human listener hearing a piece of music for the first time. The range of variation of music features becomes apparent, and templates or classes of music are formed. In the second hearing, the structure of the music can be identified in terms of the previously identified templates.

An automated system is inspired by this two-pass model. In the first pass, texture change indicates segment boundaries, and "potential" states are formed from the mean values of feature vectors within segments. In the second pass, potential states that are highly similar are merged by using the K-means algorithm. The resulting K states are called the "middle" states. Because they represent clusters with no regard for temporal contiguity, a hidden Markov model initialized with these "middle" states is then used to inhibit rapid inappropriate state transitions by penalizing them. The Baum–Welch algorithm is used to train the model on the sequence of feature vectors from the song. Viterbi decoding is used to obtain a state sequence. Figure 3 shows the result of an analysis using this smoothing technique.

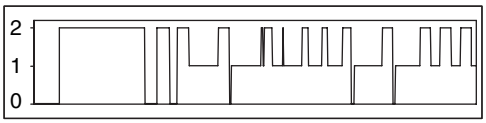


FIGURE 2 Segmentation of 20 s of a song. State 0 is silence, State 1 is voice, accordion, and accompaniment, and State 2 is accordion and accompaniment.

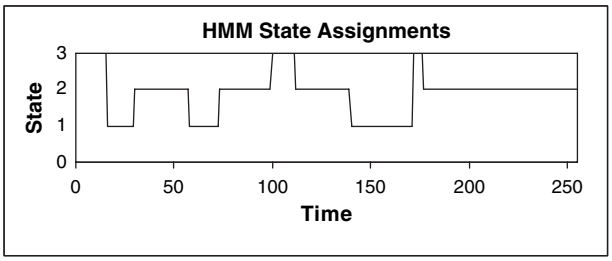


FIGURE 3 Classification of states in "Head Over Feet" from artist Alanis Morisette. (Adapted from Peeters et al. 2002).

4 • The Similarity Matrix

A concept used by many researchers is the similarity matrix. Given a sequence of feature vectors V_i and a measure of similarity $S(i, j)$, one can simply view $S(i, j)$ as a matrix. The matrix can be visualized using a grayscale image where black represents dissimilar vectors and white represents similar vectors. Shades of gray represent intermediate values. Since any vector is similar to itself, the diagonal of the similarity matrix will be white. Also, assuming the similarity measure is symmetric, the matrix will be symmetric about the diagonal. The interesting information in the matrix is in the patterns formed off the diagonal.

In very general terms, there are two interesting sorts of patterns that appear in the similarity matrix, depending on the nature of the features. The first of these appears when features correspond to relatively long-term textures. The second appears when features correspond to detailed short-term features such as pitch or harmony and where similar sequences of features can be observed. These two types of patterns are considered in the next two sections.

4.1 Texture Patterns

First, consider the case where features represent the general texture of the music, for example whether the music is primarily vocal, drum solo, or guitar solo. Figure 4 shows an idealized similarity matrix for this case. The white diagonal appears because feature vectors along the diagonal are identical. Notice that wherever there are similar textures, the matrix is lighter in color (more similar), so for example, all of the feature vectors for the vocals (V) are similar to one another, resulting in large light-colored square patterns both on and off the diagonal. Where two feature vectors correspond to different textures, for example drums and vocals, the matrix is dark.

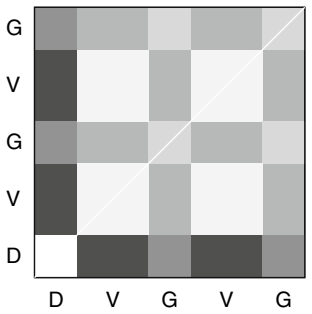


FIGURE 4 An idealized similarity matrix for segments of drum (D), vocal (V), and guitar (G) texture.

Notice that along the diagonal, a checkerboard pattern appears at segment boundaries, with darker regions to the upper left and lower right, and lighter regions to the lower left and upper right. Foote (2000) proposes the correlation of the similarity matrix S with a kernel based on this checkerboard pattern in order to detect segment boundaries. The general form of the kernel is

$$C = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}. \tag{6}$$

(Note that in Eq. (6), row numbers increase in the downward direction whereas in the similarity matrix images, the row number increases in the upward direction. Therefore the diagonal in Eq. (6) runs from upper left to lower right.) The kernel image in Figure 5 represents a larger checkerboard pattern with radial smoothing. The correlation $N(i)$ of this kernel along the diagonal



Checkerboard Kernel



Checkerboard Kernel with Radial Smoothing

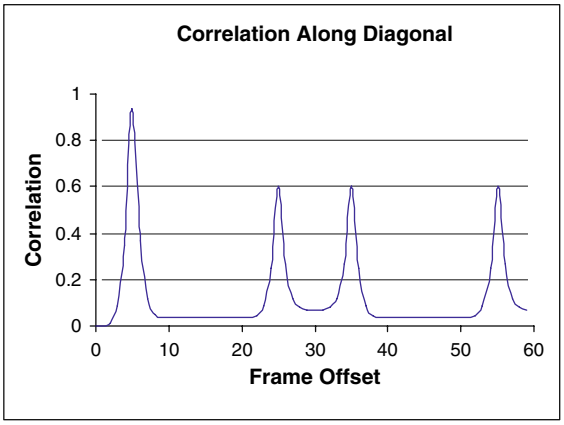


FIGURE 5 The correlation of the kernel shown at *lower left* with a similarity matrix.

of a similarity matrix Scan be considered to be a measure of novelty (Foote, 2000):

$$N(i) = \sum_{m=-L/2}^{L/2} \sum_{n=-L/2}^{L/2} C(m,n)S(i+m, i+n). \quad (7)$$

A graph of $N(i)$ for the similarity matrix in Figure 4 is shown in Figure 5. A peak occurs at each transition because transition boundaries have the highest correlation to the checkerboard pattern.

Cooper and Foote (2003) extend this technique for finding segment boundaries with a statistical method for clustering segments.

4.2 Repeating Sequence Patterns

While the texture patterns described above are most useful for detecting transitions between segments, the second kind of pattern can be used to discover repetition within a song. For these patterns to appear, it is important that features reflect short-term changes. Generally, features should vary significantly with changes in the pitch of a melody or with changes in harmony. If this condition is satisfied, then there will not be great similarity within a segment and there will not be a clear pattern of light-colored squares as seen in Figure 4. However, if a segment of music repeats with an offset of j, then $S(i,i)$ will equal $S(i,i+j)$, generating a diagonal line segment at an offset of j from the central diagonal. This is illustrated schematically in Figure 6, where it is assumed that the vocal sections (V) constitute three repetitions of very similar music, whereas the two guitar sections (G) are not so similar. Notice that each non-central diagonal line segment indicates the starting times and the duration of two similar sequences of features. Also, notice that since each pair of similar sequences is represented by two diagonal line segments, there are a total of six (6) off-central line segments in Figure 6.

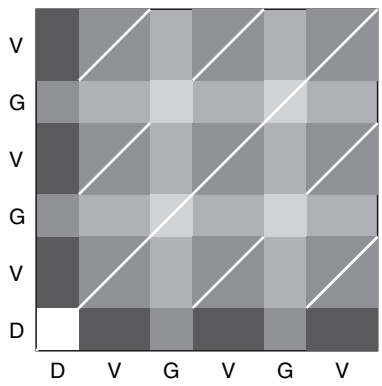


FIGURE 6 When sections of music are repeated, a pattern of diagonal line segments is generated.

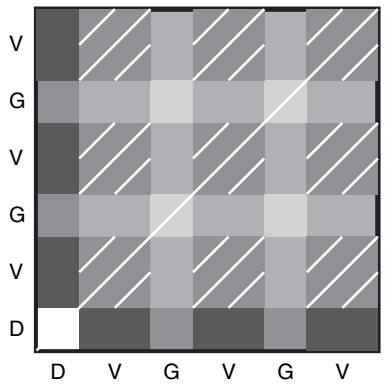


FIGURE 7 The vocal segments (V) in this similarity matrix contain a repetition, generating additional pattern that is characteristic of music structure.

Although not shown in Figure 6, the similarity matrix can also illustrate hierarchical relationships. For example, if each vocal section (V) consists of a phrase that is repeated, the similarity matrix would look like the one in Figure 7.

Figure 8 illustrates both texture patterns and repeated sequence patterns from the song "Day Tripper" by the Beatles. The bridge is displayed, starting with three repetitions of a two-measure guitar phrase in the first 11 s, followed by six measures of vocals. Notice how a checkerboard pattern appears due to the timbral self-similarity of the guitar section (0–11 s) and the vocal section (11–21 s). Finer structure is also visible. A repeated sequence pattern appears within the guitar section as parallel diagonal lines. This figure uses the power spectrum below 5.5 kHz as the feature vector, and uses the cosine of the angle between vectors as a measure of similarity.

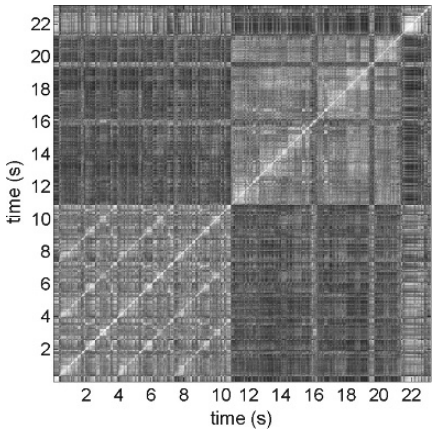


FIGURE 8 Similarity matrix using spectral features from the bridge of "Day Tripper" by the Beatles.

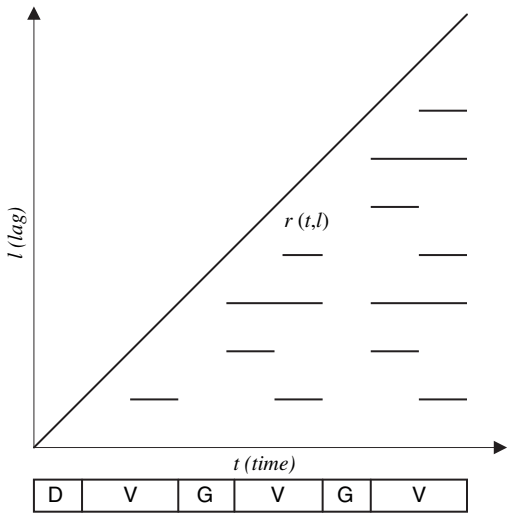


FIGURE 9 Time-lag matrix representation of the similarity matrix in Figure 7.

4.3 The Time-Lag Matrix

When the goal is to find repeating sequence patterns, it is sometimes simpler to change coordinate systems so that patterns appear as horizontal or vertical lines. The *time-lag* matrix r is defined by

$$r(t,l) = S(t, t-l), \text{where } t-l \geq 0. \tag{8}$$

Thus, if there is repetition, there will be a sequence of similar frames with a constant lag. Since lag is represented by the vertical axis, a constant lag implies a horizontal line. The time-lag version of Figure 7 is shown in Figure 9. Only the lines representing similar sequences are shown, and the grayscale has been reversed, so that similarity is indicated by black lines.

5 • Finding Repeating Sequences

Of course, with audio data obtained from real audio recordings, the similarity or time-lag matrix will be full of noise and ambiguity arising from spurious similarity between different frames. Furthermore, repetitions in music are rarely exact; variation of melodic, harmonic, and rhythmic themes is an essential characteristic of music. In order to automate the discovery of musical structure, algorithms must be developed to identify the structure that lies within the similarity matrix.

5.1 Melodic Sequence Matching

One way to find repetition is to transcribe the melody and perform matching on the resulting symbolic transcription. While extracting the melody from a polyphonic recording (Goto, 2004) is very difficult in general, an approximate transcription from an instrumental recording or from a monophonic (melody only) recording is relatively easy. Dannenberg (2002) describes a simple transcription system based on the enhanced autocorrelation algorithm (Tolonen and Karjalainen, 2000) applied to a ballad recorded by John Coltrane. The transcription results in a quantized integer pitch value p_i and real inter-onset interval d_i for each note (inter-onset intervals are typically preferred over note duration in music processing). This sequence is processed as follows:

1. First, a similarity matrix is constructed where rows and columns correspond to notes. This differs from the similarity matrix described above where rows and columns correspond to feature vectors with a fixed duration.
2. Each cell of the similarity matrix $S(i,j)$ represents the duration of similar melodic sequences starting at notes i and j. A simple "greedy" algorithm is used to match these two sequences. If note i does not match note j, $S(i,j) = 0$.
3. Simplify the matrix by removing redundant entries. If a sequence beginning at i matches one at j, then there should be another match at $i+1$ and $j+1$. To simplify the matrix, find the submatrix $S(i:u, j:v)$ where the matching sequences at i and j end at u and v. Zero every entry in the submatrix except $S(i,j)$. Also, zero all entries for matching sequences of length 1.
4. Now, any non-zero entry in S represents a pair of matching sequences. By scanning across rows of S we can locate all similar sequences. Sequences are clustered: The first non-zero element in a row represents a cluster of two sequences. Any other non-zero entry in the row that roughly matches the durations of the clustered sequences is added to the cluster. After scanning the row, all pair-wise matches are zeroed so they will not be considered again.

The result of this step is a set of clusters of similar melodic segments. Because repetitions in the music are not exact, there can be considerable overlap between clusters. It is possible for a long segment to be repeated exactly at one offset, and for a portion of that same segment to be repeated several times at other offsets. It may be desirable to simplify the analysis by labeling each note with a particular cluster. This simplification is described in the next section.

5.2 Simplification or Music Explanation

The goal of the simplification step is to produce *one* possible set of labels for notes. Ideally, the labels should offer a simple "explanation" of the music that highlights repetition within the music. The AABA structure common in songs is a typical explanation. In general, longer sequences of notes are preferable because they explain more, but when sequences are too long, interesting substructure may be lost. For example, the structure AABAAABA could also be represented as AABA repeated, i.e., the structure could be labeled AA, but most theorists would consider this to be a poor explanation. Hierarchical explanations offer a solution, but there is no formal notion as yet of the optimal simplification or explanation.

Dannenberg uses a "greedy" algorithm to produce reasonable explanations from first note to last (Dannenberg and Hu, 2002). Notes are initially unlabeled. As each unlabeled note is encountered, search the clusters from the previous section to find one that includes the unlabeled note. If a cluster is found, allocate a new label, e.g., "A", and label every note included in the cluster accordingly. Continue labeling with the next unlabeled note until all notes are processed.

Figure 10 illustrates output from this process. Notice that the program discovered a substructure within what would normally be considered the "bridge" or the B part, but this substructure is "real" in the sense that one can see it and hear it. The gap in the middle of the piece is a piano solo where transcription failed. Notice that the program correctly determines that the saxophone enters on the bridge (the B part) after the piano solo. The program also identifies the repeated two-measure phrase at the end. It fails to notice the structure of ascending pitches at the very end because, while this is a clear musical gesture, it is not based on the repetition of a note sequence.



FIGURE 10 A computer analysis of "Naima" by John Coltrane. The automatic transcription appears as a "*piano roll*" at the top, the computer analysis appears as *shaded bars*, where *similar shading* indicates similar sequences, and conventional labels appear at the bottom.

5.3 Finding Similar Sequences in the Similarity Matrix

Typically, transcription of a music signal into a sequence of notes is not possible, so similar sequences must be detected as patterns in the similarity or time-lag matrix. For example, Bartsch and Wakefield (2001) filter along diagonals of a similarity matrix to detect similarity. This assumes nearly constant tempo, but that is a good assumption for the popular music used in their study. Their objective was not to identify the beginnings and endings of repeating sequences but to find the chorus of a popular song for use as an "audio thumbnail" or summary. The thumbnail is selected as the maximum element of the filtered similarity matrix, with the additional constraints that the lag is at least one-tenth of the length of the song and the thumbnail does not appear in the last quarter of the song.

Peeters and Rodet (2003) suggest using a 2D structuring filter on the lag matrix to detect similar sequences. Their filter counts the number of values in the neighborhood to the left and right of a point that are above a threshold. To allow for slight changes in tempo, which results in lines that are not perfectly horizontal, neighbor cells above and below are also considered. Lu et al. (2004) suggest erosion and dilation operations on the lag matrix to enhance and detect significant similar sequences.

Dannenberg and Hu (2003) use a discrete algorithm to find similar sequences which is based on the idea that a path from cell to cell through the similarity matrix specifies an alignment between two subsequences of the feature vectors. If the path goes through $S(i, j)$, then vector i is aligned with j. This suggests using a dynamic time warping (DTW) algorithm (Rabiner and Juang, 1993), and the actual algorithm is related to DTW.

The goal is to find alignment paths that maximize the average similarity of the aligned features. A partial or complete path P is defined as a set of pairs of locations and is rated by the average similarity along the path:

$$q(P) = \frac{1}{|P|} \sum_{(i,j)\in P} S(i, j), \tag{9}$$

where $|P|$ is the path length using Euclidean distance. Paths are extended as long as the rating remains above a threshold. Paths are constrained to move up one cell, right one cell, or diagonally to the upper right as shown in Figure 11 (and adopting the orientation of the similarity matrix visualizations where time increases upward and to the right). Therefore, every point that is on a path can be reached from below, from the left, or from the lower left. Each cell (i, j) of an array is computed by looking at the cell below, left, and below left to find the (previously calculated) best path (highest $q(P)$) passing through those cells. Three new

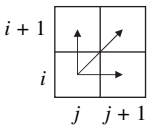


FIGURE 11 Extending a path from $S(i, j)$.

ratings of r are computed by extending each of the three paths to include (i, j). The path with the highest rating is remembered as the one passing through (i, j).

Because cells depend on previously computed values to the lower left, cells are computed along diagonals of constant $i+j$, from lower left to upper right (increasing $i+j$). When no path has a rating above some fixed threshold, the path ends. A path may begin wherever $S(i, j)$ is above threshold and no previous paths exist to be extended.

5.4 Forming Clusters

After alignment paths are found, they are grouped into clusters. So if sequence A aligns to sequence B, and sequence A also aligns to sequence C, then A, B, and C should be grouped in a single cluster. Unfortunately, it is unlikely that the alignments of A to B and A to C use exactly the same frames. It is more likely that A aligns to B and A′ aligns to C, where A and A′ are mostly overlapping. This can be handled simply by considering A to equal A′ when they start and end within some fraction of their total length, for example within 10%. Once clusters are formed, further simplification and explanation steps can be performed as described above.

5.5 Isolating Line Segments from the Time-Lag Matrix

If nearly constant tempo can be assumed, the alignment path is highly constrained and the alignment path approach may not work well. Taking advantage of the fact that similar sequences are represented by horizontal lines in the time-lag matrix, Goto (2003a) describes an alternative approach to detecting music structure. In this work, the time-lag matrix is first normalized by subtracting a local mean value while emphasizing horizontal lines. In more detail, given a point $r(t, l)$ in the time-lag matrix, six-directional local mean values along the right, left, upper, lower, upper right, and lower left directions starting from $r(t, l)$ are calculated, and the maximum and minimum are obtained. If the local mean along the right or left direction takes the maximum, $r(t, l)$ is considered a part of a horizontal line and emphasized by subtracting the minimum from $r(t, l)$. Otherwise, $r(t, l)$ is considered a noise and suppressed by subtracting the maximum from $r(t, l)$; noises tend to appear as lines along the

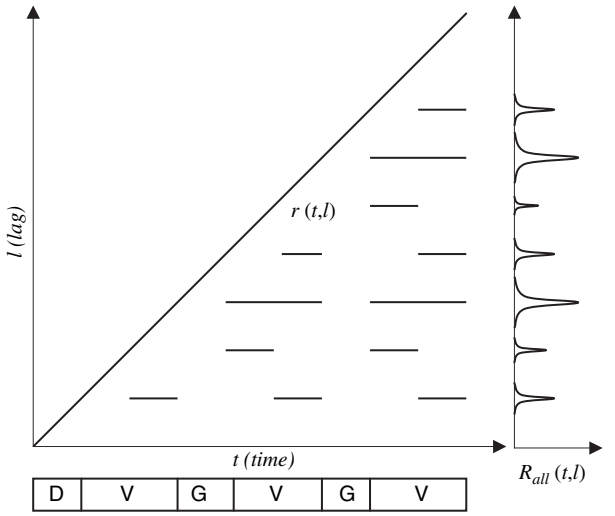


FIGURE 12 The summary $R_{\text{all}}(t, l)$ indicates the possibility that there are similar segments at a lag of l.

upper, lower, upper right, and lower left directions. Then, a summary is constructed by integrating over time:

$$R_{\text{all}}(t, l) = \int_l^t \frac{r(\tau, l)}{t - l} \mathrm{d}\tau. \tag{10}$$

R_{all} is then smoothed by a moving average filter along the lag. The result is sketched in Figure 12. R_{all} is used to decide which lag values should be considered when searching for line segments in the time-lag matrix. A thresholding scheme based on a discriminant criterion is used. The threshold is automatically set to maximize the following between-class variance of the two classes established by the threshold:

$$\sigma_{\text{B}}^2 = \omega_1 \omega_2 (\mu_1 - \mu_2)^2, \tag{11}$$

where ω_1 and ω_2 are the probabilities of class occurrence (the fraction of peaks in each class), and μ_1 and μ_2 are the means of peak heights in each class.

Each peak above threshold determines a lag value, l_{p}. For each peak, the one-dimensional function $r(\tau, l_{\text{p}})$ is searched over $l_{\text{p}} \leq \tau \leq t$. A smoothing operation is applied to this function and the discriminant criterion of Eq. (11) is again used to set the threshold. The result is the beginning and ending points of line segments that indicate repeated sections of music.

5.6 Modulation Detection

A common technique in pop music when repeating a chorus is to change the key, typically modulating upward by half-steps.

(Note that "modulation" in music is not related to amplitude modulation or frequency modulation in the signal processing sense.) Since modulation changes all the pitches, it is unlikely that a feature vector that is sensitive to pitch sequences could detect any similarity between musical passages in different keys. To a first approximation, a modulation in music corresponds to frequency scaling, as if changing the speed of a vinyl record turntable or changing the sample rate of a digital recording. On a logarithmic frequency scale, modulation is simply an offset, and when the scale is circular as with pitch classes and chroma, modulation is a rotation. To rotate a vector by ζ, the value of the ith feature is moved to become feature $(i+\zeta)$ mod 12. One would expect the chroma vectors for a modulated passage of music to be quite similar to a rotation of the chroma vectors of the unmodulated version.

Goto (2003a) exploits this property of the chroma vector by extending the time-lag matrix to incorporate chroma vector rotation by a transposition amount ζ. Denoting V_t^{ζ} as a transposed (rotated) version of a chroma vector V_t, $r_{\zeta}(t,l)$ is the similarity between V_t^{ζ} and the untransposed vector V_{t-l}^{0}. Since we cannot assume the number of semitones at the modulation in general, the line segment detection is performed on each of 12 versions of $r_{\zeta}(t,l)$ corresponding to the 12 possible transpositions (this usually does not increase harmful false matches). The segments from all 12 versions are combined to form the set of repeated sections of music, and the transposition information can be saved to form a more complete explanation of the music structure.

5.7 Chorus Selection After Grouping Line Segments

Since each line segment indicates just a pair of repeated contiguous segments, it is necessary to organize into a cluster the line segments that have mostly overlapping frames. When a segment is repeated n times ($n \geq 3$), the number of line segments to be grouped in a cluster should theoretically be $n(n-1)/2$ in the time-lag matrix. Aiming to exhaustively detect all the repeated segments (choruses) appearing in a song, Goto (2003a) describes an algorithm that redetects missing (hidden) line segments to be grouped by top-down processing using information on other detected line segments. The algorithm also appropriately adjusts the start and end times of line segments in each cluster because they are sometimes inconsistent in the bottom-up line segment detection. Lu et al. (2004) describe another approach to obtain the best overall combination of segment similarity and duration by adjusting segment boundaries.

A cluster corresponding to the chorus can be selected from those clusters. In general, a cluster that has many and long

segments tends to be the chorus. In addition to this property, Goto (2003a) uses heuristic rules to select the chorus with a focus on popular music; for example, when a segment has half-length repeated sub-segments, it is likely to be the chorus. The *choruslikeness* (chorus possibility) of each cluster is computed by taking these rules into account, and the cluster that maximizes the choruslikeness is finally selected.

5.8 Texture Sequences

Detecting repeating patterns in the similarity matrix is equivalent to finding sequences of similar feature vectors. An alternative is to find sequences of similar texture classes. Aucouturier and Sandler (2002) perform a segmentation using hidden Markov models as described earlier. The result is a "texture score", a sequence of states, e.g., 11222112200, in which patterns can be discovered. They explore two methods for detecting diagonal lines in the similarity matrix. The first is kernel convolution, similar to the filter method of Bartsch and Wakefield (2001). The second uses the Hough transform (Leavers, 1992), a common technique for detecting lines in images. The Hough transform uses the familiar equation for a line: $y = mx + b$. A line passing through the point (x, y) must obey the equation $b = -mx + y$, which forms a line in the (m, b) space. A series of points along the line $y = m_0x + b_0$ can be transformed to a series of lines in (m, b) space such that all intersect at (m_0, b_0). Thus, the problem becomes one of finding the intersection of lines. This can be accomplished, for example, by making a sampled two-dimensional image of the (m, b) space and searching for local maxima. It appears that the Hough transform could be used to find patterns in the similarity matrix as well as in the "texture score" representation.

One of the interesting features of the texture score representation is that it ignores pitch to a large extent. Thus, music segments that are similar in rhythm and instrumentation can be detected even if the pitches do not match. For example, "Happy Birthday" contains four phrases of six or seven notes. There are obvious parallels between these phrases, yet they contain four distinct pitch sequences. It seems likely that pitch sequences, texture sequences, rhythmic sequences, and other feature sequences can provide complementary views that will facilitate structure analysis in future systems.

6 • Music Summary

Browsing images or text is facilitated by the fact that people can shift their gaze from one place to another. The amount of material that is skipped can be controlled by the viewer, and in

some cases, the viewer can make a quick scan to search for a particular image or to read headlines. Music, on the other hand, exists in time rather than space. Listeners cannot time-travel to scan a music performance, or experience time more quickly to search for musical "headlines". At best, one can skip songs or use fast-forward controls with recorded music, but even this is confusing and time-consuming.

One application of music structure analysis is to enable the construction of musical "summaries" that give a short overview of the main elements of a musical work. Summaries can help people search for a particular piece of music they know or locate unfamiliar music they might like to hear in full. By analogy to low-resolution versions of images often used to save space or bandwidth, summaries of music are sometimes called "music thumbnails".

Cooper and Foote describe a simple criterion for a music summary of length L: the summary should be maximally similar to the whole. In other words, a summary can be rated by summing the similarity between each feature vector in the summary with each feature vector in the complete work. The rating for the summary beginning at feature vector i is

$$Q_L(i) = \frac{1}{NL} \sum_{m=i}^{i+L} \sum_{n=1}^{N} S(m,n). \tag{12}$$

The best summary is then the one starting at the value of i that maximizes $Q_L(i)$. The formula can be extended by weighting $S(m,n)$ to emphasize earlier or louder sections of the song.

Other approaches to summary construction are outlined by Peeters et al. (2002). Assume that music has been segmented using one of the techniques described above, resulting in three classes or labels A, B, and C. Some of the interesting approaches to musical summary are:

- Use the most common class, which in popular music is often the chorus. Some research specifically aims to determine the chorus as described earlier (Bartsch and Wakefield, 2001; Goto, 2003a).
- Use a sample of music from each class, i.e., A, B, C.
- Use examples of each class transition, i.e., A → B, B → A, A → C.

In all cases, audio segments are extracted from the original music recording. Unfortunately, artificially and automatically generated transitions can be jarring to listeners. Music structure analysis can help to pick logical points for transitions. In particular, a cut from one phrase of music to a repetition of that phrase

can be inaudible. When a cut must be made to a very different texture, it is generally best to make the cut at an existing point of strong textural change. In most music, tempo and meter create a framework that is important for listening. Cuts that jump from the end of one measure to the beginning of another preserve the short-term metrical structure of the original music and help listeners grasp the harmonic and melodic structure more easily. Segments that last two, four, or eight measures (or some duration that relates to the music structure) are more likely to seem "logical" and less disruptive. Thus, music structure analysis is not only important to determine what sections of music to include in a summary, but also to organize those sections in a way that is "musical" and easy for the listener to comprehend.

An alternative to the construction of "music thumbnails" is to provide a "smart" interface that facilitates manual browsing of entire songs. The SmartMusicKIOSK music listening station (Goto, 2003b) displays a time line with the results of an automatic music structure analysis. In addition to the common stop, pause, play, rewind, and fast-forward controls, the SmartMusicKIOSK has controls labeled "next chorus", "next section", and "prev section" (see Figure 13). These content-based controls allow users to skim rapidly through music and give a graphical overview of the entire music structure, which can be understood without listening to the entire song.

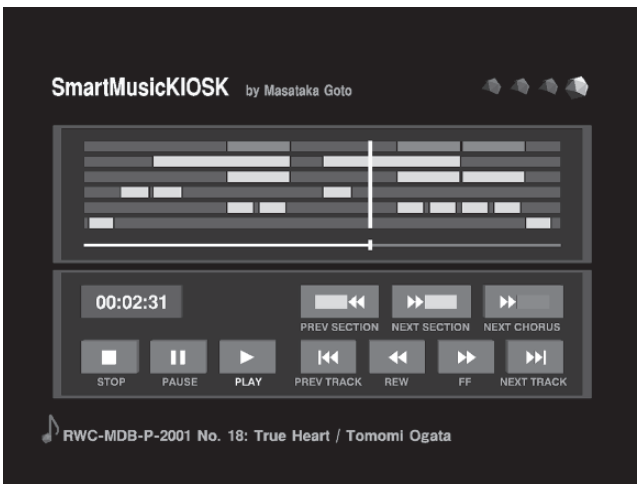


FIGURE 13 The SmartMusicKIOSK user interface showing music structure and structure-related controls.

7 • Evaluation

Most research in this area has been exploratory, with no means to evaluate whether computer-generated structures and segments are "correct". In most cases, it is simply interesting to explore what types of structures can be uncovered and what methods can be used. Quantitative evaluations will become more important as problems are better understood and when competing methods need to be compared.

Tzanetakis and Cook (1999) conducted a pilot study to compare their automatic segmentation with human segmentation. They found most human subjects agreed on more than half of the segments, and their machine segmentation found more than half of the segments that humans agreed upon.

Bartsch and Wakefield (2001) hand-selected "true audio thumbnails" from 93 popular songs and measured "recall", the fraction of true frames labeled by their program as the chorus, and "precision," the fraction of labeled chorus frames that are true frames. With the chorus length set to around 20–25 s, the average recall and precision is about 70%, compared to about 30% for a chorus interval selected at random.

Goto (2003a) also used hand-labeled choruses in 100 popular songs from the RWC Music Database, a source that enables researchers to work with common test data (Goto et al., 2002). Goto judged the system output to be correct if the F-measure was more than 0.75. The F-measure is the harmonic mean of recall rate (R) and precision rate (P): $F\text{-measure} = 2RP/(R+P)$. The system dealt correctly with 80 of 100 songs.

Evaluating music structure descriptions is difficult. Structure exists at many levels and often exhibits hierarchy. The structure intended by the composer and perhaps determined by a music theorist may not correspond to the perception of the typical listener. Nevertheless, one can ask human subjects to identify pattern and structure in music, look for consistency between subjects, and then compare human descriptions to machine descriptions of music. One can also evaluate the impact of music structure detection upon tasks such a browsing, as in SmartMusicKIOSK (Goto, 2003b).

8 • Summary and Conclusions

Knowledge of musical structure can be used to construct music summaries, assist with music classification, provide high-level interfaces for music browsing, and offer high-level top-down guidance for further analysis. Automatic analysis of music

structure is one source of music meta-data, which is important for digital music libraries.

High-level music structure is generally represented by partitioning the music into segments. Sometimes, segments are labeled to indicate similarity to other segments. There are two main principles used to detect high-level music structure. First, segment boundaries tend to occur when there is a substantial change in musical texture. In other words, this is where the music on either side of the boundary is self-similar, but the two regions differ from each other. Secondly, segments can be located by detecting patterns of repetition within a musical work.

It should be noted that the music signal, viewed as a time-domain waveform, is not directly useful for analysis because repetition in music is never exact enough to reproduce phase and amplitude relationships. Therefore, the signal is processed to obtain features that capture useful and more-or-less invariant properties. In the case of texture analysis, features should capture the overall spectral shape and be relatively insensitive to specific pitches. Low-order MFCCs are often used to measure texture similarity. To detect music repetition, features should capture changes in pitch and harmony, ignoring texture which may change from one repetition to the next. The chroma vector is often used in this case.

The similarity matrix results from a comparison of all feature vector pairs. The similarity matrix offers an interesting visualization of music, and it has inspired the application of various image-processing techniques to detect music structure. Computing the correlation with a "checkerboard" kernel is one method for detecting texture boundaries. Using filters to detect diagonal lines is one method for detecting repetition.

Detecting segment boundaries or music repetition generates individual segments or pairs of segments. Further processing can be used to merge segments into clusters. Hidden Markov models, where each hidden state corresponds to a distinct texture, have been applied to this problem. When music is analyzed using repetitions, the structure can be hierarchical, and the structure is often ambiguous. Standard clustering algorithms assume a set of distinct, fixed items, but with music analysis, the items to be clustered are possibly overlapping segments whose start and end times might be adjustable.

Music structure analysis is a rapidly evolving field of study. Future work will likely explore the integration of existing techniques, combining texture-based with repetition-based segmentation. More sophisticated features including music transcription

will offer alternative representations for analysis. Finally, there is the possibility to detect richer structures, including hierarchical patterns of repetition, rhythmic motives, harmonic progressions and key changes, and melodic phrases related by transposition.

Acknowledgments The authors wish to thank Jonathan Foote for Figure 8. Jean-Julien Aucotourier, Mark Bartsch, Jonathan Foote, Geoffroy Peeters, Ning Hu, Xavier Rodet, Mark Sandler, George Tzanetakis, and Greg Wakefield have made contributions through their work, discussions, and correspondence.

References

Aucouturier, J.-J., F. Pachet, and M. Sandler. 2005. "'The Way It Sounds': Timbre Models for Structural Analysis and Retrieval of Music Signals". *IEEE Trans. on Multimedia* 7(6), 1028–1035.

Aucouturier, J.-J., and M. Sandler. 2001. "Segmentation of Musical Signals Using Hidden Markov Models". *110th Convention of the Audio Engineering Society*, Preprint No. 5379.

Aucouturier, J.-J., and M. Sandler. 2002. "Finding Repeating Patterns in Acoustic Musical Signals: Applications for Audio Thumbnailing". In *AES22 Int. Conf. on Virtual, Synthetic and Entertainment Audio*. Audio Engineering Society, pp. 412–421.

Bartsch, M., and G. H. Wakefield. 2001. "To Catch a Chorus: Using Chroma-Based Representations for Audio Thumbnailing". *Proc. the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2001)*. New York: IEEE, pp. 15–18.

Cooper, M., and J. Foote. 2003. "Summarizing Popular Music via Structural Similarity Analysis". Proc. *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2003)*. New York: IEEE, pp. 127–130.

Dannenberg, R. B. 2002. "Listening to 'Naima': An Automated Structural Analysis from Recorded Audio". *Proc. 2002 Int. Computer Music Conf. (ICMC 2002)*. San Francisco: International Computer Music Association, pp. 28–34.

Dannenberg, R. B., and N. Hu. 2002. "Discovering Musical Structure in Audio Recordings". In Anagnostopoulou, C., et al. eds. *Music and Artificial Intelligence, Second International Conference, (ICMAI 2002)*, Berlin: Springer-Verlag, pp. 43–57

Dannenberg, R. B., and N. Hu. 2003. "Pattern Discovery Techniques for Music Audio". *J. of New Music Research, 32*(2), 153–164.

Foote, J. 1999. "Visualizing Music and Audio Using Self-Similarity". *Proc. ACM Multimedia '99*. New York: Association for Computing Machinery, pp. 77–80.

Foote, J. 2000. "Automatic Audio Segmentation Using a Measure of Audio Novelty". *Proc. Int. Conf. on Multimedia and Expo (ICME 2000)*. New York: IEEE, pp. 452–455.

Galas, T., and X. Rodet. 1990. "An Improved Cepstral Method for Deconvolution of Source–Filter Systems with Discrete Speara: Application to Musical Sounds". *Proc. 1990 Int. Computer Music Conf. (ICMC 1990)*. San Francisco: International Computer Music Association, pp. 82–84.

Goto, M. 2003a. "A Chorus-Section Detecting Method for Musical Audio Signals". *Proc. 2003 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2003)*. New York: IEEE, pp. V-437–V-440.

Goto, M. 2003b. "SmartMusicKIOSK: Music Listening Station with Chorus-Search Function". *Proc. 16th Annual ACM Symposium on User Interface Software and Technology (UIST 2003)*. New York: Association for Computing Machinery, pp. 31–40.

Goto, M. 2004. "A Real-Time Music Scene Description System: Predominant-F0 Estimation for Detecting Melody and Bass Lines in Real-World Audio Signals". *Speech Communication (ISCA Journal), 43*(4), 311–329.

Goto, M., T. Nishimura, H. Hashiguchi, and R. Oka. 2002. "RWC Music Database: Popular, Classical, and Jazz Music Databases". *Proc. 3rd Int. conf. on Music Information Retrieval (ISMIR 2002)*. Paris: IRCAM, pp. 287–288.

Hu, N., R. B. Dannenberg, and G. Tzanetakis. 2003. "Polyphonic Audio Matching and Alignment for Music Retrieval". *Proc. 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2003)*. New York: IEEE, pp. 185–188.

Leavers, V. F. 1992. *Shape Detection in Computer Vision Using the Hough Transform*. Berlin: Springer-Verlag.

Logan, B., and S. Chu. 2000. "Music Summarization Using Key Phrases". *Proc. 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings (ICASSP 2000)*. New York: IEEE, pp. II-749–II-752.

Lu, L., M. Wang, and H.-J. Zhang. 2004. "Repeating Pattern Discovery and Structure Analysis from Acoustic Music Data". *Proc. 6th ACM SIGMM Int. Workshop on Multimedia Information Retrieval*. New York: Association for Computing Machinery, pp. 275–282.

Peeters, G., A. L. Burthe, and X. Rodet. 2002. "Toward Automatic Audio Summary Generation from Signal Analysis". *Proc. 3rd Int. conf. on Music Information Retrieval (ISMIR 2002)*. Paris: IRCAM, pp. 94–100.

Peeters, G., and X. Rodet. 2003. "Signal-Based Music Structure Discovery for Music Audio Summary Generation". *Proc. 2003 Int. Computer Music Conf. (ICMC 2003)*. San Francisco: Int. Computer Music Association, pp. 15–22.

Rabiner, L., and B.-H. Juang. 1993. *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice Hall.

Shepard, R. 1964. "Circularity in Judgements of Relative Pitch". *J. Acoust. Soc. Am., 36*(12), 2346–2353.

Tolonen, T., and M. Karjalainen. 2000. "A Computationally Efficient Multi-pitch Analysis Model". *IEEE Trans. on Speech and Audio Processing, 8*(6), 708–716.

Tzanetakis, G., and P. Cook. 1999. "Multifeature Audio Segmentation for Browsing and Annotation". In *Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 1999)*. New York: IEEE.

22

Computer Music Synthesis and Composition

Lydia Ayers

Clear Water Bay, Kowloon, Hong Kong, China,
e-mail: layers@cse.ust.hk

1 • Introduction

What is computer music composition? Composers are using the computer for everything from MIDI instruments communicating with computer sequencers, pitch trackers analyzing the sounds of acoustic instruments and converting them to pitch information, live performers with recorded music, performers with interactive computer programs, computer music produced by dancers using sensors, automatic music composition with the computer programs composing the music, composing with sounds or parts of sounds rather than notes, how to structure the use of time, composing with timbres, or the colors of sounds, and timbre morphing, such as a gong morphing to a voice, composing with textures and texture morphing, such as flutter-tonguing morphing to pitch, granular synthesis, trills and convolution. Computers are especially accurate in tuning, such as just intonation, Pythagorean tuning, Indonesian gamelan tuning, many kinds of equal temperaments and pitch bend, which can

make any tuning even more expressive. The computer can control manipulation of samples in musique concrète and collage. Many composers use computer speech to tell stories, which can take place in fantasy environments with the movements of the sounds in the aural landscapes part of the composition.

What role does synthesis play in computer music composition? While it is true that composers used some of these techniques on acoustic instruments before they composed for computers, the computer has made it possible to use these techniques in new ways and with new sounds.

What makes a performance of computer music dramatic? One of the issues in computer music composition is whether the audience can watch someone perform the music. While we can watch the actions that produce the sounds on acoustical musical instruments in concerts, composers often synthesize computer music in studios using processes that don't occur in real time. Then they save the final mix of the music in a digital medium, such as a binary computer file, tape or CD recording. Some composers then play the recordings in concerts. But how can music with no human performer be interesting to "watch"? Some composers perform their compositions on mixing consoles, moving the sound around on "symphonies" of loudspeakers. Concerts may have such colorful names as the "Concerts of Nothing to See" in Montreal, the "Concerts for Closed Eyes" in Hong Kong or John Cousins' "Cinema for the Ears" at the Australasian Computer Music Conference in Wellington, New Zealand. Other composers team up with dancers or graphic artists to give the audience something to watch. This chapter gives brief descriptions of these aspects of computer music.

2 • Music with Live Performers

2.1 Compositional Interface

The compositional interface is the communication channel between the composer and the computer. The most suitable compositional interface to use for a particular composition depends, in part, on the type of performance. Writing computer code for synthesis languages such as Csound, or programming languages, such as C++, is suitable for recorded music. It is also possible to mix samples in sound editors. Music with live performers, however, requires a program that reacts to instructions quickly enough to perform music in "real time". The interface may include desktop objects in a program such as Max, which a computer performer can manipulate, and even improvise with, in real time. Other interfaces use hardware input devices, such as sensors or electronic musical instruments, to send

signals to the computer which cause the computer to produce sounds. Although some useful commercial electronic musical instruments, such as the Yamaha wind controller, already exist, many composers are inventing their own hardware devices and programming the related software.

2.2 Notation Systems

When live performers are involved in computer music, they need to know what to do, and when to do it, just as performers on acoustic instruments do. Notation should handle performer parts, and it may be helpful if it can also handle reference parts describing the computer music. Notation can also be useful for pitch input for recorded compositions. Computer notation programs, such as Finale and Sibelius, handle notation more or less suitable for traditional western instruments, but their graphics are limited in describing computer sounds. Many computer notation programs, such as Finale and Cubase Score, contain sequencers which can play back the score using synthesized instruments. These programs can store the score information as MIDI files, which can be used in other computer music systems.

Chinese and Indonesian instruments normally use a cipher notation with numbers from 1 to 7. In Chinese music, pitch 1 is the tonic of a major key. In Indonesian music, pitch 1 is the first tone of a five-tone scale which is tuned to the taste of the owner of a gamelan. One computer input method uses the numbers for the pitches, and then transposes the scale up and down by octaves. The octave is set by multiplying the pitch by another power of 2 for each higher octave, so the original octave is 1, the next octave is 2 and the next octave is 4.

2.3 Pitch Input Methods

Composers use several methods to get their musical ideas into computers. Composers using the computer as a notation tool can use the MIDI sequencers in music notation programs, such as Finale, Sibelius and Cubase, to synthesize demos of pieces composed for acoustic instruments, but the quality of the synthesis is limited to the quality of the MIDI instruments available on the computer's sound card. The musical result tends to be more about the notes than the characteristics of the sound, and it doesn't exploit many of the characteristics of computer music. Other composers play MIDI instruments to communicate with computer sequencers, or use pitch trackers to analyze the sound of acoustic instruments and convert it to pitch information. However, the MIDI instruments do not have to be commercial instruments, such as wind controllers and keyboards. Many commercial instruments arose from hardware tinkering,

and such tinkering is still going on. Using MIDI instruments to input pitch information can result in interactive music, where a computer program processes the input information in real time, possibly mixing it with sound synthesis and sending the result to the playback system.

Another computer performance method involves the composer connecting objects on the screen. Using such a system, a composer can create music in a computer environment similar to those on analog synthesizers, such as the Buchla. Max/MSP is one of the most well-known graphical programming environments for music, audio and multimedia. Max users create their own software using a visual toolkit of objects and connect them together with patch cords. The Max environment includes a user interface, MIDI control and timing objects. Hundreds of objects work on top of Max. Jitter is a set of real-time video and image processing and algorithmic image generation objects integrated with the musical capabilities of Max/MSP (Cycling 74 2004).

In Christopher Keyes' *Touched by Space & Time*, microphones capture the sounds of percussion instruments, including circular saw blades, and send them to a computer which alters and spatializes them over the multichannel sound system. Played like an "air marimba", the infrared batons trigger a note when they pass a certain point in the vertical plane, with pitch relative to the

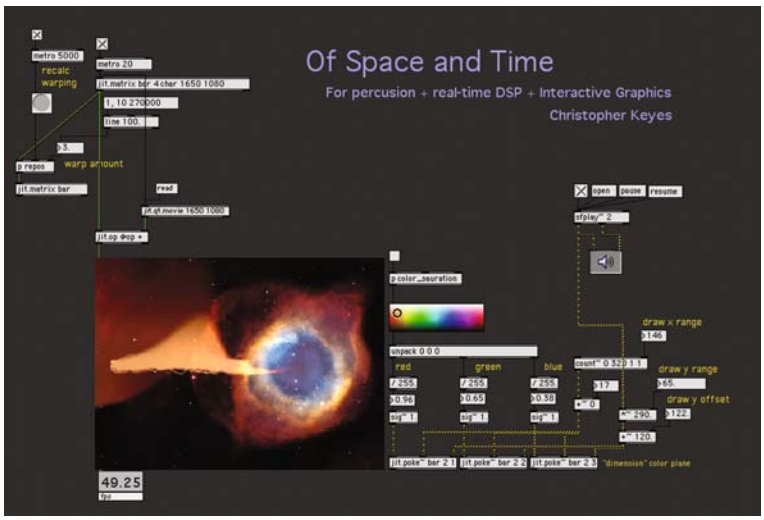


FIGURE 1 Screen shot of Jitter patch for Keyes' *Of Space and Time*.

horizontal plane and dynamics relative to the velocity at the point the baton passes the plane. Keyes wrote a Max/MSP/Jitter patch for the piece to process the incoming percussion audio signals from two microphones and output them over four channels, route MIDI messages to a sampler patch triggering notes in 24- and 36-tone equal temperament, play a four-channel prerecorded accompaniment and send the incoming audio signal to a Jitter patch which generates oscilloscope-like images (Figure 1).

2.4 Live Performers with Recorded Music

Recording a digital computer file and playing it back to accompany live performers is one of the most portable solutions to the problem of mixing performers with computer music. The live performers can play without requiring a computer operator to travel with them. However, the live performers must play in time with the accompanying recording because the recording cannot "follow" them.

2.5 Performers with Interactive Computer Programs

What if the computer could follow the performers? Max Matthews' Radio Baton (Figure 2) is a three-dimensional MIDI controller that tracks the X, Y and Z motions of the ends of two batons held in a performer's hands. At the end of each baton is a small radio-transmitting antenna that transmits a signal to five antennas arranged around the receiving surface. The batons can also send MIDI triggers by crossing invisible

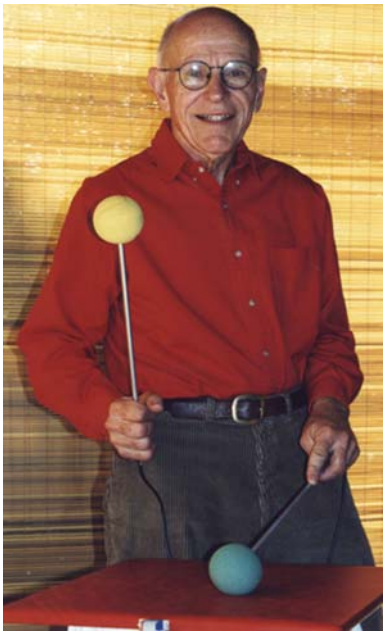


FIGURE 2 Max Matthews and Radio Baton.
http://csounds.com/mathews

planes at various heights above the antenna board. Using a real-time Csound interface on the PC or the Max/MSP interface on the MacIntosh, the performer's movements can control the computer to follow other performers (Mathews). David Jaffe's piece, *Racing Against Time*, uses a similar method with drum sticks that can perform on percussion instruments while they transmit motion data (Jaffe 2001). Todd Winkler (2001) discusses many issues related to interactive music systems in *Composing Interactive Music: Techniques and Ideas Using Max*.

2.6 Computer Music Produced by Dancers Using Sensors

Systems ranging from simple on/off triggers placed on the floor to sophisticated systems incorporating ultrasound, radio and real-time visual analysis convert the movements of dancers into signals to control musical performances. Dancers "play" "active control" systems expressively like musical instruments. With the "passive" approach, the system responds to dancers' normal movements. The audience is usually unaware of how the movements trigger the computer – dancing to a recording can produce the same effect. However, performers like the sense of control that interactive devices provide. Russell Pinkston has used the University of Texas MIDI Dance Floor (Figure 3) in interactive dance compositions, including *Dervish Sketches*, *Memory of Absence* and *Song for the Living/Dance for the Dead*. The dance floor uses force-sensing resistors which transmit the position coordinates, velocity and pressure information as MIDI messages. These are attached to heavy-duty plastic sheeting

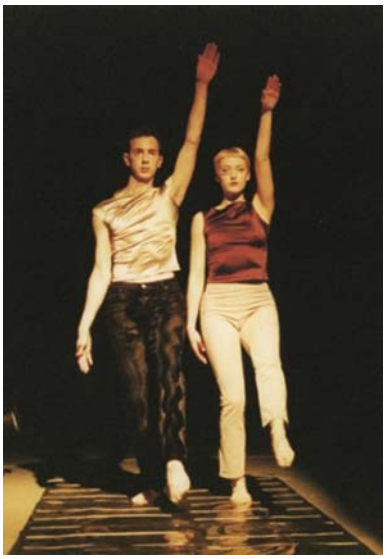


FIGURE 3 Dancers on MIDI Dance Floor.

with contact cement and covered with polyethylene foam. Used with an "intelligent" external MIDI processing system, it permits dancers to affect music and lighting with their movements and positions on its surface (Pinkston et al. 2001).

Chris van Raalte wears a BodySynth to control music, video and animation. Electrodes on different parts of his body control different synthetic and instrumental sounds. His muscle movements produce electrical energy, which a wireless transmitter sends to a computer which controls a MIDI synthesizer. In *Hondo Butoh*, after dancing covered by a large pink fabric that looks like a bedspread, van Raalte emerges looking geeky in glasses and a lab coat. A provocative "striptease" gradually reveals the BodySynth (Figure 4) (Van Raalte and Zane-Cheong 1995).

2.7 Automatic Music Composition

Automatic music composition, also called algorithmic composition or computer-assisted composition, uses computer programs to compose music, and in some cases such automatic music composition produces music for live performers without any computer synthesis at all. Using an Illiac computer at the University of Illinois, Lejaren Hiller and Leonard Isaacson created the first computer-composed composition, *The Illiac Suite for String Quartet*, in 1956 (Roads 1996). Iannis Xenakis described the use of stochastic procedures in *Formalized Music* (1963, 1992) and used such procedures in works ranging from *Eonta* (1964) to *Gendy 3* (1991) and *S.709* (1994).



FIGURE 4 Chris van Raalte wearing the BodySynth.

Xenakis created the latter works using his GENDYN program which functions as a "black box" that can create an entire composition using a small amount of user input. Barry Truax developed a PODX (Poisson Distribution) system, which he used to compose with sounds and parts of sounds, rather than notes, in pieces such as *Inside* (1995) and *Beauty and the Beast* (1989). In my pieces, *Balinese* and *Pendopos* (Ayers 2006), I wrote C++ computer programs which composed Csound scores algorithmically, loosely within the Balinese gamelan style. Some elements of this style are the rhythmic structure and pentatonic scales using Balinese tunings (McPhee 1966). *Pendopos* is a walk up an imaginary mountain past open pavilions playing gamelan music.

3 • Form, Structure and Time

How to structure the use of time is an interesting question in computer music composition. While it is possible to create a symphony in sonata form, it seems more idiomatic to use experimental forms. In concert performances, the audience cannot tell when an invisible piece will begin, nor is it obvious whether a pause in the sound is the end of the piece or a break between movements. This characteristic of computer music has encouraged one-movement works, or works where sections flow together. Everything can seem quite static, or the movement can have contrasting sections.

Composers often play with the way we perceive time. Time can move disconnectedly as points of sound or smoothly as a slowly evolving drone. Pierre Boulez coined the term, "striated time", to describe music with meter, beats or at least events that can be counted, even if they are irregularly spaced. "Smooth time" occurs in compositions in which it is difficult to count time (Boulez 1963, 1971). Spectrally evolving drones may be completely smooth, with no events marking time. In smooth time, listeners may even lose their awareness of how much (or how little) time is passing. Computer music contains many shades between completely smooth time and completely striated time. Curtis Roads' *Field* (Roads 1981) contrasts smooth time with striated time in "a flowing musical process". The piece begins with long filtered noise sounds, which seem to be nearly a textured drone. The spectra evolve as the composition unfolds. He combined individual sound objects into compound sound mixtures, using computer synthesis techniques, such as frequency modulation, waveshaping, additive and subtractive synthesis and granular synthesis, or mixing them with digitized natural sounds. He processed the compound sounds and linked

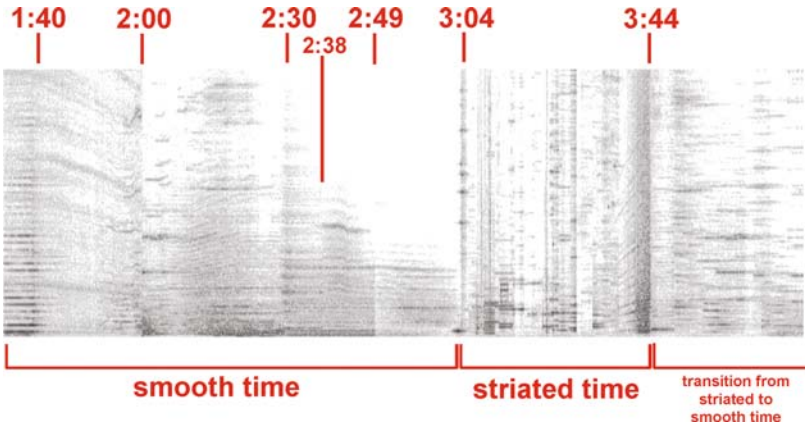


FIGURE 5 Smooth and striated time in Roads' *Field*.

them into layered melodies. Interest focuses on the qualities of the sounds as they slowly morph into other sounds. Figure 5 shows an episode of striated time interrupting the smooth time.

Striated time layers can move at different speeds. I used this technique in *Nostalgia Strata*, which remembers some familiar twentieth century melodies using some of the woodwind designs from our book, *Cooking with Csound* (Horner and Ayers 2002). Slices of the famous melodies form a "strata", which suggests both fugal layers and the name of a layered dish. Heavy-effects processing masks a lot of the melody, but the layers in this piece form a distorted fugue, with the "subject" reappearing at different pitch levels in simultaneously different tempi, stretching and contracting the stretto effect.

4 • Timbre and Texture

Timbre is the color of the sound, the quality that distinguishes between, say, a trumpet and a violin, and in computer music composition it ranges from pitched musical instrument tones to environmental sounds, such as wind, water and birds, to artificial sounds. Timbre may be independent of and more important than rhythm and pitch. Timbre morphing, such as a gong morphing to a voice, slows down spectral interpolation to the point that the focus is on the spectral change process rather than on the resulting sound goal.

Larger scale texture has been described as relationships between monophonic, homophonic, polyphonic and heterophonic melodies and harmonies. Another way to use texture is on the surface of a sound, as silk usually has a smoother texture

than wool fabric. This texture may be smooth versus striated or continuous versus discontinuous.

4.1 Granular Synthesis

Granular synthesis uses many short overlapping grains of sound. Multiplying samples or waveforms by amplitude envelopes creates grains. Mara Helmuth's StochGran stochastic granular synthesis program generates complex sounds from tiny grains of recorded sounds, which maintain some relationship to the original source (Helmuth 2002). She used granular synthesis to produce clouds of sound in *Loonspace* and *Abandoned Lake in Maine*, and at the end of *Implements of Actuation* (co-composed with Allen Otte) (1998).

Barry Truax used granular synthesis in pieces such as *Inside* (Truax 1995) and *Beauty and the Beast* (Truax 1989).

4.2 Flutter Tonguing

Flutter tonguing is a grainy texture on wind instrument tones that is accomplished by rolling the "r" at the front of the tongue or gargling in the back of the throat. The movement of the performer's throat or tongue produces amplitude modulation near 20 Hz, the border frequency between clicks and tone. We modeled it using a function that represents the average amplitude envelope of the individual modulations (Ayers 2004). Changing the speed of amplitude modulation on a synthesized bassoon tone morphs the texture from flutter tonguing to pitch (Figure 6).

4.3 Trills

Trills are a special case of a slur and are similar to flutter tonguing as well. A repeating function controlling the frequency of a single tone produces smoother trilled notes than overlapping tones does, and we have used amplitude and frequency modulation functions to cross-fade the trills in wavetable instrument designs

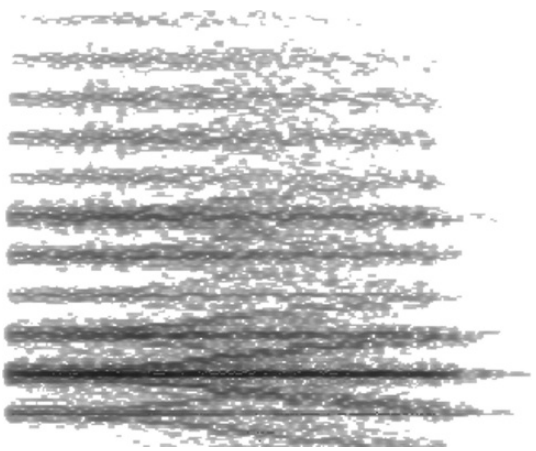


FIGURE 6 Synthesized bassoon tone with flutter tonguing morphing to pitch.

(Ayers 2003, Ayers 2004, Ayers and Horner 2004a and b). An additive synthesis design improves the transitions between the trilled notes. The additive synthesis instrument uses a single signal with a line segment on all the parameters for each of its harmonics. That is, each harmonic is one component signal for the spectrum, with a frequency and an amplitude line segment handling a varying number of notes. Then all the harmonics are added together. The repeating trill function models the frequency change for the trill. The function does not need to model pitch variation of the average tone, change of speed or jitter, so it can represent one average cycle of the trill, and adjusting the parameters randomly within their typical ranges can vary each cycle. The amplitude function modulates the difference between the harmonic amplitudes of the lower and higher notes to give each note in the trill the correct spectrum. The design controls an initial, middle and final tremolo rate and the time required to change from the first tremolo rate to the second. Figure 7 shows an example of speeding up the trill until it becomes frequency modulation. A separate score parameter controls the changing amplitude of the trill as it would for a single sustained tone (Ayers 2005, 2006, Ayers and Horner 2006, 2007).

I am using these designs in a number of compositions. I first used it for the Chinese *erhu* (Figure 8) in *Why Ask Questions*, which was composed as a background music for an installation art exhibition. At the time, the 1997 Handover was on the horizon, and many in Hong Kong were questioning what the future would bring. Others asked, "Why ask questions? Why not just accept fate?" I composed the original 30-min microtonal piece algorithmically (using the computer to make some of the compositional decisions). Then I added multi-effect processing to Chinese voices

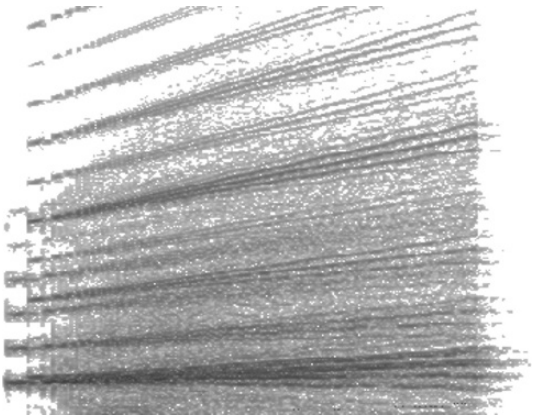


FIGURE 7 Changing speed from trill to frequency modulation.

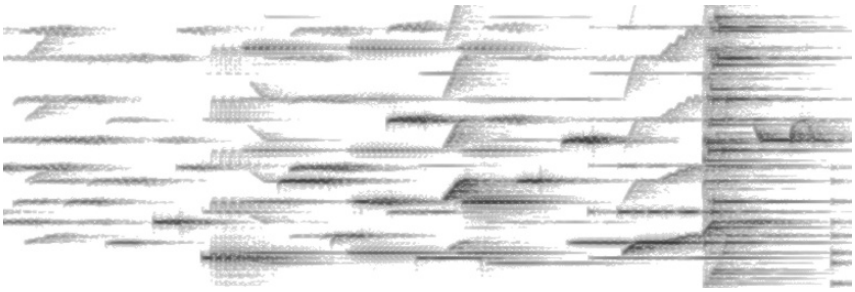


FIGURE 8 *Erhu* slurs, pitch bend and trills in Ayers' *Why Ask Questions*.

and Chinese computer instrument designs for *dizi, xiao, erhu, pipa* and gong.

4.4 Convolution

Convolution, a process of multiplying all the samples in one signal by all the samples in another signal, is often used in cross-synthesis, filters and reverb (Boulanger 2000, Roads 1996). John Rimmer's *Ancestral Voices* (Rimmer 2000) explores sampled sounds, including the Maori trumpet-like, male *putorino*, and the flute-like, female *koauau*, the Australian didgeridoo, three Asian instruments and two European instruments. Rimmer colored the samples with convolution and contrasted the active gestures and granulations of the "male" sections with the "delicate, colored and highly resonant sonorities" of the "female" sections. That is, the female sections are in smooth time, and the male sections are more striated.

5 • Tuning Systems

Computers are especially accurate with tuning. Virtually any kind of tuning is possible, paired with any timbre, so it is possible to have real or imaginary computer musical instruments play in tunings that acoustic instruments can only wish for.

5.1 Just Intonation

The just intonation tuning system uses integer ratios to calculate frequencies at intervals from an originating frequency. Opinion differs about which ratios fall into the category of just intonation. Some people limit it to only the simplest integer ratios, such as those needed for the major and minor scales. However, the major scale can require two ratios for the major second in order to have consonant intervals in typical chords. The major second, 9/8, is typically used in the just major scale, but harmonic music may add 10/9 to form a consonant minor third (6/5) interval with the perfect fourth (4/3) and major sixth (5/3) in the *ii* chord

$(10/9 \times 6/5 = 4/3)$. *The Just Intonation Primer* (Doty 1993) gives an introduction to just intonation, and *Genesis of a Music* (Partch 1974) gives a very detailed discussion.

Other people prefer a wider view of just intonation, such as simple or small integer ratios. It is a matter of opinion how small "simple" integers should be. The most unlimited definition is any integer ratios, but some intervals formed that way are quite dissonant. Depending on the individual definition, just intonation may or may not include dissonant intervals. However it is defined, the ratios should contain integer fractions, so real numbers such as the 12th root of 2 are not included.

Bill Alves piece, *In–Yo* (Alves 1999), uses just intonation with a live shakuhachi player and computer accompaniment. *In* and *yo*, the Japanese equivalent to the Chinese *yin* and *yang*, are also the names of two Japanese scales. *In*, a scale of large and small intervals, is associated with meditation, darkness and melancholy. In contrast, *yo* has no very small intervals, and expresses the extroverted joy of folk music. Alves represents these perspectives with two systems of just intonation. The *in* section uses "utonal" intervals in which the pitches are related consonantly to each other, but not necessarily to a common fundamental. In the *yo* section, Alves used "otonal" intervals, where all the pitches are harmonics related to a single fundamental (Partch 1974). Alves used Csound to manipulate the breath sounds of a bamboo flute, the deep tolling of a Buddhist temple bell (in which the partials have been altered to match the tuning system) and the human voice.

5.2 Pythagorean Tuning

The "Pythagorean" tuning method was actually first discovered by the Chinese musician Ling Lun (Partch 1974) and was later named after Pythagoras in Greece. On acoustic instruments, the first interval is tuned as a perfect 5th, then another interval is tuned to a perfect 5th from the upper note and so on until all the notes are tuned. Transposing tones up or down by octaves organizes them into scales. The computer produces Pythagorean tuning by multiplying 3/2 by 3/2 and then repeating until it produces all the desired notes.

My piece, *Bioluminescense* (Ayers 2006), uses Indian tunings from ragas of the deep night, including *Paraj*, *Sohini*, *Lalita*, *Vibhasa*, *Durga*, *Hambir* and *Khammaja*. According to Daniélou (1980), some of these ragas, along with many others, use Pythagorean tuning. *Bioluminescense* also explores the boundary between tuning and timbre. Gamelan-like timbres result from a tuning using groups of 17-denominator ratios, generally with the higher frequencies dropping out before the lower frequencies

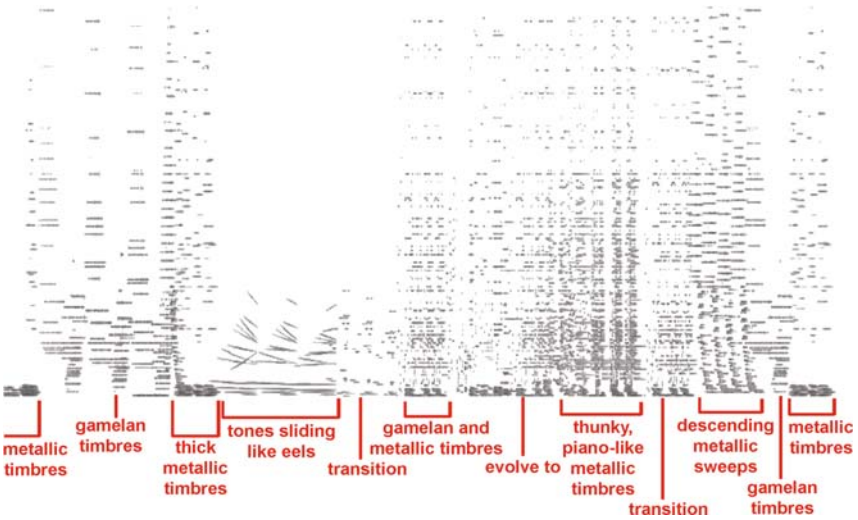


FIGURE 9 End of Ayers' *Bioluminescense*.

(Figure 9). Metallic timbres result from a relationship among high harmonics, the quality of which changes when it is detuned. In this piece, I also flirted with just intonation utonalities and combined aggregates of ratios which are not closely related to each other, such as a section combining ratios of 7, 8, 13, 19 and 23. Some of the thickest aggregates give metallic and strange, thunky timbres resulting from detuning the high harmonics.

5.3 Equal Temperaments

Computers can easily produce accurate 12-tone equal temperament, the division of the octave into 12 equal pieces, which is the hoped-for tuning in western music for about the last 150 years. And they can go beyond it into many other equal temperaments. Equal divisions of 19 and 31 are especially popular. Equal divisions of 13 sound especially strident. Bill Sethares has exploited a number of octave divisions in his *Xentonality* CD (Sethares 1997), which includes, among others, *Ten Fingers* in 10-tone equal temperament, *Seventeen Strings* in 17-tone equal temperament and *Unlucky Flutes* in 13-tone equal temperament. *Chaco Canyon Rock*, included on the CD in his book, *Tuning, Timbre, Spectrum, Scale* (Sethares 1998), is an example of Sethares' technique mapping the spectrum of a sound to the tuning of the scale he uses in the piece. He used a fast Fourier transform to analyze 12 samples of a circular rock from Chaco Canyon. Striking different positions on the rock produced different sounds, each with a unique spectrum. All of the sounds have a resonance near 4,070 Hz, and other resonances occur in

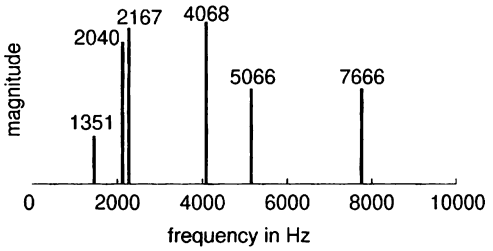


FIGURE 10 Sethares' combined line spectrum for *Chaco Canyon Rock*.

just one or two samples. Sethares combined the most prominent partials into a composite line spectrum that captures much of the acoustic behavior of the samples (Figure 10). Sethares then used a program to draw the dissonance curve for the tone. He derived his scale (Table 1) for the rock instrument from the most consonant points on the dissonance curve. The *Chaco Canyon Rock* piece uses sampled tones tuned to this scale and electronic rocks produced with additive synthesis of the inharmonic partials shown in Table 1.

5.4 Indonesian Tunings

While Bill Alves has composed music for Indonesian gamelan instruments, and Lou Harrison composed for justly tuned gamelan instruments that Bill Colvig built for him, these pieces depend on the physical instruments for which they are composed. Playing them on other instruments requires changing their tunings, which readily occurs in Indonesian gamelan music. Intonational differences, however, may not always please contemporary American composers. Virtual gamelan instruments, on the other hand, can stay in a single desired tuning or play simultaneously in multiple tunings. In *Merapi* (Ayers 2006), inspired by the name of Java's most active volcano, I used the first approach, a single hypothetical tuning derived from the tuning of the Berkeley Gamelan (Ayers 1996). The computer program randomly generated the piece within the style of Indonesian gamelan music. In addition to making Risset's bell design (Risset 1969, 1996) less inharmonic for the *bonang* (a group of tuned horizontal gongs) by tuning its partials to harmonics, I adapted Judy Klein's design for the sound of struck crystal glasses. "Waiting" opens the piece with a fairly tradi-

TABLE 1 Scale used by Sethores for the rock instrument

Ratio	1	1.063	1.17	1.24	1.25	1.37	1.385	1.507	1.612	1.69	1.77	1.88	2
Cents	0	105	272	372	386	545	564	710	827	908	988	1,093	1,200

tional rhythmic structure, “Simmering” is a slow section with strange chords, the “Eruption” is a meterless tape interlude in which the tuning gradually rises to symbolize the eruption of the volcano and the “Bubaran” is a short closing piece. I used a variety of Balinese tunings (McPhee 1966) in two other algorithmically composed pieces, *Balinese* and *Pendopos* (Ayers 2006), a walk up an imaginary mountain past open pavilions playing gamelan music. The computer program generated several short segments using Balinese tunings (McPhee 1966), loosely within the Balinese gamelan style. The synthesized gamelan reuses the *bonang* adaptation of Risset’s bell design and the Woodstock gamelan design (Ayers and Horner 1999). I also synthesized my voice and a Balinese *suling* flute. Each section of *Pendopos* represents a pavilion. As the trail curves around, you hear the first two gamelan groups again, playing the same fragments transposed to different pitch centers with a changing background full of chattering birds, including a few ducks, that I recorded in Bali. Suddenly, passing the tree line, the gamelan layer changes from the heavier Risset gongs (Risset 1969, 1996) to the bright and sparkly Woodstock gamelan, which accompanies the voice. At the summit, the voice swirls around everywhere as the bird calls fade into the distance.

5.5 Overview of How to Tune a Computer

The following example shows one way to tune a computer using Csound. First, set a global tuning frequency. One number in the note statement can control all the instruments in the orchestra. For example, for C major, middle C, at 261.6 Hz, is a good note for a tuning frequency, and for A minor, A at 440 Hz is a good tuning frequency. To modulate to another key in the middle of a piece, simply set a new global tuning frequency. The Csound orchestra code for instrument 1 uses the global frequency and attenuation values from the score. As these values are initialized at the start time of the note, the duration is irrelevant as long as it is greater than 0, so I just use a short value such as 1.

```
;-------------------------------------------------------------------------
instr 1
gifreq      =       p4      ; set global tuning frequency
giattn      =       p5      ; set global attenuation scaling
endin
;-------------------------------------------------------------------------
```

Once the global frequency is set, a few simple lines of code in the orchestra set the frequency for the individual notes.

```
;-------------------------------------------------------------------------
instr 25
. . .
ipitch      =     p8
ioct        =     p9
ipitchtable =     9                    ; table of indexes for
                                         table 10
iratiotable =     10                   ; table of frequency ratios
iprange     table ipitch*100, ipitchtable ; pitch range of note
inum        table (iprange*2), iratiotable    ; numerator of note
iden        table (iprange*2)+1, iratiotable  ; denominator of note
ifreq       =     inum/iden * gifreq          ; ratio x global freq
ifreq       =     (ioct = 0 ? ifreq/2 : ifreq * ioct)
. . .
;-------------------------------------------------------------------------
```

Function tables 9 and 10 control the tuning of the pitches. Function table 10 sets just intonation in the following example. The input pitches for a major scale consist of the integers 1–7. The input method is more intuitive if the chromatic pitches occur between the integers. For example, the input for pitch 3b, the minor third, would be 2.5. As the indexes to table 10 must be integers, table 9 provides the conversion. Csound then reads the pair of numbers occurring at the indexed position and the following position in table 10 as the numerator and denominator of the frequency ratio for the tone.

```
;-------------------------------------------------------------------------
; function table for pitch to choose ratios
f9 0 16384 -17 0 0 105 1 205 2 255 3 305 4 405 5 505 6 605 7 705 8
; function table for just intonation ratios (commented out)
f10 0 64 -2 1 1 9 8 6 5 5 4 4 3 3 2 5 3 15 8
;-------------------------------------------------------------------------
```

5.6 Pitch Bend

Pitch bend often occurs in natural sounds, such as bird songs and cat meows. My piece, *Yazi's Dream*, celebrates the life of Yazi, my first cat. "Looking at Birds" represents the birds she watched and tried to talk to through the window. The synthesized trills on the *dizi*, a Chinese membrane flute which resembles a keyless bamboo piccolo, represent Yazi's bird fantasies. The excerpt shown in Figure 11 includes a *dizi* melody and trill, lowered cat sounds with natural pitch bend, stretched (slowed) and granulated lowered cat sounds, and normal and lowered crow sounds with natural pitch bend.

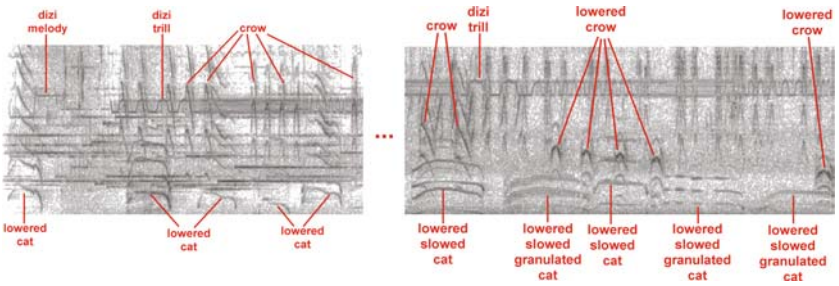


FIGURE 11 Pitch bend in Ayers' *Yazi's Dream*.

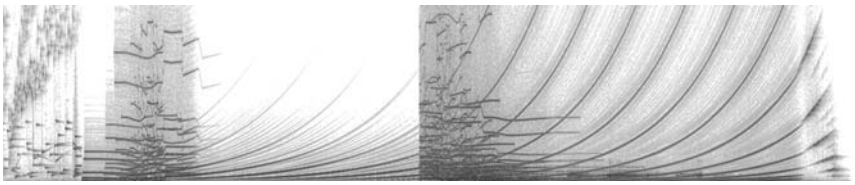


FIGURE 12 Endless ascending glissandi in Risset's *Mutations*.

The endless glissando, an illusion in which the pitch seems to keep going up (or down) but never gets anywhere (Dodge and Jerse 1984), is another use of pitch bend. French composer Jean-Claude Risset used the endless glissando in *Mutations*, synthesized using Max Matthews' Music V program at Bell Labs in 1969. The piece uses mostly additive synthesis and is the first work using FM, based on John Chowning's work. A melodic motive transforms into a chord, which stretches into a timbre. The density of the frequencies increases and gradually mutates to an endless ever-ascending glissando (Figure 12).

6 • Musique Concrète and Collage

Although most people barely notice the sounds surrounding them, some have found those sounds musical. For example, Benjamin Franklin assembled a glass harmonica from wine glasses, and, in the middle of the twentieth century, such composers as Pierre Schaeffer and Luciano Berio recorded collages of such concrete, or real, sounds as locomotives (Schaeffer 2000a,b) and voices (Berio 1999). Their compositions often took months of cutting and splicing taped sounds together. Schaeffer described these techniques extensively in the *Traité des objets musicaux* (Schaeffer 1966). Now, it is much

easier to cut and paste recorded samples together to create music with computer "instruments".

6.1 Sampling

One evening after dinner, Paul Lansky's sons played everything in their kitchen which would make noise. In *Table's Clear*, Lansky makes syncopated gamelan music mixed with a jazz chorus or two from such sounds as pots and pans, glasses, tupperware drums, water drops and the boys talking, belching, burping and clapping. In the middle, the gamelan-like patterns twist as they modulate from one pitch center to another (Lansky 1992).

Joyce Tang derived the title *Cympergon* (1996) from the names of its cymbal, paper and gong sound sources. She then transformed these "sonic seeds" into usable computer "instruments" for sampling and sound-wave editing. The imagined motions of striking, scratching, tearing and rubbing these instruments are an important part of the piece's character. Combining the gong timbres with polymetric and polyrhythmic patterns gave her piece a gamelan-like flavor, but pitch bending and time stretching transform the gongs and other timbres into an imaginary landscape (Figure 13).

In *Kasumi*, Ian Whalley sampled seven New Zealand Mäori instruments, played by Richard Nunns, and Chikako Komaki reciting a Japanese poem. The samples of the Mäori wind instruments are melody fragments, rather than isolated tones, and Whalley repeats these fragments and eventually disintegrates them. The composition is constructed in levels like a folk tale. Figure 14 illustrates the beginning of the piece, which begins with a *tumutumu* (a pilot whale jawbone struck with wood), a *pahü pounamu* (small Greenstone gong) and two higher layers of melody fragments on the *köauau köiwi kuri* and *köauau pongäihu*

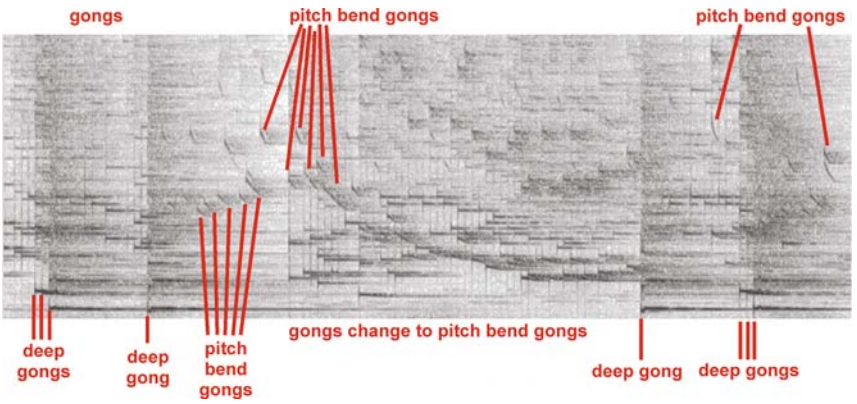


FIGURE 13 Gongs from Tang's *Cympergon*.

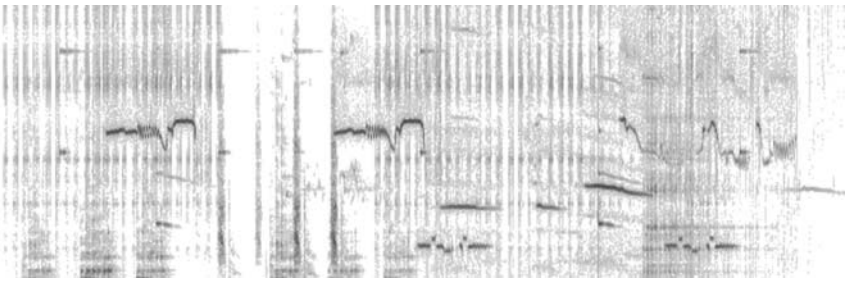


FIGURE 14 Beginning of Whalley's *Kasumi*.

(dog bone and nose flutes). Whalley manipulates the sample fragments later in the piece with a great deal of processing, and also uses many synthetic instruments (hybrids of electronics and real instruments) (Whalley 2002).

In *Implements of Actuation* (1998), Mara Helmuth and Allen Otte transformed the sounds of an electric mbira and a kid's rusty bicycle wheel into a delicate texture (Figure 15) using a Javanese scale and a computer. Implements applied to the mbira and bicycle wheel, such as a Chinese wok brush, red lentils, comb, credit card and baseball trading card, became subjects to be acted upon by the computer applications used both as a mixing tool for all of the collected fragments which became the tape part and also as an algorithmic generator of both the tape and the live performance score.

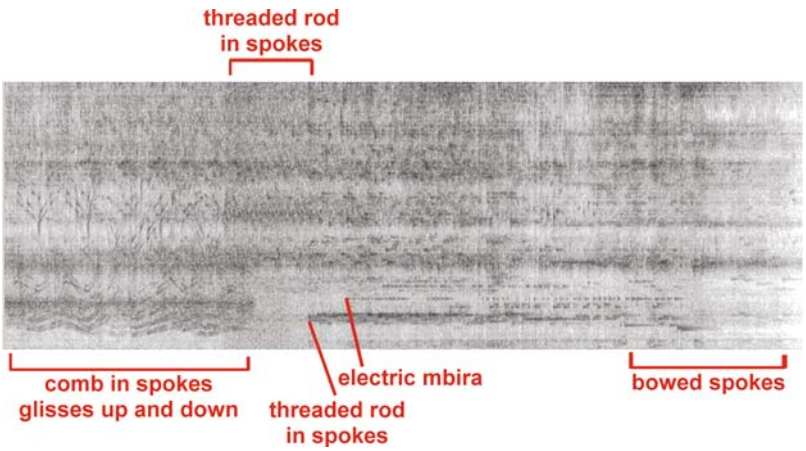


FIGURE 15 Sampling in Helmuth's *Implements of Actuation*.

6.2 Speech, the Voice and Program Music

Many composers have been inspired by the possibilities of manipulating voices on the computer. For example, in *Idle Chatter*, Paul Lansky transformed unintelligible chattering speech into rhythmic chanting (Lansky 1985). Charles Dodge was attracted to "making music out of the nature of speech itself", to "changing the timbre of the voice" in *He Destroyed Her Image*. Dodge tried to depict the reversal, from looking outside to being inwardly confused in a poem by Mark Strand, with the changes of tone quality in the voice, back and forth between a speechlike electronic phrase (you can understand the words) and a less speechlike electronic phrase (you can't understand the words). This happens even though the two have the same pitch patterns (Dodge 1972).

Program music paints a picture or tells a story. Anna Rubin's *Family Stories*, co-composed with Laurie Hollander, used Paul Koonce's phase vocoder instruments (Koonce http://silvertone.princeton.edu/winham/PPSK/koonce.html) in a text/sound piece using narrative, sampled ambient sounds and computer-generated music to tell the story of Rubin's mother, Sophie, the child of Russian Jewish immigrants in Atlanta. At seven, Sophie lost her mother and was left in the care of an African-American woman, Sally Johnson. Sophie's unbearable loss is framed by the atmosphere of racism and anti-Semitism in early twentieth-century Atlanta (Rubin and Hollander 2004).

7 • Environments

Soundscapes are landscapes "painted" with sounds. For example, while "on the ice" in Antartica, composer Chris Cree Brown (NZ) recorded the sounds of walking on snow, skua birds, polar wind, radio communications, Weddell seals, an Adelie penguin rookery and Antarctic white out. In his piece, *Under Erebus*, he sculpted his recordings into a soundscape reflecting his experiences of the Antartic environment (Brown 2000).

7.1 Spatialization and Composing an Imaginary Space

Jean-Claude Risset's *Invisible* is inspired by Calvino's book, *Le città invisibili*, in which Marco Polo describes to Kublai Khan cities from his dreams, desires, fantasies, fears and utopias. With digital synthesis and processing using MUSIC V, SYTER and Sound Mutations, Risset produced harmonically composed timbres, voice and wind hybrid sounds and illusory bells, gongs and voices that dialog with the *Invisible* real voice of soprano Irene Jarsky in imagined spaces similar to Calvino's invisible cities. Through spatial movement (Figure 16), Risset transforms the voice into the butterfly described in the texts of Taoist Chinese poet Zhuangzi (Risset 1996).

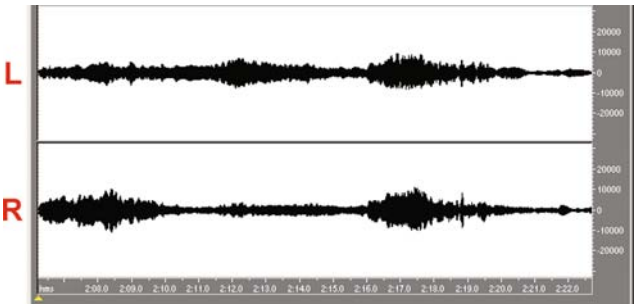


FIGURE 16 Stereo spatialization in Risset's *Invisible*.

Mark Applebaum composed *Pre-Composition* (2002) using only his voice and put each of the eight characterizations of himself in a different speaker. The result is a humorous piece about his internal dialog whenever he plans a piece of computer music.

Australian composer Jim Barbour recreated an outback waterhole in *Outback* (Barbour 2004). After converting the subwoofer into a regular channel, he used four horizontal loudspeakers and two overhead loudspeakers in the 5.1 surround sound format to create a three-dimensional sound environment. The sounds move over and around listeners with reverberation reflections from all over the vertical hemisphere. Some sources sound at ground level, some in trees while others are flying around and over a listener, as they would in their original environment. The wild budgerigar birds, for example, fly and sing in the trees only from the elevated loudspeakers (Barbour 2004, 2006).

7.2 Composing for a Space

An installation is a continuous sound design that fits a place rather than a composition for a set length of time. Installations include realistic or imaginary sound scenes all around the space. French composers Michel Redolfi, Luc Martinez and Michel Pascal composed pieces for an installation for the Nausicaa Exposition. They mixed sounds recorded in the Pacific Ocean and South China Sea into music in their studios. A central computer controls three-dimensional sound circulation around the public through all levels of the museum. The Chalutier room simulates the fury of a fish in the high sea in a sort of eight-channel acoustic "holograph". As people walk through the Remontée, their random movements interactively recompose the sound. Synthesized music and biological sounds evoke a scientific sea where 20,000 known sonorous species murmur the last siren songs of a strange opera (Redolfi 1990–91).

8 • Conclusion

This chapter has given brief examples of some of the major aspects of computer music composition. Computer music takes place, as all music does, during time and within the mind. I have described computer music which plays purely in the sonic dimension and works using additional "dimensions", such as spatialization and the visual dimension in multimedia works. Some composers perform their compositions on mixing consoles, moving the sound around on "symphonies" of loudspeakers. Other composers team up with performers with interactive computer programs, dancers using sensors or graphic artists to give the audience something to watch. I have described the use of computers in various types of tuning systems, automatic music composition with the computer programs composing the music, and how the computer has made it possible to use synthesis techniques in new ways and with new sounds. Composers are using the computer for composing with sounds or parts of sounds rather than notes, composing with timbres and timbre morphing, composing with textures and texture morphing, granular synthesis, trills and convolution. Many composers use computer speech to tell stories in fantasy environments with sounds dancing toward the horizon in the aural landscapes.

Acknowledgments Many thanks to the composers whose works are cited in this chapter. Details in many of the piece descriptions have been summarized from their program notes.

References

Publications

Ayers, L. 1996. "A Composition for Gamelan and Computer-Generated Tape: *Merapi*", *Leonardo Music Journal*. **6**: 7–14.

Ayers, L. 2003. "Synthesizing Trills for the Chinese Dizi", *Proceedings of the 2003 International Computer Music Conference*. Singapore, pp. 227–230.

Ayers, L. 2004. "Synthesizing Timbre Tremolos and Flutter Tonguing on Wind Instruments", *Proceedings of the 2004 International Computer Music Conference*. Miami, FL, pp. 390–393.

Ayers, L. 2005. "Synthesizing Chinese Flutes Using Csound", *Organised Sound*. **10**(1): 37–49.

Ayers, L. 2006. "Fantasy Birds in Yazi's Dream", *Proceedings of the 2006 International Computer Music Conference*, New Orleans, LA, p. 33.

Ayers, L. and Horner, A. 1999. "Modeling the Woodstock Gamelan for Synthesis", *Journal of the Audio Engineering Society*. **47**(10): 813–823.

Ayers, L. and Horner, A. 2004a. "Expressive Modeling of Chinese Folk Wind Instruments", *Journal of the Acoustical Society of America*. **116**(4): 2619.

Ayers, L. and Horner, A. 2004b. "Synthesis of Chinese *Dizi* Ornaments", *Ghost in the Machine: Performance Practice in Electronic Music, Australasian Computer Music Association Conference*. Wellington, New Zealand.

Ayers, L. and Horner, A. 2006. "Synthesizing the Chinese *Suona*", *Medi(t)ations, Australasian Computer Music Conference*, Adelaide, Australia, pp. 15–22.

Ayers, L. and Horner, A. 2007. "Synthesizing the Chinese *Suona*", *Asia Pacific Festival and Conference*. Wellington, New Zealand.

Barbour, J. 2004. "Exploration of the Height Dimension in Audio Reproduction", *Australasian Computer Music Conference*, Wellington, New Zealand.

Barbour, J. 2006. "Creating Acoustic Space", *Australasian Computer Music Conference*, Adelaide, Australia.

Boulanger, R., Ed. 2000. *The Csound Book: Perspectives in Software Synthesis, Sound Design, Signal Processing and Programming*. MIT Press, Cambridge, MA.

Boulez, P. 1963. *Penser la Musique Aujourd'hui*. Editions Gonthier, Geneve.

Boulez, P. 1971. *Boulez on Music Today*. Bradshaw, S. and Bennett, R.R., trans. Faber, London.

Cycling 74. 2004. http://www.cycling74.com/products/maxmsp.html.

Daniélou, A. 1980. *The Ragas of North Indian Music*. New Delhi: Munshiram Manoharlal.

Dodge, C. and Jerse, T. 1984. *Computer Music*. 1st ed., Schirmer, New York.

Dolson, M. 1986. "The Phase Vocoder: A Tutorial", *Computer Music Journal*. **10**(4): 14–27.

Doty, D. 1993. *The Just Intionation Primer: An Introduction to the Theory and Practice of Just Intonation*. The Just Intonation Network, San Francisco, CA.

Helmholtz, H. 1954. *On the Sensations of Tone* 2nd ed., Dover Publications, New York, trans. Alexander J. Ellis. p. 90.

Helmuth, M. 2002. "StochGran on OSX", *Proceedings of the 2002 International Computer Music Conference*, Gotheborg, pp. 77–78.

Helmuth, M. StochGran available at: http://meowing.ccm.uc.edu/softmus.htm.

Horner, A. and Ayers, L. 1998a. "Modeling Acoustic Wind Instruments with Contiguous Group Synthesis", *Journal of the Audio Engineering Society*. **46**(10): 868–879.

Horner, A. and Ayers, L. 1998b. "Modeling Chinese Musical Instruments", *Proceedings of the 135th Meeting of the Acoustical Society of America*. Seattle, WA, **4**: 2541–2542.

Horner, A. and Ayers, L. 2002. *Cooking with CSound, Part 1: Woodwind and Brass Recipes*. AR Editions.

Horner, A., Ayers, L. and Law, D. 1999. "Synthesis Modeling of the Chinese *Dizi*, *Bawu*, and *Sheng*". *Journal of the Audio Engineering Society*. **47**(12): 1076–1087.

Koonce, P. "PVC", http://silvertone.princeton.edu/winham/PPSK/koonce.html.

Mathews, M. *The Radio Baton*. http://csounds.com/mathews/

McPhee, C. 1966. *Music in Bali*. Yale University Press, New Haven.
Partch, H. 1974. *Genesis of a Music*, 2nd ed. Da Capo, New York.
Pinkston, R., Kerkhoff, J. and McQuilken, M. 2001. *The U. T. Touch-Sensitive Dance Floor and MIDI Controller*. http://www.utexas.edu/cofa/music/ems/research/midifloor/midifloor.html.
Risset, J-C. 1996. "Introductory Catalogue of Computer Synthesized Sounds" and "My 1969 Sound Catalogue: Looking back from 1992", *Historical CD of Digital Sound Synthesis (Computer Music Currents 13)*. Wergo, Germany.
Roads, C. 1996. *The Computer Music Tutorial*. MIT Press, Cambridge, MA.
Schaeffer, P. 1966. *Traité des Objets Musicaux*, Seuil, Paris, France.
Sethares, W. 1998. *Tuning, Timbre, Spectrum, Scale*. Springer, London.
Winkler, T. 2001. Composing Interactive Music. MIT Press, Cambridge, MA.
Xenakis, I. 1963. *Musiques Formelles: Nouveaux Principes Formels de Composition Musicale*. (original French edition), republished by Stock, Paris, 1981.
Xenakis, I. 1992. *Formalized Music* (updated from first English edition, 1971). Pendragon Press, Stuyvesant, NY.
Zakros InterArts, http://www.zakros.com/archives/Echoes.html.

Music References

Alves, B. 1999. "In-Yo" on *ICMC 99* conference disk, Tsinghua University, Beijing China.
Applebaum, M. 2002. "Precomposition" on *Music from SEAMUS*, Vol. 13, SEAMUS, Los Angeles, CA.
Ayers, L. 2006a. "Balinese" on *Virtual Gamelan*, Albany Records, Albany, NY, TROY874.
Ayers, L. 2006b. "Bioluminescense" on *Virtual Gamelan*, Albany Records, Albany, NY, TROY874.
Ayers, L. 2006c. "Merapi" on *Virtual Gamelan*, Albany Records, Albany, NY, TROY874.
Ayers, L. 2006d. "Pendopos" on *Virtual Gamelan*, Albany Records, Albany, NY, TROY874.
Barbour, J. 2004. "Outback", diffused at *Australasian Computer Music Conference*, Concert 7, Wellington, New Zealand.
Berio, L. 1999. "Visage" on *Passaggi/Visage*, Bmg Ricordi, B000005SSX.
Brown, C.C. 2000. "Under Erebus" on *New Zealand Sonic Art*, Vol. II (2001), The University of Waikato, Hamilton, New Zealand, UWMD1201.
Dodge, C. 1972. "He Destroyed Her Image" on *Any Resemblance Is Purely Coincidental*. New Albion, San Francisco, CA.
Helmuth, M. 1998. *Implements of Actuation*, Electronic Music Foundation, Albany, NY, EMF 023.
Hiller, L. and Isaacson, L. 1956. *The Illiac Suite for String Quartet*, excerpt in *The Voice of the IBM 7090 Computer*.

Jaffe, D. 2001. *Racing Against Time* performed at the 2004 International Computer Music Conference, Miami, Florida

Lansky, P. 1985. *Idle Chatter*. Wergo Compact Disc 2010–50, Bridge Records, 9050CD.

Lansky, P. 1992. "Table's Clear" on *Home Brew*, Bridge Records, BCD 9035.

Norris, M. 1998. "Aquarelle" on *New Zealand Sonic Art III* (2002), The University of Waikato, Hamilton, New Zealand.

Redolfi, M. 1990–91. *Nausicaa, Bande Originale de L'exposition*. Nausicaa, Boulogne-sur-Mer, France.

Rimmer, J. 2000. "Ancestral Voices" on *New Zealand Sonic Art*, Vol. II (2001), The University of Waikato, Hamilton, New Zealand, UWMD1201.

Risset, J-C. 1969. "Mutations" on *OHM: The Early Gurus of Electronic Music: 1948–1980*. Elipsis Arts, Roslyn, New York, CD3670.

Risset, J-C. 1996. *Invisible*. GMEM, Marseille, France.

Roads, C. 1981. "Field" on *Digital Rewind 25: Music for Instruments and Computer, Twenty-Fifth Anniversary of The MIT Experimental Music Studio, 1974–1999*. The Media Laboratory, Massachusettes Institute of Technology.

Rubin, A. and Hollander, L. 2004, "Family Stories: Sophie, Sally" on *Electric Music*, Capstone Records, Broolyn, NY, CPS-8743.

Schaeffer, P. 2000a. "Etude aux Chemins de Fer" on *Pierre Schaeffer: L'Oeuvre Musicale*, INA/EMF Media EM155–3.

Schaeffer, P. 2000b. "Etude aux Chemins de Fer" on *OHM: The Early Gurus of Electronic Music*, Ellipsis Arts, B00004T0FZ.

Sethares, W. 1997. *Xentonality*. Frog Peak Music. Lebanon, NH.

Tang, J.W.C. 1996. "Cympergon" on 1996 *International Computer Music Conference CD*, Hong Kong. October Music, Taiwan, 96081–10004.

Truax, B. 1989. "Beauty and the Beast" on *Inside*, Cambridge Street Records, Burnaby, BC, Canada.

Truax, B. 1995. "Inside" on *Inside*, Cambridge Street Records, Burnaby, BC, Canada.

Whalley, I. 2002. "Kasumi" on *New Zealand Sonic Art III* (2002), The University of Waikato, Hamilton, New Zealand.

Xenakis, I. 1964. "Eonta" on *Metastasis*, Chant du Monde, France.

Xenakis, I. 1991. "Gendy 3" on *Xenakis*. Harmonia Mundi, Los Angeles, CA; Neuma Records, Chelmsford, MA.

Xenakis, I. 1994. "S.709" on *Electronic Music Foundation*. EMF, Albany, NY.

Video References

Van Raalte, C. and Zane-Cheong, J. 1995. "The BodySynth" on *International Computer Music Association Video Research Reports*. V. II, The International Computer Music Association, San Francisco, California.

23

Singing Voice Analysis, Synthesis, and Modeling

Youngmoo E. Kim

Electrical & Computer Engineering, Drexel University, Philadelphia, PA, USA, e-mail: ykim@drexel.edu

1 • Introduction

The singing voice is the oldest musical instrument, but its versatility and emotional power are unmatched. Through the combination of music, lyrics, and expression, the voice is able to affect us in ways that no other instrument can. The fact that vocal music is prevalent in almost all cultures is indicative of its innate appeal to the human aesthetic. Singing also permeates most genres of music, attesting to the wide range of sounds the human voice is capable of producing. As listeners we are naturally drawn to the sound of the human voice, and, when present, it immediately becomes the focus of our attention.

To meet the demands of vocal communication, the vocal apparatus is extremely flexible, yet also remarkably self-consistent. An endless variety of sounds can be produced by a fairly simple physical system of vibration and resonance. Describing the distinctive character of a voice, however, is difficult without resorting to vague and subjective terms (e.g., "squeaky") that have no objective correlates. These qualities are believed to be a combination of innate physical characteristics and learned patterns of expression, but quantifying, extracting, and modeling these features have proven to be a difficult task. Likewise, the singing voice has also proven difficult to simulate

convincingly because of its greater physical variation compared to other instruments. In order to pronounce different words, a singer must move his or her jaw, tongue, teeth, etc., changing the shape and thus the acoustic properties of the vocal mechanism. This range of acoustic variation is difficult to capture in a low-dimensional model, and analysis and synthesis techniques which have been successful for other musical instruments often do not apply well to speech or singing.

Numerous models for analysis and synthesis of singing have been designed for a variety of applications. Some applications, such as singing detection and singer identification, focus only on analysis, while others, such as singing voice coding, require both analysis and synthesis components. This chapter explores the prevalent models used in modern singing voice analysis and synthesis and highlights the features and targeted applications of each.

Section 2 discusses background principles of singing voice modeling, such as vocal physiology and the source–filter formulation. Sections 3 and 4 examine multiple models used for the glottal source and vocal tract filter, respectively. This is followed by a description of various approaches to overall singing synthesis in Section 5.

2 • Principles of Singing Voice Modeling

2.1 The Anatomy of the Singing Voice

The anatomy of the voice consists of three primary collections of organs: the *lungs*, the *larynx*, and the *oropharynx*. Exhalation from the lungs results in the air pressure changes necessary for singing. The larynx consists of a skeleton of cartilage enclosing and supporting two structures of muscle and ligaments covered by mucous membranes known as the *vocal folds*. The oropharynx is more commonly known as the *vocal tract*. The key characteristic of the vocal tract is its ability to assume a wide range of different shapes, which are easily altered by articulating the positions of the jaw, tongue, and lips.

The process of singing begins with breath pressure produced by the lungs. In the case of *voiced* sounds (vowels, e.g., [a], [i], and semivowels, e.g., [m], [r]), muscles initially adduct (close) the vocal folds, but the breath pressure forces them to open. The airflow through the opening is uneven, with the air adjacent to the folds traveling a greater distance than the unimpeded air flowing through the opening. The result is a pressure differential, which causes the vocal folds to be sucked back together by the Bernoulli force. Rapid repetition of this process is called *phonation*, and the frequency of this repetition correlates to our

perception of the pitch. In voiced sounds, phonation results in a largely harmonic sound source. For *unvoiced* sounds (e.g., [f], [t]), the vocal folds remain open and the breath pressure results in free airflow through the larynx into the mouth, where it is impeded by a constriction (caused by the tongue, soft palate, teeth, or lips), resulting in a sound source caused by air turbulence. Some vocal sounds require both phonation and turbulence sound sources and are referred to as *mixed* sounds (e.g., [v], [z]). For a detailed description of voice anatomy and the process of singing, see [49].

In all three cases, the *source* (phonation, turbulence, or both) is modified by the shape of the vocal tract (throat, mouth, nose, tongue, teeth, and lips). Since the acoustic properties of an enclosed space follow directly from the shape of that space, the physical flexibility of the vocal tract lends itself to tremendous acoustic flexibility. From a signal processing standpoint, the vocal tract represents a time-varying acoustic *filter*, which colors an input signal. This description of the human voice is called the linear *source–filter* model (also known as the *excitation-resonance* model), which is the basis of the majority of voice models.

2.2 Speech Versus Singing

Historically, speech and singing research have been closely linked, but there are important differences between the two methods of vocal production. The vast majority of sounds generated during singing are voiced (approximately 90%), whereas speech contains a much larger percentage of unvoiced sounds (about 60% voiced and 40% unvoiced for English) [8]. In the most common classical singing technique, known as *bel canto*, singers are taught to sustain vowels as long as possible between other phonemes because they are the most efficient and audible sounds. Classical singers usually employ a technique in which they lower the larynx, creating an additional high-frequency resonance (around 2–4 kHz) not present in other types of vocal production. This resonance, known as the *singer's formant*, is especially important for being heard in the presence of other instruments, for example allowing an opera singer to be heard over an entire orchestra [49]. Because of these differences, speech models do not always translate directly into accurate models for singing.

2.3 The Signal View of the Singing Voice

The process of singing results in a time-varying acoustic signal that is periodic over short intervals. The short-term periodicity relates to the pitch, and the amplitude envelope roughly corresponds to the perceived dynamic level. The more slowly varying spectral features, such as the spectral envelope, contribute to the

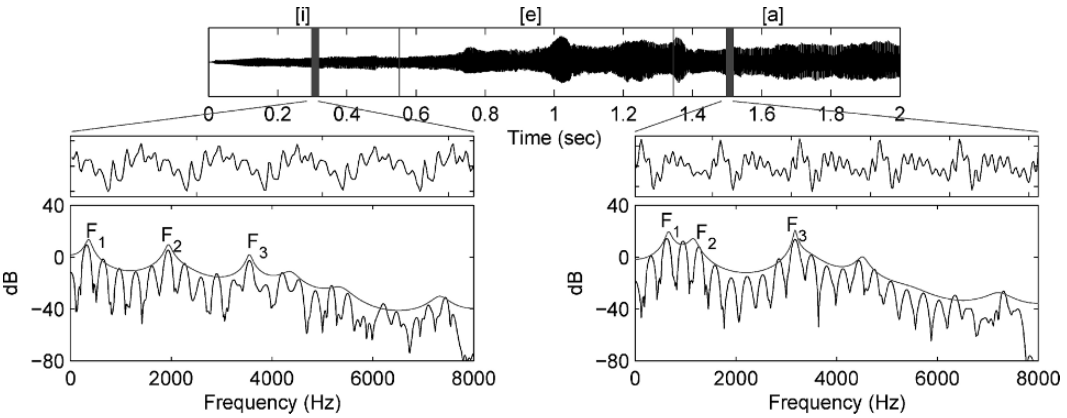


FIGURE 1 *Top*: time-domain signal of sung vowel sequence [i][e][a]. *Bottom*: magnified time-domain signals and short-time frequency analysis of phonemes [i] and [a]. The spectral envelope is depicted by the gray line, and the first three formants are labeled F_1, F_2, and F_3.

other components of singing (e.g., words, timbre, and musical expression). The relationship between spectral features and perceived expressive features is much less robust, but some general associations emerge. The spectral envelope contains several areas of increased energy (resonances) which may span across a range of harmonics. These resonances are called formants and have a great effect on vocal timbre [49]. The general position of the formants determines the phoneme of the singing. More subtle variations in formant frequency and bandwidth reflect the individual characteristics of the singer. An illustration of the singing signal in the time domain and frequency domain for two different phonemes is shown in Figure 1. Models of the singing voice have approached the problem from varying levels of abstraction: from physical models of vocal physiology to signal models of the acoustic waveform to symbolic or semantic models based on phonemes and musical notation. Several models are hybrids between different abstraction levels.

3 • Models of the Glottal Source

Because of the convenience of the linear source–filter decomposition, many computational representations of singing employ separate models for the source excitation signal. The vast majority of singing is voiced, and much of the work on excitation modeling has therefore focused on accurate modeling of the glottal flow. In voicing, the periodic but abrupt opening and

closing of the vocal folds produces a glottal flow waveform rich in harmonics. Spectrally, the glottal flow tends to fall off smoothly as frequency increases, and this slope is known as *spectral tilt*.

For convenience in representation, models often instead target the derivative of the glottal flow, which has a generally flat spectrum. This accommodation is often made because some vocal tract models can easily incorporate the spectral tilt parameter into their representation. The abrupt opening and closing of the folds leads to an impulsive quality in the derivative glottal waveform, a characteristic that has been used to great advantage in early source models. The glottal flow and its derivative have been shown to have significant correlation to an individual's vocal quality, and glottal models have been used in systems for individual voice identification [35].

3.1 Signal Models

The earliest computational voice model, Homer Dudley's *Vocoder* (a contraction of voice coder) in 1936 [9], used excitations of a periodic impulse train to approximate the derivative glottal waveform and white noise to approximate the air turbulence of unvoiced sounds. These excitations are still used in some modern very low bitrate speech codecs because of the small number of parameters needed. Reconstructed speech and singing using the impulse train excitation, however, tend to sound overly buzzy, and more recent work has moved toward more complex models.

Modern signal models for the glottal flow include the KLGLOTT88 model [18] and the transformed Liljencrants–Fant (LF) model [12], each of which model a single period of the derivative glottal waveform. The LF model has gained wide acceptance because of its accuracy while still employing a reasonably small number of parameters.

3.2 Physical Models of the Vocal Folds

Researchers have also explored physical models of the vocal folds in order to accurately synthesize glottal flow. Initial models used a fixed number of lumped masses and springs to represent the mechanics of the vocal folds. The first models [15] used two masses and springs per fold, but were limited in their flexibility. More recently, the 16-mass model proposed by Titze [50] with extensions by Kob [19] has found wider acceptance. These models are able to simulate wide-range vocal-fold mechanics, and the computational requirements have become less of a restriction on modern computers. Recent work by Hunter et al. [14] pursues the even more complex approach of a finite element model of the vocal folds. This model no longer assumes lumped masses and springs but uses a large number (1,721) of very small elements to calculate the geometry and movement of the folds. Real-time

calculation of this model, however, is still out of reach of today's computers.

These physical models can exhibit great authenticity when compared to the actual movement of the vocal folds and have been used to great effect in simulations, gaining traction in research on voice pathologies. Another benefit of these physical models is that they can be combined with physical models of the vocal tract in a nonlinear manner to model source–filter coupling and boundary effects. The greatest deficiency of physical models lies in the difficulty of parameterizing them especially from measured data. The mass values and spring constants in these models do not directly map to perceptually intuitive controls, though this is an area of continuing research [51].

4 • Models of the Vocal Tract

Dudley's Vocoder initially modeled the vocal tract response using a set of bandpass filters with amplitude controls to roughly approximate the spectral envelope of a voice signal. A modern extension of this principle is the *phase vocoder* (a short-time Fourier transform decomposition of the input signal), which has also found widespread use beyond voice coding in applications such as pitch- and time-scaling of audio [33]. Other modeling techniques have taken varied approaches to modeling the vocal tract, and some combine both source and filter elements to model the overall signal.

4.1 Formant Vocoder

A significant branch of vocoder research has focused on modeling the individual formants resulting from the shape of the vocal tract. The formant vocoder models each formant using a resonant filter (typically a second-order complex pole pair), resulting in a compact and efficient parameter set (see Figure 2b for an illustration). The formant filters are applied to an excitation signal to produce a resynthesized signal [39]. Difficulties are encountered, however, in the accurate estimation of the formant parameters (center frequency, amplitude, and bandwidth). In coding speech, the overall sound quality proved to be fairly low, and eventually the formant vocoder was eclipsed by other coding techniques that provided better sound quality at comparable bitrates.

The heightened role of formants in singing, however, has led to greater success for the formant vocoder in parametric singing synthesis. One of the earliest synthesizers designed specifically for singing voice is the Music and Singing Synthesis Equipment (MUSSE) developed at the Royal Institute of Technology (KTH) in Stockholm [21]. In this system, a simulated glottal pulse is sent

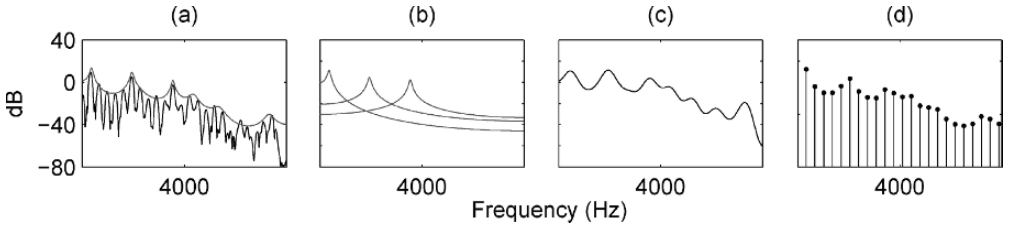


FIGURE 2 Graphical depiction in the frequency domain of several vocal tract models: (**a**) input signal, (**b**) formant vocoder, (**c**) cepstral model, and (**d**) sinusoidal model.

through a series of resonant filters to model the formants characteristic of a bass/baritone singer. Johan Sundberg and his group at KTH have determined general formant frequencies for all of the vowels and consonants. The most recent version of this synthesizer is MUSSE DIG, a digital implementation of the same structure that includes improvements to the glottal pulse model [4].

4.2 Homomorphic Vocoder

The homomorphic vocoder, based on an analysis by Alan Oppenheim [34], is an extension of the principles of the channel vocoder using homomorphic transformations to the *cepstral* (inverse log-Fourier) domain:

$$C_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |S(\omega)| \, \mathrm{e}^{\mathrm{j}\omega k} \mathrm{d}\omega, \tag{1}$$

where C_k are the cepstral coefficients and $S(\omega)$ is the Fourier transform of the voice signal $s(t)$. In this domain, what was multiplication in the frequency domain (and thus convolution in the time domain) becomes a simple addition through the nonlinearity of logarithms. The original voice signal (a convolution of the vocal excitation function and the vocal tract response function – assuming a linear model) is represented as an addition of excitation and source in the cepstral domain making estimation of the separated functions easier. This technique was implemented by Tom Stockham and Neil Miller for the separation of voice from orchestra and subsequent sound restoration in recordings of Enrico Caruso [32].

The cepstral representation and its variants (e.g., MFCCs – Mel-frequency cepstral coefficients) tend to characterize gross features of the spectrum, which makes them useful for tasks requiring generalization, such as speech recognition [38] and speaker (talker) identification [27]. Some of the individual distinctiveness in singing, however, is characterized by finer

spectral features that are not captured by low-order MFCCs. Figure 2c demonstrates how the cepstral representation results in a highly smoothed spectral envelope.

4.3 Sinusoidal Voice Coding

Sinusoidal analysis/synthesis for voice coding was first demonstrated by McAulay and Quatieri [29]. The voice signal is modeled as a number of sinusoidal partials:

$$s(t) = \sum_{k=1}^{M} A_k e^{j\omega_k t}. \quad (2)$$

For voiced sound segments the partials will be largely harmonic, while for unvoiced segments they will be inharmonic. Sinusoidal parameter estimation is typically performed using STFT analysis, and careful peak selection, tracking, and phase matching across analysis frames result in high sound quality at reduced bitrates. Mellody et al. [30] have proposed a sinusoidal analysis technique for singing using a high-resolution time–frequency distribution, called the *modal distribution*. By maintaining high resolution in both time and frequency, precise parameterizations of the individual harmonic partials of the voice can be obtained. The benefit of this analysis was demonstrated in an experiment in which individual soprano singers were automatically identified.

An extension to the sinusoidal representation using a parametric noise model was developed by Serra and Smith [43]. Since this model makes no assumptions about the input signal, it has also been applied to the synthesis of musical instruments [43] and the coding of general audio [36]. The sinusoidal decomposition also lends itself well to modifications, such as pitch- and time-scaling and has been demonstrated to be very effective in modifying speech and singing voice recordings [37] as well as arbitrary audio signals [20].

Primarily grounded in the frequency domain, the sinusoidal model benefits from the intuitions of traditional frequency domain signal processing as related to pitch and timbre. A deficiency of the sinusoidal representation has been the difficulty in relating sinusoidal parameters to the mechanics of voice production. Hybrid models have been built, however, incorporating aspects of the source–filter representation into the sinusoidal framework. At the simplest level, the sinusoidal amplitudes form a spectral envelope roughly corresponding to the vocal tract filter response, implying an excitation signal of equal-amplitude sinusoids. More sophisticated approaches have attempted to capture the amplitude and phase relationships implied by a true glottal excitation with a sinusoidal excitation

model. Researchers at the Music Technology Group of the Pompeu Fabra University (UPF) have used such a model, called EpR (excitation-plus-resonance), for singing synthesis and voice morphing in karaoke applications [5]. Recent work by McAulay [28] incorporating a sinusoidal excitation phase model has led to improvements in the quality of pitch, time scale, and spectral envelope modifications of voice signals.

4.4 Linear Predictive Coding

Perhaps the most influential advance in the history of voice coding after Dudley's work was the development of *linear predictive coding* (LPC) of speech, first proposed by Atal and Schroeder in 1970 [1]. LPC estimates a voice signal $s[n]$, using a linear combination of its p past samples. This is illustrated in the following difference equation relating the glottal source $g[n]$ and the voice output:

$$s[n] = \sum_{k=1}^{p} \alpha_k s[n-k] + g[n]. \tag{3}$$

Taking the z-transform of this difference equation, the transfer function $H(z)$ relating the voice output to the glottal source in the frequency domain can be derived:

$$H(z) = \frac{S(z)}{G(z)} = \frac{1}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}} = \frac{1}{A(z)}. \tag{4}$$

Equation (4) demonstrates that the transfer function is always an all-pole filter, and for convenience the denominator polynomial is often labeled separately as $A(z)$. This derivation further demonstrates the inherent source–filter formulation of the LPC model where the transfer function $H(z)$ filters the glottal source $G(z)$, resulting in the voice output $S(z)$. For this reason, $H(z)$ is defined as the vocal tract filter. To determine the filter coefficient values $alpha_k$ the error is minimized between the actual signal $s[n]$ and the predicted values from the previous p samples. The two most common techniques for solving this minimization are the autocorrelation and covariance methods, named for their use of the autocorrelation and covariance matrices, respectively [26], which both satisfy the minimum least squared error criterion. Vocal tract filter estimation via LPC also has the benefit of being a closed-form computation, requiring no heuristics for the determination of parameters. A disadvantage of standard LPC is that all frequencies are weighted equally on a linear scale while the frequency sensitivity of the human ear is close to logarithmic. The LPC coefficients, however, can also be found in a warped

frequency domain [13,45,48] that is more faithful to human perception.

When coupled with an appropriate excitation model, LPC can result in very low-bitrate transmission of speech. An extension, *code-excited linear prediction* (CELP), which is the basis of most speech codecs used in today's digital communications devices, employs a closed-loop analysis-by-synthesis to find the best fit within a codebook of excitation vectors. The codebook usually consists of an adaptive component from past excitations along with a fixed component allowing for further modification of the excitation.

By design, LPC finds the best fit of the vocal tract filter to an impulse excitation, which can be a poor rendition of the derivative glottal flow. Closed-phase LPC, using data only from the closed phase of the glottal flow (the time of least source–tract interaction) has been used to calculate glottal flow mode parameter estimates used for talker identification [35]. A recent extension to LPC modeling of singing has been joint optimization with a derivative glottal flow model, minimizing the error across a KLGLOTT88 model and LPC vocal tract filter, accounting for some of the inherent source–tract interaction [24]. After the joint optimization, the glottal flow is re-parameterized using the LF model, which is more intuitive for manipulation and synthesis. The joint source–filter optimization has been further enhanced to use warped LPC [17]. Synthesis quality using the jointly optimized system is improved using a parametric pulsed noise model in [24] and CELP-style residual codebooks in [17].

4.5 Physical Modeling of the Vocal Tract

Kelly and Lochbaum first modeled the vocal tract as a series of cylindrical tube sections represented by a digital ladder filter in 1962 [16]. This model of sound propagation has come to be known as a waveguide digital filter and has been used as the basis for a variety of musical instrument models [44]. Liljencrants first coupled the Kelly–Lochbaum model with the LF glottal excitation model in 1985 [22].

A *singing physical articulatory synthesis model* (SPASM) was created by Perry Cook at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University in 1990 [8]. SPASM used a digital waveguide to model the physical acoustics of the vocal tract as well as the nasal cavity and throat radiation, and the system was driven by a frequency domain excitation model. An integrated graphical composition system was provided to control the synthesis parameters. In 1994, Välimäki and Karjalainen extended the vocal tract physical model by using variable length conical sections [52]. A system using coupled

physical models of both vocal folds and vocal tract has been developed by Kob [19].

As with the physical models of the vocal folds, it is difficult to estimate model parameters from an audio signal, though several studies at KTH and the University of Arizona have extracted vocal tract area parameters from three-dimensional magnetic resonance images (MRI) [11,47]. Engwall [10] and Story [46] have developed 3D shape models of the vocal tract from measured MRI data. Articulation can be modeled in 3D, but synthesis is performed by calculating area functions from the 3D model. True 3D models, such as *finite element modeling*, are an active area of research.

4.6 PSOLA and WSOLA

Pitch-synchronous overlap-add (PSOLA) [7] and *waveform similarity overlap-add* (WSOLA) [53] are techniques that are commonly used in commercial products for voice processing and modification. Both techniques operate by identifying and manipulating time-domain regions of similarity (by estimating pitch or comparing waveforms). Deletion or replication of segments results in time compression or expansion, respectively. By overlapping and summing the windowed regions, these effects can be achieved while maintaining high sound quality. Pitch manipulation is also possible by altering sampling rates or changing the amount of overlap between windows. The use of devices based on PSOLA/WSOLA is almost ubiquitous in popular music recording, and they are also commonly used to provide pitch correction for live performances as well as karaoke machines. But the separation of parameters using this approach is somewhat limited, e.g., the pitch cannot be varied independently from the vocal tract shape.

4.7 Formant Wave Functions

Formant wave functions (FOFs, from the French) are time-domain functions with a particular resonance characteristic used to model individual formants [41]. Each function is repeated at the desired fundamental period of voicing, and the functions are then summed for synthesis. As with the formant vocoder, parameter estimation for the FOF functions is challenging because of the difficulty in identifying and characterizing formants. Because of this, FOFs have been used primarily for singing synthesis.

Singing voice synthesis with FOFs is implemented in the CHANT system, developed at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) by Xavier Rodet et al. [42]. It uses five formant wave functions to generate an approximation of the resonance spectrum of the first five formants of a singer. The results of the original system for vowel sounds are impressive

(including a famous rendition of Mozart's Queen of the Night aria from *Die Zauberflöte* [2]) but involve a great deal of hand adjustment of parameters. Later work has extended the use of FOFs to consonants and unvoiced phonemes [40]. The CHANT system was primarily designed as a tool for composers.

5 • Synthesis of the Singing Voice

Direct synthesis of singing, as opposed to resynthesis or encoding of an existing signal, adds even more challenges. Most languages can be represented in terms of a limited set of *phonemes* (basic linguistic units of sound), but the rules governing pronunciation and inflection are only guidelines to the actual speaking or singing of that language. The problem of interpreting a musical score in a musically proper and convincing fashion is difficult in the case of any instrument, and adding proper controls to manipulate the parameters of a singing voice model increases the difficulty of the problem.

5.1 Rules-Based Approaches

Several approaches have formulated rules for singing voice performance based on analytical and anecdotal data. The KTH Rule System [3], designed for the MUSSE DIG synthesizer, contains rules for numerous parameters such as pitch change glissando and timing, dynamics, vibrato, and formant transitions. A rule system was also used in the CHANT project to model formant trajectories. Rules-based systems have also been used to enhance other synthesis systems (see below).

5.2 Concatenation Synthesis

An approach favored in speech synthesis uses models of phonemes concatenated together to create words and phrases. The phonemes may be stored simply as short recordings which are smoothly concatenated and pitch-shifted to create a continuous pitch contour. Alternative models can also be used to generate the phonemes. One example is the LYRICOS project [25], which uses a set of phonemes stored as sinusoidal model parameters, incorporated with control data (such as pitch, dynamics, and vibrato) from MIDI. But single-phoneme models generally do not have the flexibility to accommodate the wide variation present in speech or singing, and without dynamic control manipulation, the synthesis is rather limited.

Multi-phoneme models have been used with a greater amount of success. A recent example at UPF uses their EpR sinusoids-plus-noise representation to model a large database of phoneme combinations from an individual singer at a variety of pitches [5]. A rule system was used to enhance the overall quality of the

phoneme transitions [6]. This work has been incorporated into a commercial product, the Yamaha Vocaloid [54], a singing voice synthesizer that makes use of "voice libraries" designed to model an individual singer's voice.

5.3 Data-Driven Approaches

Systems have also been designed to use statistical methods to model parameter or phoneme transitions for singing voice models. One project, by Ken Lomax at Oxford University, trained neural networks on input data from famous singers to control the necessary synthesis parameters [23]. This system was limited to mostly vowels, with a few voiced consonants. Yoram Meron at the University of Tokyo used *hidden Markov models* (HMMs) to control a mapping from sinusoidal phoneme models of an individual singer to a target sound, which could be from the same singer or a different singer [31]. This enabled a voice transformation in which one singer's voice model (stored phonemes) is driven by another (target sound). A recent system by Kim [17] uses HMMs to model joint source–filter (LF and warped LPC) model dynamic parameter motion within single phonemes to analyze an individual singer's time-varying expression. This system allows the source–filter model characteristic of one singer to be driven by the expressive characteristics of another, achieving a form of separation of the instrument from the instrumentalist that is not possible with real singers.

6 • Conclusion

This chapter has provided a brief overview of some of the myriad models and techniques used in applications of analysis and synthesis of singing. Some of the signal processing models have been in existence for quite some time, and their utility and robustness are well documented. Appropriate model representations and mappings from control to model parameters remain an active area of research. Dynamic modeling of expressive performance parameters is a relatively new endeavor, but new methods in pattern recognition and machine learning along with the continuing advance of computational power hold great promise for the future.

References

1. B. S. Atal and M. R. Schroeder. Adaptive predictive coding of speech signals. *Bell System Tech. J.*, pp. 1973–1986, 1970.
2. G. Bennett and X. Rodet. Synthesis of the Singing Voice, CD of Sound Examples. *Current Directions in Computer Music Research*, pp. 19–44, MIT Press, Cambridge, MA, 1989.

3. G. Berndtsson. The KTH rule system for singing synthesis. *Computer Music J.*, 20(1):76–91, 1996.
4. G. Berndtsson and J. Sundberg. The MUSSE DIG singing synthesis. *Proc. Stockholm Music Acoustics Conf. (SMAC)*, pp. 279–281, 1993.
5. J. Bonada and A. Loscos. Sample-based singing voice synthesizer by spectral concatenation. *Proc. Stockholm Music Acoustics Conf. 2003*, pp. 439–442, 2003.
6. J. Bonada, O. Celma, A. Loscos, J. Ortolà, and X. Serra. Singing voice synthesis combining excitation plus resonance and sinusoidal plus residual models. *Proc. Int. Computer Music Conf.,* Havana, pp. 139–146, 2001.
7. F. J. Charpentier and M. G. Stella. Diphone synthesis using an overlap-add technique for speech waveforms concatenation. *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Tokyo, pp. 2015–2018, 1986.
8. P. R. Cook. *Identification of Control Parameters in an Articulatory Vocal Tract Model, with Applications to the Synthesis of Singing*. PhD thesis, Stanford University, 1990.
9. H. Dudley. Synthesizing speech. *Bell Laboratories Record*, 15: 98–102, 1936.
10. O. Engwall. Synthesising static vowels and dynamic sounds using a 3d vocal tract model. *Proc. Fourth ISCA Tutorial and Research Workshop on Speech Synthesis*, Perthshire, Scotland, pp. 38–41, 2001.
11. O. Engwall and P. Badin. Collecting and analysing two- and three-dimensional MRI data for Swedish. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)*, 3–4:11–38, 1999.
12. G. Fant, J. Liljencrants, and Q.-G. Lin. A four parameter model of glottal flow. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)*, 4:1–13, 1985.
13. A. Härmä. A comparison of warped and conventional linear predictive coding. *IEEE Trans. on Speech and Audio Processing*, 9(5):579–588, 2001.
14. E. J. Hunter, I. R. Titze, and F. Alipour. A three-dimensional model of vocal fold abduction/adduction. *J. Acoust. Soc. Am.*, 115(4): 1747–1759, 2004.
15. K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell System Tech. J.*, 51: 1233–1268, 1972.
16. J. L. Kelly and C. C. Lochbaum. Speech synthesis. *Proc. Fourth Int. Congress on Acoustics*, pp. 1–4, 1962.
17. Y. E. Kim. *Singing Voice Analysis/Synthesis*. PhD thesis, Massachusetts Institute of Technology, 2003.
18. D. H. Klatt and L. C. Klatt. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.*, 87(2):820–857, 1990.
19. M. Kob. *Physical Modeling of the Singing Voice*. PhD thesis, University of Technology Aachen, 2002.

20. J. Laroche, Time and Pitch Scale Modification of Audio Signals. *Applications of Signal Processing to Audio and Acoustics*. Kluwer Academic Publishers, Dordrecht, 1998.
21. B. Larsson. Music and Singing Synthesis Equipment (MUSSE). *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)*, 1:38–40, 1977.
22. J. Liljencrants. *Speech Synthesis with a Reflection-Type Line Analog*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 1985.
23. K. Lomax. *The Analysis and Synthesis of the Singing Voice*. PhD thesis, Oxford University, 1997.
24. H.-L. Lu. *Toward a High-Quality Singing Synthesizer with Vocal Texture Control*. PhD thesis, Stanford University, 2002.
25. M. W. Macon, L. Jensen-Link, J. Oliverio, M. A. Clements, and E. B. George. Concatenation-based midi-to-singing voice synthesis. *Audio Eng. Soc. Preprint 4591*, 1997.
26. J. Makhoul. Linear prediction: A tutorial review. *Proceedings of the IEEE*, 63:1973–1986, 1975.
27. R. Mammone, X. Zhang, and R. P. Ramachandran. Robust speaker recognition: A feature-based approach. *IEEE Signal Processing Magazine*, 13(5):58–71, 1996.
28. R. J. McAulay. Voice Modification using the Sinusoidal Analysis/ Synthesis Method, *Keynote Presentation at 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, October, 2003.
29. R. J. McAulay and T. Quatieri. Speech analysis/synthesis based on a sinusoidal representation. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 34(4):744–754, 1986.
30. M. Mellody. G. H. Wakefield, and F. Herseth, *Model distribution analysis and Synthesis of a Soprano's Sung vowels. J. Voice*, 15(4): 469–482, 2001.
31. Y. Meron. *High Quality Singing Synthesis Using the Selection-Based Synthesis Scheme*. PhD thesis, University of Tokyo, 1999.
32. N. J. Miller. *Filtering of Singing Voice Signal from Noise by Synthesis*. PhD thesis, University of Utah, 1973.
33. J. A. Moorer. The use of the phase vocoder in computer music applications. *J. Audio Eng. Soc.*, 26(1):42–45, 1978.
34. A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1989.
35. M. D. Plumpe, T. F. Quatieri, and D. A. Reynolds. Modeling of the glottal flow derivative waveform with application to speaker identification. *IEEE Trans. Speech and Audio Processing*, 7(5):569–586, September 1999.
36. H. Purnhagen and N. Meine. HILN – the MPEG-4 parametric audio coding tools. *Proc. IEEE Int. Symposium on Circuits and Systems* Geneva, Switzerland, pp. 201–204, 2000.
37. T. F. Quatieri and R. J. McAulay. Speech transformations based on a sinusoidal representation. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 34(6):1449–1464, 1986.

38. L. R. Rabiner and B.-H. Juang. *Fundamentals of Speech Recognition*. Prentice Hall, Englewood Cliffs, NJ, 1993.
39. L. R. Rabiner and R. W. Schafer. *Digital Processing of Speech Signals*. Prentice-Hall, Englewood Cliffs, NJ, 1978.
40. G. Richard, C. d'Alessandro, and S. Grau. Unvoiced speech synthesis using poissonian random formant wave functions. *Signal Processing VI: European Signal Processing Conf.*, pp. 347–350, 1992.
41. X. Rodet. Time-domain formant-wave-function synthesis. *Computer Music J.*, 8(3):9–14, 1984.
42. X. Rodet, Y. Potard, and J.-B. Barrière. The CHANT project: From the synthesis of the singing voice to synthesis in general. *Computer Music J.*, 8(3):15–31, 1984.
43. X. Serra and J. O. Smith. Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music J.*, 14(4):12–24, 1990.
44. J. O. Smith. *Techniques for Digital Filter Design and System Identification with Application to the Violin*. PhD thesis, Stanford University, 1985.
45. K. Steiglitz. A note on variable recursive digital filters. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 28(1):111–112, February 1980.
46. B. H. Story. Using imaging and modeling techniques to understand the relation between vocal tract shape to acoustic characteristics. *Proc. Stockholm Music Acoustics Conf. 2003*, pp. 435–438, 2003.
47. B. H. Story, I. R. Titze, and E. A. Hoffman. The relationship of vocal tract shape to three voice qualities. *J. Acoustical Soc. Am.*, 109(4):1651–1667, 2001.
48. H. W. Strube. Linear prediction on a warped frequency scale. *J. Acoust. Soc. Am.*, 68(4):1071–1076, 1980.
49. J. Sundberg. *The Science of the Singing Voice*. Northern Illinois University Press, Dekalb, IL, 1987.
50. I. R. Titze. Phonetica. *The Human Vocal Cords: A Mathematical Model. Part 1*, 28:129–170, 1973.
51. I. R. Titze and B. H. Story. Voice quality: What is most characteristic about "you" in speech. *Echoes: The Newsletter of The Acoustical Society of America*, 12(4):1, 4, Fall 2002.
52. V. Välimäki and M. Karjalainen. Improving the Kelly–Lochbaum vocal tract model using conical tube sections and fractional delay filtering techniques. *Proc. 1994 Int. Conf. on Spoken Language Processing (ICSLP)*, Yokohama, Japan, pp. 615–618, 1994.
53. W. Verhelst and M. Roelands. An overlap-add technique based on waveform similarity (WSOLA) for high-quality time-scale modifications of speech. *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing*, Minneapolis, Vol. 2, pp. 554–557, 1993.
54. Yamaha Corporation Advanced System Development Center. *New Yamaha VOCALOID Singing Synthesis Software Generates Superb Vocals on a PC*, 2003. http://www.global.yamaha.com/news/20030304b.html.

24

Instrument Modeling and Synthesis

Andrew B. Horner[1] *and James W. Beauchamp*[2]

[1] *Department of Computer Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, e-mail: horner@cs.ust.hk*
[2] *School of Music and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, e-mail: jwbeauch@uiuc.edu*

1 • Introduction

During the 1970s and 1980s, before synthesizers based on direct sampling of musical sounds became popular, replicating musical instruments using frequency modulation (FM) or wavetable synthesis was one of the "holy grails" of music synthesis. Synthesizers such as the Yamaha DX7 allowed users great flexibility in mixing and matching sounds, but were notoriously difficult to coerce into producing sounds like those of a given instrument. Instrument design wizards practiced the mysteries of FM instrument design. Wavetable synthesis was less mysterious, but was limited to rather static organ-like sounds. With a single wavetable, you could easily attain a trumpet-like sound, but a realistic trumpet was still out of reach. Then, sampling came along and soon even cheap synthesizers could sound like realistic pianos and trumpets, as long as the desired articulations happened to match those of the original recorded samples. Sample libraries quickly replaced sound wizards.

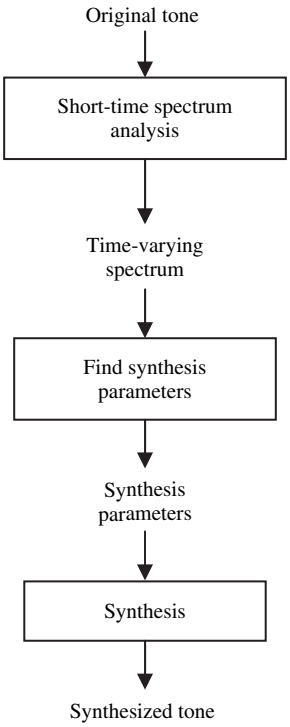


FIGURE 1 Overview of the analysis/matching/resynthesis procedure.

Ironically, at about the time that FM declined, researchers developed algorithms to optimize FM and wavetable parameters. The results were every bit as realistic as sampled sounds, with the added benefits of increased spectral and temporal control.

The basic procedure for FM and wavetable matching is shown in Figure 1. Matching generally begins with a short-time (aka time-varying) spectrum analysis of the original sound. Typically, 500–2,000 consecutive "spectral snapshots" or frames of the tone result from this analysis, each consisting of the instantaneous amplitudes and frequencies of the harmonics of the tone's time-varying fundamental frequency ($f_0(t)$). The next step is to find parameters for the synthesis algorithm that provide the average "best fit" of the spectrum of the tone over the set of these snapshots. The final step is to resynthesize the sound using the matched parameters.

The next section describes the short-time spectral analysis methods the authors have used, the first step of Figure 1. Then, we describe the fundamentals of wavetable synthesis, followed by an overview of wavetable matching procedures. Next, we describe

FM synthesis and FM matching. Finally, we conclude with some practical guidelines for selecting among these methods.

2 • Spectral Analysis

Spectral analysis methods used for wavetable and FM matching synthesis are of necessity pitch-synchronous and are described by Beauchamp (2007). The most often used method is a variation on the *phase vocoder* (Flanagan and Golden 1966; Portnoff 1976). With our method, a digital single-voice signal, which is assumed to have a nearly constant fundamental frequency, is broken into overlapping segments, called windows or frames. For analysis purposes, the fundamental is chosen to be f_a, an approximation to the actual time-varying fundamental frequency $f_0(t)$. The window lengths are taken to be twice the period corresponding to the fundamental (i.e., $T = 2/f_a$), and the overlap time or time-between-frames is taken to be a half period (i.e., $\Delta T = 0.5/f_a = 0.25T$). The windowed signal segments are multiplied by a Hamming window function ($w(t) = 0.5 + 0.426\cos(2\pi t/T)$) before applying a fast Fourier transform (FFT) to the signal samples. However, the signal must generally be upsampled to provide an even power-of-two samples per window in order to satisfy the requirements of our version of the FFT. If the original sample rate is f_s, the new sample rate is $f_s' = 2^{\text{ceil}(\log_2(Tf_s))}/T$. After the FFT is taken, only the even components are retained, with the second component corresponding to the first harmonic, the fourth component corresponding to the second harmonic, and so forth. For each spectral snapshot or frame j, the FFT returns the real and imaginary parts of the analysis, and these must be converted to harmonic amplitudes and frequencies. For each harmonic k, the estimated amplitude $A_{k,j}$ is computed as the square root of the sum of the squares of the real and imaginary parts (i.e., the complex magnitude), whereas the estimated instantaneous frequency $f_{k,j}$ is computed using an arctangent formula based on the real and imaginary parts for two successive frames (Beauchamp 2007). The number of harmonics K analyzed and kept is given by the Nyquist frequency divided by f_a; thus, $K = \text{floor}(0.5f_s/f_a)$. The total number of data points kept is twice the number of samples in the sound signal analyzed.

An alternative analysis method is based on frequency tracking of spectral components (McAulay and Quatieri 1986; Smith and Serra 1987). This method works for signals whose fundamental frequency variation exceeds a small percentage of its mean value. A fixed window of 2^N sample length is chosen to be greater than three times the period of the lowest expected frequency.

After a Kaiser-windowed FFT processes the signal, the magnitude spectrum is computed, and spectral peaks above a specified amplitude threshold are retained. Spectral peaks are connected from frame to frame using a nearest neighbor algorithm to form tracks. Next, fundamental frequency vs. time is computed using a "two-way mismatch algorithm" (Beauchamp et al. 1993; Maher and Beauchamp 1994). Finally, harmonics are separated from the total spectrum based on the computed time-varying f_0. For tones with little frequency variation, the result is very similar to that obtained using the phase vocoder, but the frequency-tracking method works much better for sounds with significant frequency deviation.

Although in general these methods result in harmonics which do not track each other perfectly (i.e., $f_{k,j}/k_1 \neq f_{k,j}/k_2$ for $k_1 \neq k_2$), the methods described in this chapter require a common fundamental frequency such that $f_{k,j} = kf_{0,j}$. For the normally used phase-vocoder method, Calculation of $f_{0,j}$ is described by Beauchamp (2007). The result is a signal model given by

$$s[n] = \sum_{k=1}^{K} A_k[n] \cos(k \int f_0[n] \mathrm{d}n/f_\mathrm{s} + \phi_k), \tag{1}$$

where $n =$ sample number, $k =$ harmonic number, and $A_k[n]$ and $f_\mathrm{o}[n]$ are given as sample-time harmonic amplitude and fundamental frequency values interpolated from neighboring frame-time values $A_{k,j}$ and $f_{o,j}$ obtained from the spectral analysis. f_s is the sampling frequency in Hz.

3 • Wavetable Synthesis

The music industry currently uses the term *wavetable synthesis* synonymously with *sampling synthesis*. However, in this article, *sampling* means recording an entire note and playing it back, whereas *wavetable synthesis* refers to storing only one period of a waveform in an oscillator table and synthesizing a signal by applying amplitude- and frequency-vs.-time envelopes (Mathews 1969). The waveform is generated by a sum of harmonically related sine waves, whose amplitudes and phases define the wavetable's spectrum. In our methods of wavetable synthesis, several wavetables, each corresponding to a *basis spectrum*, are scaled by time-varying amplitude envelopes (or weight functions) and summed to form the output signal. Figure 2 shows a block diagram for this *multiple-wavetable synthesis* model.

Prior to synthesis, one cycle of each basis waveform is stored in a wavetable. The waveform is generated by a sum of harmonically

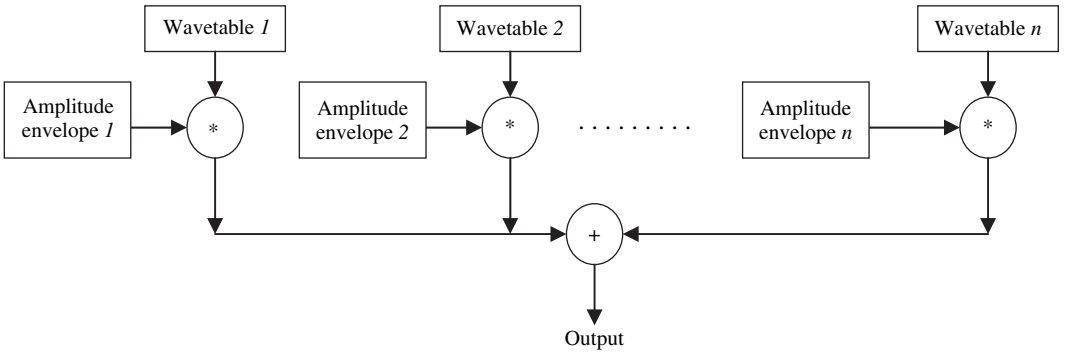


FIGURE 2 A multiple wavetable synthesis block diagram.

related sine waves, i.e., for the mth waveform or table, the table entries are given by

$$\text{table}_m[i] = \sum_{k=1}^{K} A_{k,m} \cos\left(\frac{2\pi k i}{\text{tableLength}} + \phi_{k,m}\right), \quad \text{for } 0 \le i < \text{tableLength}, \tag{2}$$

where $\text{table}_m[i]$ is the ith entry of the mth wavetable, tableLength is the number of entries in the table, and $A_{k,m}$ and $\phi_{k,m}$ are the amplitude and phase of the kth harmonic for the table. The A_ks corresponding to each table form the wavetable's basis spectrum. The phases generally do not usually affect the sound and are often set to zero, but in any case, for a given harmonic, they should be the same for all tables in order to avoid phase cancellations. Then, for M equal-length wavetables, the output signal to be matched with the original of Eq. (1) is given by

$$\hat{s}[n] = \sum_{m=1}^{M} \alpha_m[n] table_m\left[\int f_o[n]\mathrm{d}n/(f_s \text{tableLength})\right], \tag{3}$$

where $\alpha_m[n]$ is the amplitude envelope of the mth wavetable and $f_o[n]$ is the frequency envelope common to all wavetables. Although table entries are specified to be integers in Eq. (2), we assume here that the table values are interpolated according to the integer and fraction values of its argument and that its argument is taken modulo tableLength.

The main advantage of multiple-wavetable synthesis is its efficiency for generating periodic waveforms. A possible disadvantage of this method is that each wavetable only produces

a static spectrum, while real sounds produce continuously changing spectra. However, if we carefully select basis spectra and control amplitudes, we can fully reproduce the dynamic spectra of the original tone in wavetable synthesis.

Another possible disadvantage is that it is necessary that harmonics be strictly locked together in frequency so that $f_k[n] = kf_1[n]$. Since phase is the integral of frequency, this implies that relative phases are also locked together. Frequency unlocking (inharmonicity) or phase unlocking certainly occurs in freely vibrating tones, such as struck or plucked string tones (Fletcher et al. 1962), and there is some evidence that phase unlocking (or aperiodicity) also occurs momentarily in sustained tones, especially during vibrato and transients (Beauchamp 1974; Schumacher 1992; Brown 1996; Dubnov and Rodet 2003). Extension of the multiple-wavetable synthesis method for approximately handling inharmonic cases with slightly mistuned wavetables has been explored for the piano (Lee and Horner 1999; Zheng and Beauchamp 1999). Bristow-Johnson (1996) has proposed a wavetable method for changing relative phases as time progresses, which is tantamount to synthesizing unlocked phases. However, the perceptual advantage of unlocked phases in sustained tones has not been fully explored. In any case, synthesis with unlocked phases is beyond the scope of this chapter.

There are two basic types of multiple-wavetable synthesis. *Wavetable interpolation synthesis* is a method whereby tables are interpolated from one (or possibly a few) to the next in succession, usually using straight-line interpolation (Horner and Beauchamp 1996). With *wavetable-index synthesis* all tables are active simultaneously, and the amplitude envelope weights applied to the tables are computed using *least squares* (Horner et al. 1993a). In this case, the amplitude envelopes tend to be quite complex with no restriction on their polarities. Thus, when basis spectra are combined (assuming they have identical phases), their weights may cause them to cancel each other, and, in any case their combination will converge to the original time-varying spectrum as more tables are added.

Basis spectra can be taken directly from the original time-varying spectrum at certain points in time, from spectral averages using a clustering method, or by using a method such as principal components analysis (PCA) to compute basis spectra which decrease in importance as more are added. Optimal matching algorithms, such as genetic algorithms (GA) (Goldberg 1989), can be used to search for the "best" set of basis spectra.

4 • Wavetable Matching

The objective of wavetable matching is to find the best set of parameters to synthesize a musical instrument tone using wavetable synthesis. A number of papers have explored methods for optimizing wavetable basis spectra and their amplitude-vs.-time envelopes (Horner et al. 1993a; Sandell and Martens 1995; Horner and Beauchamp 1996; Horner 2001). Usually a *cost metric* is used to determine the efficacy of the match and to steer the matching process. One metric we have found successful, called the *relative amplitude spectral error*, appears to correspond closely to perceptual similarity (Horner et al. 2006); it is given by

$$\varepsilon = \frac{\sum_{j=0}^{J-1} \sum_{k=1}^{K} \left| A_{k,j} - A'_{k,j} \right|^p}{\sum_{j=0}^{J-1} \sum_{k=1}^{K} \left| A_{k,j} \right|^p}, \tag{4}$$

where J is the number of analysis frames used for the average. The error ε can be an average over all of the frames of the tone or a suitable subset of the frames. The power p can be either 1 or 2, and good results are obtained for either case.

Maher and Beauchamp (1990) used a wavetable matching synthesis method in their investigation of vocal vibrato. They selected their basis spectra at the low and high points of the vibrato of sung tones and crossfaded the wavetables as a function of the vibrato. Since the spectra at these points virtually repeated in cyclic fashion during the tone, only two wavetables were required for the synthesis.

With wavetable interpolation synthesis (Chamberlin 1980; Serra et al. 1990; Horner and Beauchamp 1996), the signal is represented by a consecutive series of basis spectra, which are generally spectral snapshots take from the tone itself at particular points in time or *breakpoints*, and synthesis proceeds by gradually crossfading from one spectrum (or waveform) to the next. While one basis spectrum ramps down to zero, a new basis spectrum ramps up to take its place.

Serra et al. (1990) give two algorithms for determining spectral interpolation basis spectra. The first draws basis spectra from the original tone. The second uses a least-squares algorithm to compute the basis spectra which minimize the mean-squared error between the original and synthetic spectra. Basis spectra are added until the maximum mean-squared error is brought below a prescribed threshold. Interpolation between basis spectra is done using a linear or nonlinear method. These algorithms cycle through basis spectra between 5 and 20 times per second.

Rather than first specifying a error threshold, Horner and Beauchamp (1996) used a genetic algorithm and a sequential enumeration (aka greedy) method to select a predetermined number of best times (breakpoints) to take basis spectra from the signal. Several questions were explored. It was found that the GA method was slightly better than the greedy method in terms of relative error vs. number of breakpoints, but it was substantially better than random breakpoints or equally spaced breakpoints. While breakpoints optimized independently for each harmonic performed somewhat better than common breakpoints in terms of error vs. the number of breakpoints per harmonic, common breakpoints won in terms of error vs. data storage requirements. Also, independent harmonic breakpoints would not allow wavetable interpolation. Another method tried used quadratic rather than linear curves to interpolate between basis spectra. Like the independent harmonic breakpoint attempt, this resulted in an improvement for the same number of breakpoints, but when total data and computation requirements were taken into account, improvement was negligible or nil.

Genetic algorithms have also been used to optimize wavetable interpolation with multiple (more than two at a time) wavetables (Horner 1996b). In this case, each breakpoint spectrum is the weighted sum of two or more spectra (wavetables) which are crossfaded to the same number of tables at the next breakpoint.

Group additive synthesis (Kleczkowski 1989) is another wavetable variant at the opposite extreme of spectral interpolation. Group additive synthesis divides the spectrum into non-overlapping subsets of harmonics for the various wavetables. As an example, one wavetable might contain only the even harmonics while a second only the odd. Subsequent to Kleczkowski's initial study, researchers have optimized group additive synthesis parameters using an automated clustering scheme (Oates and Eaglestone 1997) and genetic algorithms (Cheung and Horner 1996; Horner and Ayers 1998; Lee and Horner 1999).

For wavetable-index synthesis, researchers have used genetic algorithms to match the multiple-wavetable model shown in Figure 2 (Horner et al. 1993a; Horner 1995). Like the interpolation synthesis method, one approach is to use a genetic algorithm to select the best time values (breakpoints) for spectral snapshots to be taken from the original tone as the basis spectra (Horner et al. 1993a). However, unlike the interpolation method, a least-squares algorithm is used to compute the time-varying basis spectra amplitude weights which minimize the squared error between the original and synthetic spectra. This method

generates an exact match at the time points of the selected snapshots and usually results in excellent matches at neighboring points as well. The relative amplitude spectral error (Eq. 4) between the original and matched spectra typically serves as a fitness function to guide the GA's search for the best solution. Most matched instruments require 3–5 wavetables for a good match (less than 5% relative amplitude spectral error), a considerable savings compared to sinusoidal additive synthesis or interpolation synthesis. However, in terms of total computation, interpolation synthesis can be just as efficient as wavetable-index synthesis and is more intuitive in terms of the amplitude weight envelopes. This is because the amplitude weights for wavetable-index synthesis can be either positive or negative and thus the sum of the amplitude-controlled basis spectra only *converges* to the least-error values.

Another approach for selecting basis spectra is based on principal components analysis (PCA). PCA is a statistical procedure for optimizing both weights and basis vectors (Dunteman 1989) and has been applied to the wavetable matching problem (Zahorian and Rothenberg 1981; Sandell and Martens 1992; Horner et al. 1992, 1993a; Sandell and Martens 1995). In our case, PCA first finds a basis spectrum and time-varying amplitude control to minimize the average mean-squared spectral error (rather than the relative amplitude spectral error). It then finds an amplitude and control to further minimize the error. It usually converges to a good approximation with five basis spectra (wavetables). If the number of PCA basis spectra equals the number of harmonics, convergence is perfect. Also, an alternative way to compute the control functions, which we actually use, is to first compute the basis spectra and then employ the same method we used for wavetable-index synthesis, which is the least-squares solution. However, there are a couple of disadvantages to the PCA method in comparison to using spectral snapshots: First, in selecting basis spectra, PCA strongly weighs the high-amplitude steady-state spectra, which is unfortunate because the low-amplitude attack and release spectra are perceptually very important. However, this can be overcome by judicious use of spectra used in computing the basis spectra. Second, except for the first basis spectrum, the PCA basis spectra do not resemble real spectra because they are used to correct the first one (in analogous fashion to Fourier sine waves correcting each other to form a square wave) and are therefore less intuitive to work with than spectral snapshots. Nevertheless, PCA provides an interesting alternative approach to basis spectra generation.

Another method which uses basis spectra and wavetables is *cluster synthesis*. With this method snapshot spectra are sorted into clusters according to some measure of the spectra. One method is to sort the spectra according to their spectral centroids (Beauchamp and Horner 1995; Horner 1995). The spectra within each cluster are amplitude-normalized and then averaged to produce a basis spectrum. For synthesis the control functions are amplitude, centroid, and f_0 vs. time. For each time frame, the centroid value is used to interpolate between two basis spectra whose centroids straddle this value to get an output spectrum which is then synthesized at the appropriate amplitude and f_0. Like wavetable interpolation, it can be shown that interpolation in the frequency domain coupled with additive synthesis is equivalent to interpolation in the time domain with basis-spectra-loaded wavetables.

An alternative method of cluster synthesis comes from the technique of vector quantization (Ehmann and Beauchamp 2002). Here the spectra are clustered using a K-means algorithm (Rabiner and Juang 1993). The basic synthesis method is similar to the centroid-clustering method above. However, with vector quantization, each cluster is represented by an average spectrum which is labeled by a "code-book number" or index rather than by its average spectral centroid. So the basis spectra must be retrieved by index-vs.-time data. For synthesis, the basis spectra can be ordered according to their spectral centroids or corresponding amplitudes in the original signal. Then for each synthesis frame, the best rms basis spectrum match is found, and the index-vs.-time data are generated. To avoid discontinuities, the index-vs.-time data are smoothed, and then interpolation is used to ensure smooth transitions between the basis spectra. With one version of the mehod, the control parameters are time-varying amplitude, index, and f_0. Alternatively, the time-varying spectrum is first normalized in terms of both amplitude and spectral tilt, where tilt is the slope p of a straight line that forms the best least-squares fit to the spectrum plotted as log-amplitude-vs.-log-frequency. The K-means algorithm is then used to cluster the resulting flattened spectra. Synthesis parameters now consist of amplitude, tilt, index, and f_0 vs. time, and synthesis proceeds in a fashion described above, except that now both amplitude and spectral tilt have to be imposed on the retrieved spectra before synthesis. Therefore, synthesis has to be done in the frequency domain before conversion to signal. However, if tilt normalization is omitted, wavetables can be used as in the centroid-clustering case.

Wavetable synthesis is an inherently harmonic synthesis method, so replicating sounds that are nearly harmonic, such as piano and plucked string tones, require some enhancements of the method. By grouping partials with similar frequency stretch factors (Fletcher et al. 1962), genetic algorithms have successfully optimized group additive synthesis parameters to simulate piano tones (Lee and Horner 1999; Zheng and Beauchamp 1999) and string tones (So and Horner 2002).

Sometimes using a complex method can lead to a simpler one. We have found that insights gained from exploring a problem first with the GA method often led to finding better or simpler solutions. Wavetable matching is such an example. Instead of using the GA to approximate the best match to all or a subset of the spectral snapshots of the original tone, an alternative method is to use a combinatorial method to find the best match for a subset of the spectral snapshots (Horner 2001). It turns out that out that this approach is as effective and efficient as the GA method and much simpler.

5 • Frequency Modulation (FM) Synthesis Models

Like wavetable synthesis, FM synthesis can efficiently generate interesting sounds. There are several possible FM configurations, including those with multiple parallel modulators, nested (serial) modulators and feedback (see Figure 3). Several of these "modules" can be combined to form complex FM synthesizers. For example, as with multiple-wavetable synthesis, several carrier oscillators can be combined in parallel to form a single-modulator/multiple-carrier FM synthesizer (see Figure 4). During the height of FM's popularity in the 1980s, synthesizers such as the Yamaha DX7 allowed users great flexibility in mixing and matching models like these.

5.1 Single-Modulator/ Single-Carrier Model

The original FM equation John Chowning used for music synthesis consisted of a single sine wave modulating a carrier sine wave in a vibrato-like fashion (Chowning 1973):

$$A\sin(2\pi)\int (f_c + \Delta f_m \cos(2\pi f_m t))dt = A\sin(2\pi f_c t + \alpha_m \sin(2\pi f_m t)), \quad (5)$$

where A is the amplitude, f_c is the carrier frequency, Δf_m is the modulator amplitude, f_m is the modulation frequency, and $\alpha_m = \Delta f_m / f_m$ = modulation index. Vibrato results when the

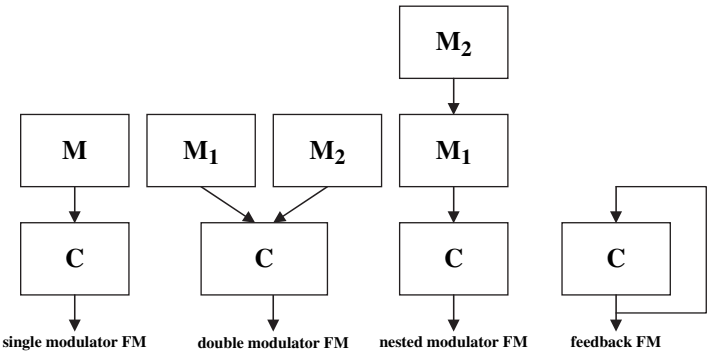


FIGURE 3 Block diagrams for several basic types of FM synthesis.

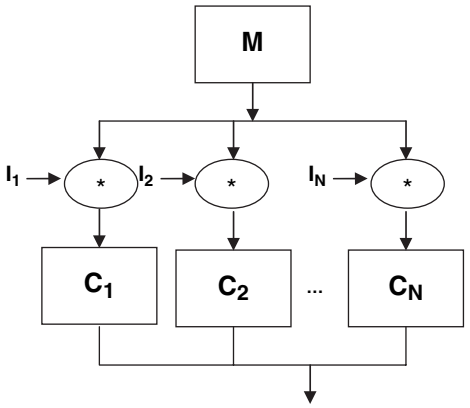


FIGURE 4 Block diagram of single-modulator/multiple-carrier (formant) FM synthesizer.

modulator frequency is low (< 20 Hz). However, with an audio-rate modulator frequency, a spectrum is heard whose frequencies depend on the carrier and modulator frequencies and whose amplitudes depend on the modulation index. This is made obvious by expanding Eq. (5) in terms of Bessel functions:

$$\sin(2\pi f_c t + \alpha_m \sin(2\pi f_m t)) = \sum_{k=-\infty}^{\infty} J_k(\alpha_m) \sin(2\pi(f_c + k f_m)t). \quad (6)$$

We see that the spectrum frequencies are given by $f_k = |f_c + kf_m|$, for $k = \ldots -3, -2, -1, 0, 1, 2, 3 \ldots$, which makes it clear that f_c is the center frequency with amplitude $J_0(\alpha_m)$ and that this component is surrounded by positive and negative side bands

of frequency $f_c \pm |k|f_m$ and amplitudes $J_k(\alpha_m)$. As k increases, frequency eventually a point is reached (when $k > f_c/f_m$) where the negative-k side-band frequencies become negative, causing the sine function to flip its sign. Negative frequencies beyond this point are said to "fold over zero". Thus, the frequencies are effectively just $|f_c \pm |k|f_m|$ with amplitudes of $\pm|J_{|k|}(\alpha_m)|$. Because, for the same value of $|k|$, the negative frequency component amplitudes are given by $J_{-k}(\alpha) = (-1)^k J_k(\alpha)$, Eq. (6) can be rewritten as

$$\sin(2\pi f_c t + \alpha_m \sin(2\pi f_m t)) = J_0(\alpha_m)\sin(2\pi f_c t) + \sum_{k=1}^{\infty} J_k(\alpha_m)[\sin(2\pi(f_c + kf_m)t) + (-1)^k \sin(2\pi(f_c - kf_m)t)] \quad , \quad (7)$$

which shows a clear separation between the positive and negative side-bands. The actual sign of each side-band amplitude depends on a combination of (1) whether it is a positive or negative side-band, (2) whether the Bessel function itself is positive or negative at a particular value of α_m, and (3) whether the component has folded-over (i.e., $k > f_c/f_m$). Also, note that the left sides of Eqs. (6) and (7) are actually in the form of phase modulation (PM). FM and PM are closely related because the phase is the integral of the frequency, i.e., $\alpha_m \sin(2\pi f_m t) = \int 2\pi\Delta f_m \cos(2\pi f_m t)\mathrm{d}t$. If some other sinusoid phase inside the integral is used (e.g., $\sin(2\pi f_m t)$), the result will be different, although basically the same type of spectra will result. For details see Beauchamp (1992).

The modulation index controls the amount of modulation and the precise shape of the resulting spectrum. Keeping f_m fixed, the spectrum bandwidth generally increases as the modulation index α_m increases. This effect is shown in Figure 5. For large α_m, the -40 dB bandwidth approaches $2\Delta f_m$. For small α_m, the bandwidth-to-Δf_m ratio gets larger; for example, at $\alpha_m = 1$, the ratio is approximately 6.2. Thus, a time-varying modulation index produces a dynamically changing spectrum from a single FM carrier–modulator oscillator pair. By contrast, single wavetable synthesis lacks this flexibility. Still, this simple FM model has the unfortunate property that, due to the oscillatory nature of the Bessel functions, as the modulation index changes, individual spectral components fade in and out dramatically, in a way that is not characteristic of typical musical tone spectral evolution.

A special case of Eq. (7) called *formant FM* occurs when the carrier frequency is an integer multiple of the modulator frequency. First, the spectrum is harmonic with fundamental frequency f_m, and second, for limited values of α_m, the spectrum tends to form a symmetrical formant band around the carrier frequency (see Figure 5).

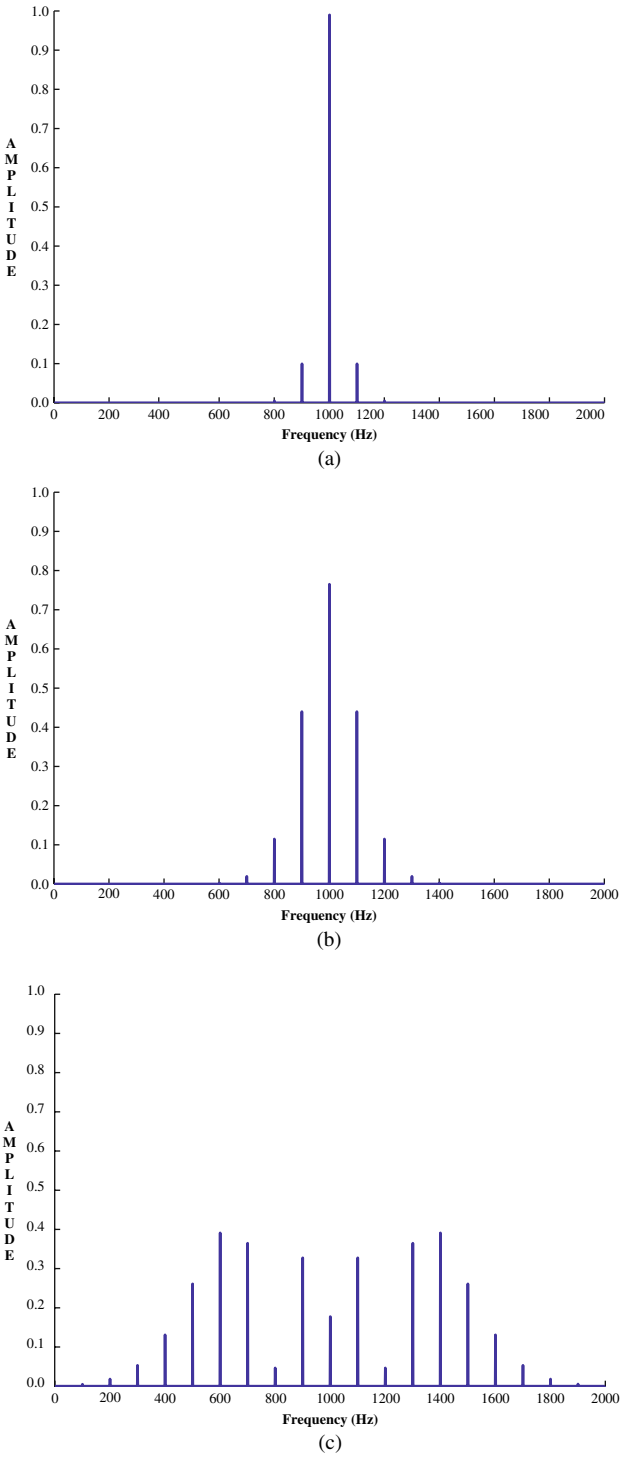


FIGURE 5 FM magnitude spectra for carrier frequency f_c, modulator frequency f_m, and index $\alpha_m =$ (**a**) 0.2, (**b**) 1.0, and (**c**) 5.0.

5.2 Single-Modulator/ Multiple-Carrier Synthesis Model

Like wavetable synthesis, FM synthesis is very efficient in terms of computation and storage: A single carrier–modulator FM instrument requires about the same amount of computation as a pair of wavetables. However, only a single sine wavetable is required for all the modulators and carriers of a complex FM synthesizer; thus, FM is more storage-efficient than wavetable synthesis. Also, assuming that modulation indices are not time-varying, each carrier's output has a static spectrum, and so the FM model provides time-varying spectrum control in similar fashion as multiple-wavetable synthesis (see Figure 2), i.e., by means of the time-varying amplitudes of the carriers (see Figure 4). However, unlike wavetable synthesis, the spectrum produced by FM is not arbitrary, but is restricted to the subset of possible FM spectra, so that more modules are generally needed for FM to produce a result of the same quality.

An equation for single-modulator/multiple-carrier FM synthesis, as depicted in Figure 4, is given by

$$s(t) = \sum_{n=1}^{N} A_n(t) \sin(2\pi(r_n f_{\mathrm{m}} t + \alpha_{\mathrm{m}_n} \sin(2\pi f_{\mathrm{m}} t))). \tag{8}$$

For each carrier oscillator n, the time-varying amplitude is given by $A_n(t)$, the modulator frequency by f_{m}, the carrier frequency by $r_n f_{\mathrm{m}}$ (where r_n is an integer), and the modulation index by α_{m_n}. The time-varying harmonic spectrum of $s(t)$ can be calculated by expanding each term of Eq. (8) using Eq. (7) and combining corresponding frequencies.

5.3 More Complex FM Models

A few years after Chowning's original work, Schottstaedt (1977) introduced a double-modulator/single-carrier (DMSC) FM model (see Figure 3), where the outputs of two oscillators are summed to modulate the carrier. Thus,

$$s(t) = A \sin(2\pi f_{\mathrm{c}} t + \alpha_{\mathrm{m}_1} \sin(2\pi f_{\mathrm{m}_1} t) + \alpha_{\mathrm{m}_2} \sin(2\pi f_{\mathrm{m}_2} t)). \tag{9}$$

If the carrier and modulator frequencies are all related by integer multiples of the fundamental, a harmonic tone results. As shown by Le Brun (1977), the DMSC FM model with N modulators and a single carrier produces frequencies of $|f_{\mathrm{c}} + k_1 f_{\mathrm{m}_1} + \cdots + k_N f_{\mathrm{m}_N}|$ with Bessel function product amplitudes $J_{k1}(\alpha_{\mathrm{m}_1}) J_{k_2}(\alpha_{\mathrm{m}_2}) \cdots J_{k_N}(\alpha_{\mathrm{m}_N})$ for all combinations of (positive, negative, or zero) integers $k_1, \ldots, k_N$. This is a significantly more complicated model, even with just two modulators, than the single-modulator case, as described above, thus making parameter optimization more difficult than for the single-modulator case.

Justice (1979) introduced a two-modulator nested FM model (see Figure 3), where one modulator modulates another, which in turn modulates the carrier:

$$s(t) = \sin(2\pi f_c t + \alpha_{m_1} \sin(2\pi f_{m_1} t + \alpha_{m_2} \sin(2\pi f_{m_2} t))). \quad (10)$$

Like double FM, if the carrier and modulator frequencies are all related by integer multiples of the fundamental, a harmonic tone results. By treating the modulator m_1 as a carrier, we see it can be expanded in terms of Bessel functions as in Eqs. (6) and (7) so that Eq. (10) could be expanded in the form of Eq. (9) with an infinite sum of modulators. Taking this a step further, as for the finite modulator sum case, results in an equation with infinite sums and products. This is a much more complicated relationship than the Bessel function expansion of double FM, where the modulators are summed instead of nested. However, if time samples of Eq. (10) are computed uniformly over the fundamental period (the inverse of the largest common divisor of f_c, f_{m_1}, and f_{m_2}), a discrete Fourier transform can be used to compute the harmonic amplitudes for any combination of the five parameters, f_c, f_{m_1}, f_{m_2}, α_{m_1}, and α_{m_2}.

Another FM variant that proved useful in the 1980s' synthesizers was *feedback FM* (Tomisawa 1981; Mitsuhashi 1982). The output of the carrier at sample n modulates the following sample, scaled by a modulation index:

$$s(n+1) = \sin(2\pi f_c n / f_s + \alpha_m s(n)). \quad (11)$$

Unlike the other FM methods discussed above, this is probably intractable for analytic solution. However, given a starting value, e.g., $s(0) = 0$, it can be computed and converted into a spectrum. It turns out that when the modulation index is less than about 1.5, a monotonically decreasing spectrum results (Tomisawa 1981). Because of this, feedback FM is potentially more easily controlled than the other forms of FM, where the harmonics oscillate in amplitude as the modulation index changes. Another advantage of feedback FM over other FM models is that its harmonic amplitudes are strictly positive when the modulation index is less than 1.5 (other forms of FM produce both positive and negative amplitudes). This avoids cancellation when adding multiple carriers together. Still, the monotonically decreasing spectrum of feedback FM has a disadvantage: Many musical instruments which have strong formants at upper harmonics cannot be directly modeled by the monotonic spectra inherent with feedback FM. Nevertheless, amplitude modulation of the feedback FM carrier signal by a sine wave can shift frequencies upward to overcome this limitation.

6 • Frequency Modulation Matching

One of the factors leading to FM's decline in popularity in the late 1980s was that matching an arbitrary musical instrument tone by FM synthesis is difficult, much more difficult than with wavetable matching. A closed-form analytical solution for determining the best set of FM parameters does not exist, and some form of general-purpose optimization is necessary. Most previous work on FM utilized ad hoc and semi-automated techniques for matching instrument tones. However, hand tuning of multiple carriers quickly exceeds the limits of human ability and endurance. In the early 1990s Horner et al. (1993b) introduced systematic techniques for matching tones to FM models based on genetic algorithm (GA) optimization.

Several researchers have emulated particular instruments by tailoring the time-varying FM indices by hand, starting with Chowning's original FM paper (Chowning 1973). In addition to applying FM to music synthesis, Chowning gave parameters appropriate to various classes of instruments based on simulating various properties of those instruments. For instance, a brass tone's spectral centroid (a strong correlate of perceptual brightness) is usually proportional to its overall amplitude. Chowning simulated this behavior by taking advantage of the fact that the centroid of an FM spectrum generally increases as its modulation index increases. He then varied the modulation index in direct proportion to the amplitude of the carrier to approximate a brass instrument. He produced woodwind-like tones and percussive sounds using similar methods. Chowning also discussed a double-carrier instrument near the end of his 1973 paper. In a later study, Chowning (1980) designed a double-carrier FM instrument to simulate a singing soprano voice. Chowning centered the first carrier at the fundamental frequency and the second at an upper formant, intuitively deriving the parameters of the instrument. He identified vibrato as critically important for achieving a voice-like sound.

Morrill (1977) followed Chowning's lead in trying to determine FM parameters based on detailed knowledge of trumpet sounds. Morrill outlined single- and double-carrier instrument designs for the trumpet and clearly identified the limitations of single-carrier instruments. His double-carrier instrument set carrier frequencies to the fundamental and its sixth harmonic, the latter corresponding to a known upper formant region in the trumpet. Morrill also pointed out the difficulty in predicting the spectral output of the double-carrier instrument.

Schottstaedt (1977) changed the basic FM instrument design by using two modulators to simultaneously frequency-modulate a single carrier. After describing the spectral behavior of the double-modulator FM model, Schottstaedt gave parameters for simulating the piano and string instruments. He used instrument characteristics and trial and error to find the parameters. The technique found small modulation indices most useful.

Justice (1979) introduced the nested modulator FM model and outlined a Hilbert transform procedure to decompose a signal into parameters for a single-carrier FM instrument. The procedure attempted to produce a matched FM signal close to the original, leaving the user to tweak the parameters as desired. However, Justice matched FM-generated signals and not those of acoustic musical instruments. Payne (1987) extended Justice's technique to a pair of carriers with nested modulators. Each carrier contributed to an independent frequency region, giving a more accurate match than with a single carrier. In addition to matching contrived FM signals, Payne matched a cello sound. The result was reportedly string-like, but lacking properties of "liveliness". Payne reported that the matching procedure was computationally very expensive.

More recent work by Delprat and his collaborators used a wavelet analysis and a Gabor transform to find spectral trajectories (Delprat et al. 1990; Delprat 1997). They used these trajectories to estimate modulation parameters. This approach was similar to that used by Justice and Payne except that it broke the frequency range into more component parts. Thus, it was also computationally expensive. They gave examples for a saxophone and trumpet using five carrier–modulator pairs, indicating the growing awareness that precise spectral control requires several carriers.

Meanwhile, Beauchamp (1982) developed a frequency-domain method to find FM parameters as part of a larger study on nonlinear synthesis based on time-varying spectral centroid (aka "brightness") matching. He used a single-modulator/single-carrier model with a centroid-controlled modulation index to match the time-varying spectral centroid of the original signal. Though the level of control was too coarse to provide a convincing synthesis, the technique was notable in its attempt to perform an automated spectral match.

In the early 1990s researchers introduced evolutionary matching techniques for the various FM models, first applying them to the single-modulator/multiple-carrier (formant FM) model (Horner et al. 1992, 1993b). A genetic algorithm (GA) procedure

was used to optimize fixed modulation indices and modulator-to-carrier frequency ratios for various numbers of carriers in an effort to best match particular instrument sounds. Fixed rather than time-varying modulation indices were used because allowing them to vary resulted in radically different frame-to-frame indices with their accompanying audible artifacts. Using invariant indices also avoided the considerable extra expense of optimizing time-varying indices. As in wavetable matching, the relative amplitude spectral error between the original and matched spectra served as a fitness function in guiding the GA's search for the best FM parameters. Matching results in terms of relative amplitude spectral error (see Eq. (4)) vary with instrument. Recent measurements on a trumpet tone showed that five formant-FM carriers were needed to achieve an error of 10%, whereas for a Chinese pipa tone an error of 25% required six carriers (Horner 2007). This performance was substantially worse than for the comparable wavetable-index method where only three tables were needed to yield an error of 7.5% for the trumpet and six tables for an error of 10% for the pipa.

Tan et al. (1994) introduced an enumerative procedure for optimizing a steady-state double-modulator FM model. Because this model only produced static spectra, it did not effectively match instruments with dynamic spectra. Since then, genetic algorithms have successfully optimized the double FM problem (Horner 1996a; Tan and Lim 1996; Lim and Tan 1999). GA methods were used to optimize invariant modulation indices in the double FM model, and the modulation indices it found were relatively small. Double FM matches were worse than formant-FM matches when compared with the same number of table lookups. However, double FM matches were better than formant-FM matches for the same number of carriers, an advantage when double FM hardware is available.

Finally, the GA method was applied to nested modulator and feedback FM matching (Horner 1998). Fixed modulation indices were used for nested modulator FM, but time-varying modulation indices were allowed for feedback FM. Like double FM matching, the optimized parameters for nested modulator and feedback FM had relatively small modulation indices. Feedback FM often gave the best matches of all the FM models when compared against the same number of table lookups, indicating feedback FM is a good choice for software synthesis where computation is the main factor. However, if nested modulator FM hardware is already available, then double or triple modulator FM gave the best results of all the FM models for the same number of carriers.

7 • Conclusions

We have reviewed several techniques for instrument modeling/synthesis based on spectral analysis, with a particular focus on wavetable and FM matching. Among the various types of wavetable synthesis, the best method depends on the given situation. For simplicity, group additive synthesis has the advantage of being intuitive, since each harmonic is only in one wavetable. For memory-constrained systems where instruments have to compete for limited wavetable space, wavetable matching is a very good choice. Conversely, for real-time systems where memory is not a problem, wavetable interpolation is a good choice.

Although wavetable matching is simpler and more effective than FM matching in general (Horner 1997), FM synthesis provides real-time flexibility over wavetable synthesis when wavetable memory is limited. Among the various types of FM, the best method again depends on the given situation. For simplicity and ease of control, formant FM is best. For software synthesis where computation is the main factor, feedback FM is best. If FM hardware is available, nested modulator FM is best.

In any case, the various wavetable and FM matching procedures provide an interesting point of departure for instrument designers in applications such as *timbral interpolation* (Grey 1975; Beauchamp and Horner 1998), where the parameters of one spectral match are crossfaded to that of another. The smoothness of the transformation depends on the synthesis technique. Wavetable synthesis gives a smoother interpolation than FM synthesis. Interpolating distantly spaced FM index values will likely produce wildly changing spectral results during the interpolation due to oscillation of the Bessel functions. However, such interpolations may still be interesting and musically useful.

Acknowledgments The Hong Kong Research Grant Council's Projects 613505 and 613806 supported this work.

References

Beauchamp, J.W. (1974), "Time-Variant Spectra of Violin Tones", *J. Acoust. Soc, Am.* **56**(3), 995–1004.

Beauchamp, J. (1982). "Synthesis by Amplitude and 'Brightness' Matching of Analyzed Musical Instrument Tones", *J. Audio Eng. Soc.* **30**(6), pp. 396–406.

Beauchamp, J. (1992). "Will the Real FM Equation Please Stand Up"(letter), *Comput. Music J.* **16**(4), pp. 6–7.

Beauchamp, J.W., R.C. Maher, & R. Brown (1993). "Detection of Musical Pitch from Recorded Solo Performances", *Audio Eng. Soc.* Preprint No. 3541.

Beauchamp, J.W., & A, Horner (1995). "Wave table Interpolation Synthesis Based on Time-Variant Spectral Analysis of Musical Sounds", *Audio Eng. Soc.* Preprint No. 3960.

Beauchamp, J., & A. Horner (1998). "Spectral Modeling and Timbre Hybridization Programs for Computer Music", *Organ. Sound* **2**(3), pp. 253–258.

Beauchamp, J.W. (2007). "Analysis and Synthesis of Musical Instrument Sounds", in Analysis, Synthesis, and Perception of Musical Sounds; J.W. Beauchamp, ed., New York: Springer, pp. 1–89.

Bristow-Johnson, R. (1996). "Wavetable Synthesis 101, A Fundamental Perspective", *Audio Eng. Soc.* Preprint No. 4400.

Brown, J.C. (1996). "Frequency ratios of Spectral Components of Musical Sounds", *J. Acoust. Soc. Am.* **99**(2), 1210–1218.

Chamberlin, H. (1980). "Advanced Real-Timbre Music Synthesis Techniques", *Byte Magazine* April, pp. 70–94, 180–196.

Cheung, N.-M., & A. Horner (1996). "Group Synthesis with Genetic Algorithms", *J. Audio Eng. Soc.* **44**(3), pp. 130–147.

Chowning, J. (1973). "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation", *J. Audio Eng. Soc.* **21**(7), pp. 526–534.

Chowning, J. (1980). "Computer Synthesis of the Singing Voice", *Sound Generation in Wind, Strings, Computers*, Stockholm: The Royal Swedish Academy of Music, pp. 4–13.

Delprat, N., P. Guillemain, & R. Kronland-Martinet (1990). "Parameter Estimation for Non-linear Resynthesis Methods with the Help of a Time–Frequency Analysis of Natural Sounds", *Proc. 1990 Int. Computer Music Conf.*, Glasgow, Scotland, pp. 88–90.

Delprat, N. (1997). "Global Frequency Modulation Law Extraction from the Gabor Transform of a Signal: A First Study of the Interacting Components Case", *IEEE Trans. Speech Audio Process.* **5**(1), pp. 64–71.

Dubnof, S. & X. Rodet (2003), "Investigation of phase Coupling Phenomend in Sustained Portion of Musical Instrument Sound", *J. Acoust, Soc.* **113**(1), 348–359.

Dunteman, G. (1989). *Principal Components Analysis*, Newbury Park, CA: Sage.

Ehmann, A. F., & J. W. Beauchamp (2002) "Musical sound analysis/synthesis using vector-quantized time-varying spectra" (abstract), *J. Acoust. Soc. Am.* **112** (5, pt. 2), p. 2239.

Flanagan, J. L., & R. M. Golden (1966). "Phase Vocoder", *Bell Syst. Tech. J.* **45**, pp. 1493–1509.

Fletcher, H., E. D. Blackham, & R. Stratton (1962). "Quality of Piano Tones", *J. Acoust. Soc. Am.* **34**(6), pp. 749–761.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Reading, MA: Addison-Wesley.

Grey, J. (1975). "An Exploration of Musical Timbre", Ph.D. Dissertation, Stanford: Dept. of Music, Stanford Univ.

Horner, A., J. Beauchamp, & L. Haken (1992). "Wavetable and FM Matching Synthesis of Musical Instrument Tones", *Proc. 1992 Int. Computer Music Conf.*, San Jose, CA, pp. 18–21.

Horner, A., J. Beauchamp, & L. Haken (1993a). "Methods for Multiple Wavetable Synthesis of Musical Instrument Tones", *J. Audio Eng. Soc.* **41**(5), pp. 336–356.

Horner, A., J. Beauchamp, & L. Haken (1993b). "Machine Tongues XVI: Genetic Algorithms and Their Application to FM Matching Synthesis", *Comput. Music J.* **17**(4), pp. 17–29.
Horner, A. (1995). "Wavetable Matching Synthesis of Dynamic Instruments with Genetic Algorithms", *J. Audio Eng. Soc.* **43**(11), pp. 916–931.
Horner, A. & J. Beauchamp (1995), "Synthesis of Trumpet Tones Using a Wavetable and a Dynamic Filter", *J. Audio Eng. Soc.* **43**(10), 799–812.
Horner, A. (1996a). "Double Modulator FM Matching of Instrument Tones", *Comput. Music J.* **20**(2), pp. 57–71.
Horner, A. (1996b). "Computation and Memory Tradeoffs with Multiple Wavetable Interpolation", *J. Audio Eng. Soc.* **44**(6), pp. 481–496.
Horner, A., & J. Beauchamp (1996). "Piecewise Linear Approximation of Additive Synthesis Envelopes: A Comparison of Various Methods", *Comput. Music J.* **20**(2), pp. 72–95.
Horner, A. (1997). "A Comparison of Wavetable and FM Parameter Spaces", *Comput. Music J.* **21**(4), pp. 55–85.
Horner, A. (1998). "Nested Modulator and Feedback FM Matching of Instrument Tones", *IEEE Trans. Speech Audio Process.* **6**(4), pp. 398–409.
Horner, A., & L. Ayers (1998). "Modeling Acoustic Wind Instruments with Contiguous Group Synthesis", *J. Audio Eng. Soc.* **46**(10), pp. 868–879.
Horner, A. (2001). "A Simplified Wavetable Matching Method Using Combinatorial Basis Spectra Selection", *J. Audio Eng. Soc.* **49**(11), pp. 1060–1066.
Horner, A., J. Beauchamp, & R. So (2006). "A Search for Best Error Metrics to Predict Discrimination of Original and Spectrally Altered Musical Instrument Sounds", *J. Audio Eng. Soc.* **54**(3), 140–156.
Horner, A. (2007). "A Comparison of Wavetable and FM Data Reduction Methods for Resynthesis of Musical Sounds", in *Analysis, Synthesis, and Perception of Musical Sounds*, J. W. Beauchamp, ed., New York: Springer, pp. 228–249.
Justice, J. (1979). "Analytic Signal Processing in Music Computation", *IEEE Trans. Acoust. Speech Signal Process.* **27**(6), pp. 670–684.
Kleczkowski, P. (1989). "Group Additive Synthesis", *Comput. Music J.* **13**(1), pp. 12–20.
Le Brun, M. (1977). "A Derivation of the Spectrum of FM with a Complex Modulating Wave", *Comput. Music J.* **1**(4), pp. 51–52.
Lee, K., & A. Horner (1999). "Modeling Piano Tones with Group Synthesis", *J. Audio Eng. Soc.* **47**(3), pp. 101–111.
Lim, S. M., & B. T. G. Tan (1999). "Performance of the Genetic Annealing Algorithm in DFM Synthesis of Dynamic Musical Sound Samples", *J. Audio Eng. Soc.* **47**(5), pp. 339–354.
Maher, R., & J. Beauchamp (1990). "An Investigation of Vocal Vibrato for Synthesis", *Appl. Acoust.* **30,** pp. 219–245.
Maher, R.C. & J.W. Beauchamp (1994). "Fundamental Frequency Estimation of Musical Signals Using a Two-way Mismatch Procedures", *J. Acoust. Soc. Am.* **95**(4), 2254–2263.
Mathews, M. V. (1969). *The Technology of Computer Music*, Cambridge, MA: MIT Press, p. 56.
McAulay, R. J., & T. F. Quatieri (1986). "Speech Analysis/Synthesis Based on a Sinusoidal Representation", *IEEE Trans. Acoust. Speech Signal Process.* **34**(4), pp. 744–754.

Mitsuhashi, Y. (1982). "Musical Sound Synthesis by Forward Differences", *J. Audio Eng. Soc.* **30**(1/2), pp. 2–9.

Morrill, D. (1977). "Trumpet Algorithms for Computer Composition", *Comput. Music J.* **1**(1), pp. 46–52.

Oates, S., & B. Eaglestone (1997). "Analytic Methods for Group Additive Synthesis", *Comput. Music J.* **21**(2), pp. 21–39.

Payne, R. (1987). "A Microcomputer Based Analysis/Resynthesis Scheme for Processing Sampled Sounds Using FM", *Proc. 1987 Int. Computer Music Conf.*, Urbana, IL, pp. 282–289.

Portnoff, M. R. (1976). "Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform", *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-24**(3), pp. 243–248.

Rabiner, L., & B.-H. Juang (1993). *Fundamentals of Speech Recognition*, New Jersey: Prentice Hall, pp. 122–132.

Sandell, G., & W. Martens (1992). "Prototyping and Interpolation of Multiple Musical Timbres Using Principal Component-Bases Synthesis", *Proc. 1992 Int. Computer Music Conf.*, San Jose, CA, pp. 34–37.

Sandell, G., & W. Martens (1995). "Perceptual Evaluation of Principal-Component-Based Synthesis of Musical Timbres", *J. Audio Eng. Soc.* **43**(12), pp. 1013–1028.

Schottstaedt, B. (1977). "The Simulation of Natural Instrument Tones Using Frequency Modulation with a Complex Modulating Wave", *Comput. Music* J. **1**(4), pp. 46–50.

Schumacher, R.T. (1992). "Analysis of Aperiodicities in Nearly Periodic Waveforms", *J. Acoust. Soc. Am.* **91**(1), 438–451.

Serra, M.-H., D. Rubine, & R. Dannenberg (1990). "Analysis and Synthesis of Tones by Spectral Interpolation", *J. Audio Eng. Soc.* **38**(3), pp. 111–128.

Smith, J. O., & X. Serra (1987). "PARSHL: An Analysis/Synthesis Program for Non-harmonic Sounds Based on a Sinusoidal Representation", *Proc. 1987 Int. Computer Music Conf.,* Urbana, IL, pp. 290–297.

So, K. F., & A. Horner (2002). "Wavetable Matching of Inharmonic String Tones", *J. Audio Eng. Soc.* **50**(1/2), pp. 47–57.

Tan, B. T. G., S. L. Gan, S. M. Lim, & S. H. Tang (1994). "Real-Time Implementation of Double Frequency Modulation (DFM) Synthesis", *J. Audio Eng. Soc.* **42**(11), pp. 918–926.

Tan, B. T. G., & S. M. Lim, (1996). "Automated Parameter Optimization for Double Frequency Modulation Synthesis Using the Genetic Annealing Algorithm", *J. Audio Eng. Soc.* **44**(1/2), pp. 3–15.

Tomisawa, N. (1981). "Tone Production Method for an Electronic Music Instrument", U.S. Patent 4,249,447.

Zahorian, S., & M. Rothenberg (1981). "Principal-Components Analysis for Low Redundancy Encoding of Speech Spectra", *J. Acoust. Soc. Am.* **69**(3), pp. 832–845.

Zheng, H., & J. Beauchamp (1999). "Analysis and Critical-Band-Based Group Wavetable Synthesis of Piano Tones", *Proc. 1999 Int. Computer Music Conf.*, Beijing, China, pp. 9–12.

25

Digital Waveguide Architectures for Virtual Musical Instruments

Julius O. Smith

Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA 94305, USA,
Website: http://ccrma.stanford.edu/~jos/

1 • Introduction

Digital sound synthesis has become a standard staple of modern music studios, videogames, personal computers, and hand-held devices. As processing power has increased over the years, sound synthesis implementations have evolved from dedicated chip sets, to single-chip solutions, and ultimately to software implementations within processors used primarily for other tasks (such as for graphics or general purpose computing). With the cost of implementation dropping closer and closer to zero, there is increasing room for higher quality algorithms. A particularly fertile source of natural sound synthesis algorithms is the mathematical models of musical instruments developed within the science of musical acoustics [1–3, 72]. To realize practical instrument voices from these models, it is helpful to develop robust and efficient signal processing algorithms which retain the audible physical behavior while minimizing computational cost [4].

In this chapter, a number of cost-effective synthesis models will be summarized for various musical instrument families, including strings and winds. Emphasis is placed on techniques adapted from the field of digital signal processing [5,6]. Notably absent is any discussion of percussion instruments, which are normally handled via sample-based methods [7–9], although some model-based methods have been proposed based e.g., on the digital waveguide mesh [10].

2 • Vibrating Strings

In a stringed musical instrument, most of the sound energy is stored in the vibrating string at any given time. The expressive range of sound from a stringed instrument arises primarily from an interaction of the vibrating string with the performer. The body of the instrument functions as a passive resonator which is well modeled in most cases, in principle, by a linear, time-invariant filter [11,12].

2.1 Wave Equation

The starting point for a stringed instrument model is typically a *wave equation* for transverse vibrations of the vibrating string [13–16]. For example, a recently proposed [14] partial differential equation (PDE) governing the motion of a *piano string* is given by

$$f(t,x) = \epsilon\ddot{y} - Ky'' + EIy'''' + R_0\dot{y} + R_2\ddot{y}' \tag{1}$$

where symbols are defined as follows:

$y = y(t,x) =$ string displacement at position x and time t
$\dot{y} = \frac{\partial}{\partial t}y(t,x)$, $y' = \frac{\partial}{\partial x}y(t,x)$, etc.
$f(t,x) =$ driving force density (N/m)
$\epsilon =$ mass density (kg/m)
$K =$ tension force along the string axis (N)
$E =$ Young's modulus (N/m^2)
$I =$ radius of gyration of the string cross-section (m)

The basic *lossless* wave equation $\epsilon\ddot{y} = Ky''$ is derived in most textbooks on acoustics, e.g., [15].[1] The term $\epsilon\ddot{y}$ is the mass-per-

[1] For an online derivation, see, e.g., http://ccrma.stanford.edu/~jos/pasp/String_Wave_Equation.html.

unit-length ϵ times the transverse acceleration $\ddot{y}$, and Ky'' approximates the transverse restoring force (per unit length) due to the string tension K.

The term EIy'''' in Eq. (1) models the transverse restoring force exerted by a *stiff string* in response to being bent. In an ideal string, with zero diameter, this force is zero. Stiffness is normally neglected in models for guitars and violins, but modeled in instruments with larger-diameter strings, such as the piano and cello. The test for whether stiffness is needed in the model for plucked or struck strings (any freely vibrating string) is whether the ear can hear the "stretching" of the partial overtones due to stiffness [17]; for bowed strings, dispersion due to stiffness can affect *bow–string dynamics* [18]. In the context of a digital waveguide string model (described in Section 2.3), the dispersion associated with stiff strings is typically modeled indirectly by designing an *allpass filter* for the string model.

The final two terms of Eq. (1) provide *damping*, which is required in any string practical model. Papers including damping models for piano strings include [13,14]. Instead of introducing such terms into the wave equation based on physical considerations, these terms may be determined implicitly by *digital filter design* techniques [11,19,29], as discussed further in Section 2.5.

2.2 Finite Difference Models

The original approach to digitally modeling vibrating strings was by means of *finite difference schemes* (FDS) [13,20–24]. Such models are also called finite difference time domain (FDTD) methods [25,26]. In these models, partial derivatives are replaced by finite differences, e.g.,

$$\ddot{y}(t,x) \approx \frac{y(t+T,x) - 2y(t,x) + y(t-T,x)}{T^2} \tag{2}$$

$$y''(t,x) \approx \frac{y(t,x+X) - 2y(t,x) + y(t,x-X)}{X^2}. \tag{3}$$

These finite differences are substituted into the string wave equation, Eq. (1), from which a time-recursive finite difference scheme is derived. Given boundary conditions characterizing the string endpoints, the finite difference scheme may be used to recursively compute the time evolution of the string state.

2.3 Digital Waveguide Models

More recently, the *digital waveguide* (DW) approach has been developed for modeling vibrating strings [27,28,41,53]. The DW approach is compared quantitatively with FDS approaches in

[14,31]. For strings used in typical musical instruments, the digital waveguide method generally provides a more efficient simulation for a given sound quality level. A combination of digital waveguides and finite differences may be preferred, however, for nonlinear string simulation [26,32,33].

The digital waveguide formulation can be derived by simply *sampling* the *traveling-wave* solution to the ideal wave equation:

$$Ky'' = \epsilon\ddot{y}.$$

It is easily checked that the lossless 1D wave equation is solved by any string shape y which travels to the left or right with speed $c = \sqrt{K/\epsilon}$ [34]. Denote *right-going* traveling waves in general by $y_{\mathrm{r}}(t - x/c)$ and *left-going* traveling waves by $y_{\mathrm{l}}(t + x/c)$, where y_{r} and y_{l} are assumed twice-differentiable. Then, as is well known, the general class of solutions to the lossless, 1D, second-order wave equation can be expressed as

$$y(t,x) = y_{\mathrm{r}}\left(t - \frac{x}{c}\right) + y_{\mathrm{l}}\left(t + \frac{x}{c}\right). \tag{4}$$

Sampling these traveling-wave solutions using sampling intervals T and X along the time and position coordinates, respectively, yields

$$\begin{aligned} y(nT,mX) &= y_{\mathrm{r}}(nT - mX/c) + y_{\mathrm{l}}(nT + mX/c) \\ &= y_{\mathrm{r}}[(n-m)T] + y_{\mathrm{l}}[(n+m)T] \\ &= y^{+}(n-m) + y^{-}(n+m), \end{aligned} \tag{5}$$

where n and m are integers. Note that a "+" superscript denotes a "right-going" traveling-wave component, while "−" denotes propagation to the "left". This notation is similar to that used for acoustic-tube modeling of speech [5].

Figure 1 shows a signal flow diagram for the computational model of Eq. (5), which is often called a digital waveguide model (for the ideal string in this case) [28,41]. Note that, by the sampling theorem [35, Appendix G],[2] it is an exact model so long as the initial conditions and any ongoing additive excitations are bandlimited to less than half the temporal sampling rate $f_{\mathrm{s}} = 1/T$.

Note also that the position along the string, $x_m = mX = mcT$ meters, is laid out from left to right in the diagram, giving a

[2] http://ccrma.stanford.edu/~jos/mdft/Sampling_Theorem.html.

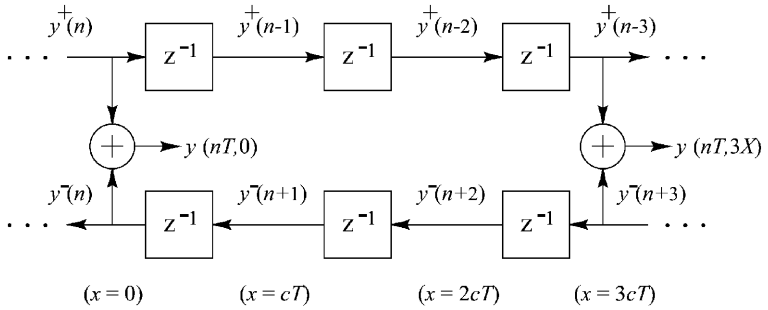


FIGURE 1 Digital simulation of the ideal, lossless, 1D waveguide with observation points at $x = 0$ and $x = 3X = 3cT$. (The symbol "z^{-1}" denotes a one-sample delay.) [Reprinted with permission from [41].]

physical interpretation to the horizontal direction in the diagram, even though explicit spatial samples have been eliminated in the translation of physical variables to traveling-wave components. In Figure 1, "transverse displacement outputs" have been arbitrarily placed at $x = 0$ and $x = 3X$. The diagram is similar to that of well-known ladder and lattice digital filter structures [5], except for the delays along the upper rail, the absence of "scattering junctions", and the direct physical interpretation. Also closely related is the *Kelly–Lochbaum model* for voice synthesis [5,36–39].

2.4 Bowed Strings

An example DW model for a bowed-string instrument is shown in Figure 2 [40,41]. The main control is bow velocity, but bow force and position also have an effect on the tone produced. The digital waveguide simulates traveling velocity-wave components. The left- and right-going traveling-wave components on the left of the bow are denoted $v^+_{s,l}(n)$ and $v^-_{s,l}(n)$, respectively, where n denotes time in samples. To the right of the bow, the components are $v^+_{s,r}(n)$ and $v^-_{s,r}(n)$. The (abstract) "incoming string velocity" is defined as

$$v^+_s(n) = v^+_{s,l}(n) + v^+_{s,r}(n) \tag{6}$$

and the "incoming differential velocity" is defined as

$$v^+_\Delta(n) = v_b(n) - v^+_s(n), \tag{7}$$

where $v_b(n)$ denotes the bow velocity at sample time n. The incoming differential velocity v^+_Δ can be interpreted physically as

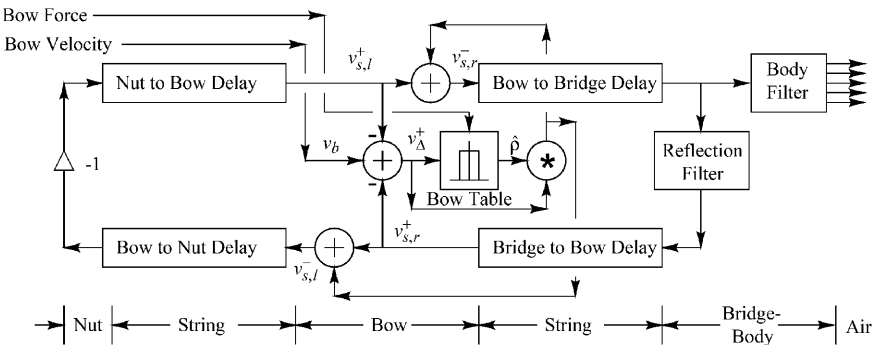


FIGURE 2 Digital waveguide bowed-string model. [From [41].]

the physical differential velocity (bow minus string) that would occur if the bow–string friction were zero (ideal, frictionless "slipping" of the bow along the string). A table lookup (or other nonlinear function implementation) gives the *reflection coefficient* of the bow–string contact point, as seen by traveling waves on the string (labeled "Bow Table" in the figure). This coefficient is then applied to v_Δ^+ and added to the left- and right-going traveling-wave paths. The bow table is derived from the bow–string friction curve characteristic, such as the one shown in Figure 3. Details of this derivation may be found in [41].[3]

The delay lines are drawn in "physical canonical form" for ease of physical interpretation. We see that the string is modeled using two ideal (lossless) digital waveguides, one to the left and one to the right of the bowing point. (A 1D *digital waveguide* is formally defined as a pair of delay lines flowing in opposite directions — a *bidirectional delay line*.) In practice, only two delay lines are normally implemented, one on each side of the bowing point.

Note that delay lines require $\mathcal{O}(1)$ operations per sample, i.e., the number of operations per sample does not increase as the delay-line length is increased.[4] This is the heart of the reason digital waveguide models are more efficient than finite difference models. At present, there is no known $\mathcal{O}(1)$ finite difference model for vibrating strings.

---

[3]Available online at
http://ccrma.stanford.edu/~jos/pasp/BSSJ.html.

[4]The notation $\mathcal{O}(K)$ denotes "computational complexity of order K". This means that the computational complexity is bounded by cN^K for some constant c, as $N \to \infty$, where N is the size of the problem (delay-line length in this case).

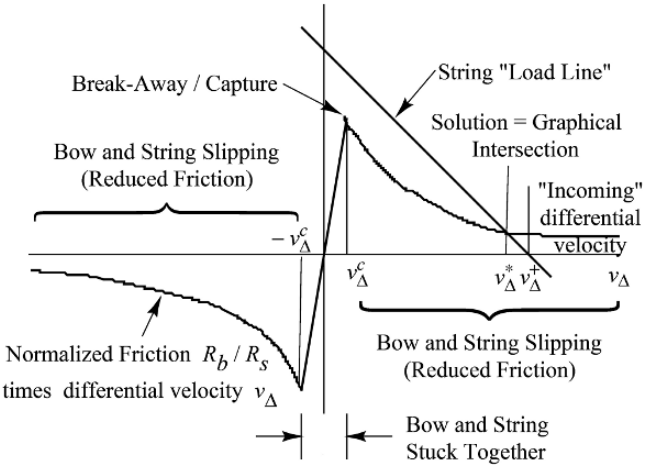


FIGURE 3 Overlay of normalized bow–string friction curve $R_b(v_\Delta)/R_s$ with the string "load line" $v_\Delta^+ - v_\Delta$. The "capture" and "break-away" differential velocity is denoted v_Δ^c. Note that increasing the bow force increases v_Δ^c as well as enlarging the maximum force applied (at the peaks of the curve). [From [41].]

The reflection filter in Figure 2 implements *all losses in one period of oscillation* due to the yielding bridge, absorption by the bow and finger, string losses, etc. Since the string model is linear and time invariant, i.e., Eq. (1) is linear with constant coefficients, superposition applies, and loss/dispersion filtering within the string may be *commuted* to concentrated points. In principle, such filters should appear on either side of the bow, and prior to each output signal extracted. However, because the difference is perceptually moot, normally only one loss/dispersion filter is employed per string loop. For multiple coupled strings, all loss/dispersion filtering may be implemented within the *bridge* at which they share a common termination [41,42].[5]

The bow–string junction is typically implemented as a *memoryless* lookup table ("Bow Table" in Figure 2) or a segmented polynomial. Preferably, however, a *thermodynamic model* should be employed for bow friction, since the bow rosin is known to have a time-varying viscosity due to temperature variations within a period of sound [43]. In [44], thermal models of dynamic friction in bowed strings are discussed, and such models have been incorporated into more recent synthesis models [45–47].

[5]http://ccrma.stanford.edu/~jos/pasp/Two_Ideal_Strings_Coupled.html.

A real-time software implementation of a bowed-string model similar to that shown in Figure 2 is available in the Synthesis Tool Kit (STK) distribution [48,49] as `Bowed.cpp`. This prototype can serve as a starting framework for more elaborate models.

2.5 Filter Design Methods for String Models

Methods for designing the string loop filter ("Reflection Filter" in Figure 2) are described in, [11,29,50–53]. For all types of waveguide string models, the loop-filter design software should ideally minimize *audibility of the error*. The loop-filter error naturally divides into two components: the *amplitude response* of the loop filter determines the partial overtone *decay rates*, while the *phase response* determines the precise partial overtone *tuning* [11, pp. 182–184]. These error components are normally addressed separately. That is, string loop filters are normally designed as a series connection of a low-order filter having an inaudible phase response (the *damping filter*) and an allpass filter (the *dispersion filter*), if any. Only stiff strings require an allpass filter component, and only if the stiffness is sufficient to audibly "stretch" the tuning of the partial overtones. An elementary introduction to digital filters and their analyses is given in [12].[6]

In [54], the damping in real piano strings was modeled using a length 17 linear-phase FIR filter for the lowest strings and a length 9 linear-phase FIR filter for the remaining strings. It is not necessary to restrict consideration to linear-phase filters, since a minimum-phase design may sound just as good, and any undesirable aspects of the phase response may be compensated in the dispersion-filter design.

In the case of stiff piano strings, it is possible to well tune the first several tens of partials with total allpass order of 20 or less [55]. Additionally, minimization of the L^∞ norm [56] has been used to calibrate a series of allpass-filter sections [57]. A summary of methods for dispersion-filter design is given in [51].

Filter design methods applicable to acoustic resonators (the "Body Filter" in Figure 2) are discussed in [11,58]. The *commuted synthesis* technique, described in Sect. 2.8 below, replaces a parametric body filter with its nonparametric impulse response (filtered by excitation characteristics). The commuted synthesis technique is applicable to plucked and struck strings, and also to *linearized* bowed-string models [42].

2.6 Electric Guitars

While most musical vibrating strings are well approximated as linear, time-invariant systems, there are special cases in which

[6] http://ccrma.stanford.edu/~jos/filters/.

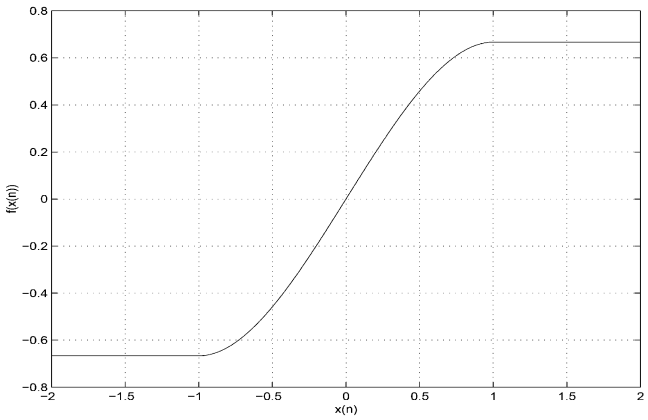


FIGURE 4 Soft clipper defined by Eq. (8). [From [41].]

nonlinear behavior is desired. One example is the distorted electric guitar [30].

A *soft clipper* is similar to a hard clipper (saturation on overflow), but with the "corners" smoothed. A common choice of soft clipper is the *cubic nonlinearity*, e.g., [59],

$$f(x) = \begin{cases} -\frac{2}{3}, & x \le -1 \\ x - \frac{x^3}{3}, & -1 \le x \le 1 \\ \frac{2}{3}, & x \ge 1. \end{cases} \tag{8}$$

This particular soft-clipping characteristic is diagrammed in Figure 4. An analysis of its spectral characteristics, along with some discussion regarding how to avoid the aliasing it can cause, is given in [41].[7] An input gain may be used to set the desired degree of distortion.

A cubic nonlinearity, as well as *any* odd distortion law,[8] generates only odd-numbered harmonics (like in a square wave). For best results, and in particular for *tube distortion* simulation [60,61], it can be argued that some amount of even-numbered harmonics should also be present. Breaking the odd symmetry in any way will add even-numbered harmonics to the output as well. One simple way to accomplish this is to add an *offset* to the input signal, obtaining

$$y(n) = f[x(n) + c], \tag{9}$$

[7] http://ccrma.stanford.edu/~jos/pasp/Nonlinear_Elements.html.
[8] A function $f(x)$ is said to be *odd* if $f(-x) = -f(x)$.

where c is some small constant. (Signals $x(n)$ in practice are typically constrained to be zero mean by one means or another.)

Another method for breaking the odd symmetry is to add some square-law nonlinearity to obtain

$$f(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta, \tag{10}$$

where β controls the amount of square law distortion. This is then a more general third-order polynomial. A square law is the gentlest nonlinear distortion, as can be seen by considering the Taylor series expansion of a general nonlinearity transfer characteristic $f(x)$. The constant δ can be chosen to zero the mean, on average; if the input signal $x(n)$ is zero mean with variance 1, then $\delta = -\beta$ compensates the nonzero mean introduced by the squaring term. The term γ can be modified to adjust the "effect mix".

2.7 Amplifier Feedback

A nonlinear feedback effect used with distorted electric guitars is *amplifier feedback*. In this case, the amplified guitar signal couples back into the strings with some gain and delay, as depicted schematically in Figure 5 [59]. The feedback delay can be adjusted to cause different partial overtones to be amplified relative to others.

2.8 Commuted Synthesis

Figure 6 shows a schematic diagram of a plucked or struck stringed instrument, such as a guitar. When the string and resonator may be assumed linear and time invariant, they may be *commuted*, i.e., connected in the opposite series ordering [12], as

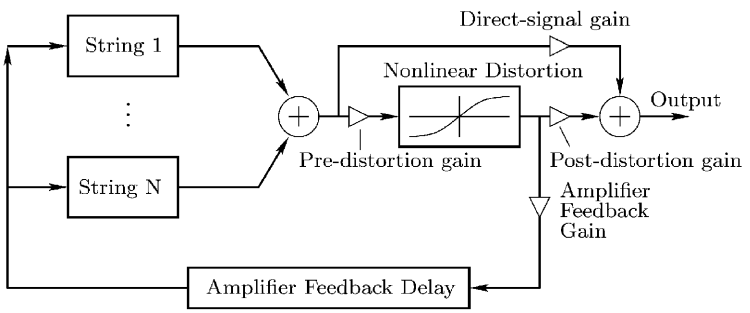


FIGURE 5 Simulation of a basic distorted electric guitar with amplifier feedback. [From [41].]



FIGURE 6 Schematic diagram of a stringed musical instrument. [From [29].]



FIGURE 7 Schematic diagram of commuted synthesis of plucked/struck stringed instruments. [From [29].]

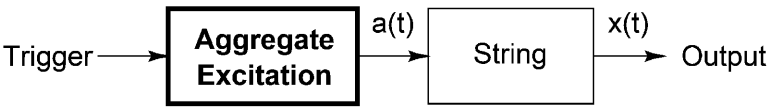


FIGURE 8 Use of an aggregate excitation given by the convolution of original excitation with the resonator impulse response. [From [41].]

shown in Figure 7. Sound synthesis algorithms based on this arrangement are typically called *commuted waveguide synthesis* algorithms [41,42,62,63].

Instead of plucking the string and filtering the string output with a digital filter of extremely high order (to capture the many resonances in the range of human hearing), the "pluck response" of the guitar body (a filtered impulse response) can be fed to the string instead, as shown in Figure 8. In a typical implementation, the guitar-body impulse response (or some filtering of it) is stored in table, just as in sampling synthesis, and a low-order filter is applied to the table playback in order to impart details of the plucking excitation. This simplification exchanges an expensive body filter for an inexpensive "pluck filter". In addition to body resonances, the excitation table may include characteristics of the listening space as well. Commuted synthesis of the piano has also been developed [14,57,64,65].

3 • Wind Instruments

A basic DW model for a single-reed woodwind instrument, such as a clarinet, is shown in Figure 9 [40,41,66].

When the bore is cylindrical (plane waves) or conical (spherical waves), it can be modeled quite simply using a bidirectional delay line [66]. Because the main control variable for the instrument

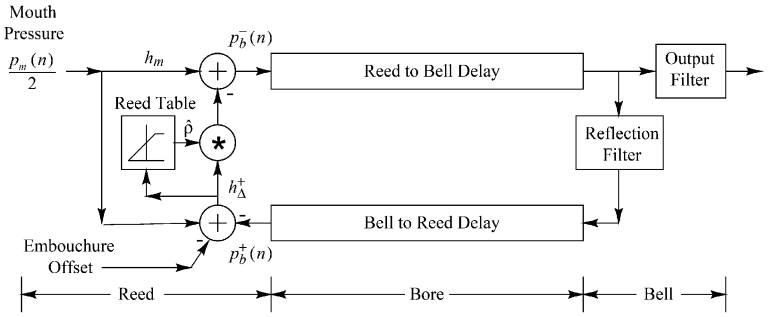


FIGURE 9 Waveguide model of a single-reed, cylindrical-bore woodwind, such as a clarinet. [From [41].]

is air pressure in the mouth at the reed, it is convenient to choose pressure-wave variables. Thus, the delay lines carry left-going and right-going *pressure* samples p_{b}^+ and p_{b}^- (respectively) which represent the traveling pressure-wave components within the bore.

To first order, the bell passes high frequencies and reflects low frequencies, where "high" and "low" frequencies may be divided by the wavelength which equals the bell's diameter. Thus, the bell can be regarded as a simple "cross-over" network, as is used to split signal energy between a woofer and tweeter in a loudspeaker cabinet. For a clarinet bore, the nominal "cross-over frequency" is around 1500 Hz [67].

The reflection filter at the right of the figure implements the bell or tone-hole losses as well as the round-trip attenuation losses from traveling back and forth in the bore. The bell output filter is highpass and *power complementary* with respect to the bell reflection filter. This must be true because the incident wave is assumed to either reflect or transmit, and all wave energy must be accounted for. (We assume that the bell itself does not vibrate or otherwise absorb sound.) From continuity of pressure and conservation of flow, we also have that the bell and reflection filters are *amplitude complementary* [41]. Therefore, given the reflection filter $H_{\mathrm{r}}(z)$, the transmission filter is simply $H_{\mathrm{t}}(z) = 1 - H_{\mathrm{r}}(z)$ for velocity waves, or $H_{\mathrm{t}}(z) = 1 + H_{\mathrm{r}}(z)$ for pressure waves. That such filters are also power complementary may be seen by summing the transmitted and reflected power waves:

$$\begin{aligned} P_{\mathrm{t}}U_{\mathrm{t}} + P_{\mathrm{r}}U_{\mathrm{r}} &= (1+H_{\mathrm{r}})P \cdot (1-H_{\mathrm{r}})U + H_{\mathrm{r}}P \cdot (-H_{\mathrm{r}})(-U) \\ &= [1 - H_{\mathrm{r}}^2 + H_{\mathrm{r}}^2]PU \\ &= PU, \end{aligned}$$

where P and U denote the Fourier transform of the incident pressure and volume-velocity waves, respectively.

The reed may be modeled as a signal- and embouchure-dependent *nonlinear reflection coefficient* terminating the bore. Such a model is possible because the reed mass is neglected.[9] The player's embouchure controls damping of the reed, reed aperture width, and other parameters, and these can be implemented as parameters on the contents of the lookup table or nonlinear function ("Reed Table" in Figure 9).

The controlling mouth pressure is denoted p_{m}. The *reflection coefficient* of the reed is denoted $\rho(h_\Delta^+)$, where $h_\Delta^+ \triangleq p_{\mathrm{m}}/2 - p_{\mathrm{b}}^+$ ("incoming half-pressure drop"). A simple choice of *embouchure control* is a simple additive offset in the reed-table address. Since the main feature of the reed table is the pressure drop at which the reed begins to open, such a simple offset can implement the effect of biting harder or softer on the reed, or changing the reed stiffness.

The following qualitatively chosen reed table has been used with good results [40]:

$$\hat{\rho}(h_\Delta^+) = \begin{cases} 1 - m(h_\Delta^{\mathrm{c}} - h_\Delta^+), & -1 \le h_\Delta^+ < h_\Delta^{\mathrm{c}} \\ 1, & h_\Delta^{\mathrm{c}} \le h_\Delta^+ \le 1 \end{cases}. \tag{11}$$

This function is shown in Figure 10 for $m = 1/(h_\Delta^{\mathrm{c}} + 1)$. The corner point h_Δ^{c} is the smallest pressure difference giving reed

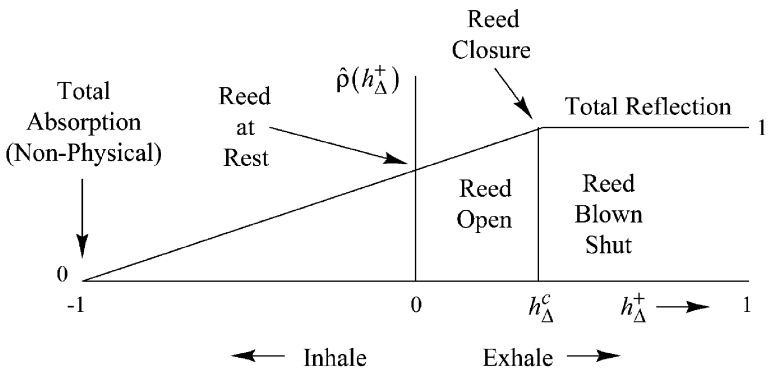


FIGURE 10 Simple, qualitatively chosen reed table for the digital waveguide clarinet. [From [40].]

[9]It is valid to neglect the reed mass when the first reed resonance is well above the fundamental frequency of the played note, as is normally the case.

closure. Having the table go all the way to zero at the maximum negative pressure $h_\Delta^+ = -1$ is nonphysical (0.8 would be a better physical estimate), but zero reflectance gives the practical benefit of opening the feedback loop when the lookup-table input signal is too large. Embouchure and reed stiffness correspond to the choice of offset h_Δ^c and slope m. Brighter tones are obtained by increasing the curvature at the point corresponding to the reed opening; for example, one can use $\hat{\rho}^k(h_\Delta^+)$ for increasing $k \geq 1$.

Another variation is to replace the table-lookup contents by a piecewise polynomial approximation. While less general, good results have been obtained in practice with this method [68].

An intermediate approach between table lookups and polynomial approximations is to use interpolated table lookups. Typically, linear interpolation is used, but higher order polynomial interpolation can also be considered [41].[10]

STK software [49] implementing a model as in Figure 9 can be found in the file `Clarinet.cpp`.

4 • Recommended Reading

An excellent musical acoustics textbook is Fletcher and Rossing [1]. For considerable detail on the physics of vibrating strings in musical acoustics, see the chapter by Vallette in **Mechanics of Musical Instruments** [16]. Classic references on acoustics include Pierce [69] and Morse and Ingard [70]. The original monograph **Vibration and Sound** by Morse [15] is still very interesting to read, and contains musical acoustics material that was dropped from Morse and Ingard.

References

1. N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, Second Edition, Springer Verlag, New York, 1998.
2. A. Hirschberg, J. Kergomand, and G. Weinneich, eds., *Mechanics of Musical Instruments*, Berlin: Springer-Verlag, 1995.
3. T. D. Rossing, F. R. Moore, and P. A. Wheeler, *The Science of Sound*, Third Edition, Addison-Wesley, Reading, MA, 2003.
4. J. O. Smith, "Virtual acoustic musical instruments: Review and update", *J. New Music Research*, 33(3):283–304, 2004.
5. J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, Springer Verlag, New York, 1976.
6. A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1999.

[10] http://ccrma.stanford.edu/~jos/pasp/Delay_Line_Interpolation.html.

7. D. Massie, "Wavetable sampling synthesis", in *Applications of DSP to Audio & Acoustics*, M. Kahrs and K. Brandenburg, Eds, pp. 311–341. Kluwer Academic Publishers, Boston/Dordrecht/London, 1998.
8. A. Horner, J. Beauchamp, and L. Haken, "Methods for multiple wavetable synthesis of musical instrument tones", *J. Audio Eng. Soc.*, 41(5):336–356, May 1993.
9. A. Ng and A. Horner, "Iterative combinatorial basis spectra in wavetable matching", *J. Audio Eng. Soc.* 50(12):1054–1063, 2002.
10. F. Fontana and D. Rocchesso, "Physical modeling of membranes for percussion instruments", *Acta Acustica*, 84(3):529–542, May/June 1998.
11. J. O. Smith, *Techniques for Digital Filter Design and System Identification with Application to the Violin*, PhD thesis, Elec. Engineering Dept., Stanford University (CCRMA), June 1983, CCRMA Technical Report STAN-M-14, http://ccrma.stanford.edu/STANM/stanms/stanm14/.
12. J. O. Smith, *Introduction to Digital Filters*, http://ccrma.stanfo - rd.edu/~jos/filters/, September 2005.
13. A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings, parts I and II", *J. Acoust. Soc. Am.*, 95(2–3):1112–1118, 1631–1640, Feb.–March 1994.
14. J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. O. Smith, "The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides", *J. Acoust. Soc. Am.*, 114(2):1095–1107, 2003.
15. P. M. Morse, *Vibration and Sound*, American Institute of Physics, for the Acoustical Society of America, http://asa.aip.org/publications.html, 1948, 1st edition 1936, author's last edition 1948, ASA edition 1981.
16. ibid [2].
17. H. Järveläinen, V. Välimäki, and M. Karjalainen, "Audibility of inharmonicity in string instrument sounds, and implications to digital sound synthesis", in *Proc. 1999 Int. Computer Music Conf., Beijing*, pp. 359–362, 1999. http://www.acoustics.hut.fi/~hjarvela/publications/.
18. M. E. McIntyre and J. Woodhouse, "On the fundamentals of bowed string dynamics", *Acustica*, 43(2):93–108, Sept. 1979.
19. T. W. Parks and C. S. Burrus, *Digital Filter Design*, Wiley, New York, June 1987, contains FORTRAN software listings.
20. B. Fornberg, *A Practical Guide to Pseudo-spectral Methods*, Cambridge University Press, 1998.
21. P. M. Ruiz, *A Technique for Simulating the Vibrations of Strings with a Digital Computer*, Music Master Thesis, University of Illinois, Urbana, 1970.
22. L. Hiller and P. Ruiz, "Synthesizing musical sounds by solving the wave equation for vibrating objects. Part II", *J. Audio Eng. Soc.*, Part I: 19(6):462–470, June 1971, Part II: 19(7):542–551, July/Aug. 1971.

23. J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth and Brooks, Pacific Grove, CA, 1989.
24. S. Bilbao, *Wave and Scattering Methods for Numerical Simulation*, Wiley, New York, July 2004.
25. M. Karjalainen, "Mixed physical modeling: DWG + FDTD + WDF", *Proc. 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY*, Oct. 2003, pp. 225–228, IEEE Press.
26. M. Karjalainen and C. Erkut, "Digital waveguides vs. finite difference schemes: Equivalance and mixed modeling", *EURASIP J. Applied Signal Processing*, 2004(7):978–989, June 15, 2004.
27. J. O. Smith, "Music applications of digital waveguides", Tech. Rep. STAN-M-39, CCRMA, Music Department, Stanford University, 1987, CCRMA Technical Report STAN-M-39, line feel http://ccrma.stanford.edu/STANM/stanms/stanm39/.
28. J. O. Smith, "Physical modeling using digital waveguides", *Computer Music J.*, 16(4):74–91, Winter 1992, special issue: Physical Modeling of Musical Instruments, Part I, http://ccrma.stanford.edu/~jos/pmudw/.
29. N. Lee and J.O. Smith, *"Virtual Stringed instruments"*, Dec. 2007, http://ccrma.stanford.edu/realsimple/phys_mod_overview/.
30. J. O. Smith, "Virtual Electric Guitar and Effects Using Faust and Octave", *Proc. 6th Int. Linux Audio Conf. (LAC 2008), paper:* http:// lac.linuxaudio.org/download/papers/zz.pdf supporting website: http://ccrma.stanford.edu/realsimple/faust_strings.
31. M. J. Beeson and D. T. Murphy, "RoomWeaver: A digital waveguide mesh based room acoustics research tool", *Proc. 2004 Conf. Digital Audio Effects (DAFx-04), Naples, Italy*, Oct. 2004, http://www.dafx.de/.
32. A. Krishnaswamy and J. O. Smith, "Methods for simulating string collisions with rigid spatial objects", *Proc. 2003, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY*, Oct. 2003, IEEE Press.
33. R. Pitteroff and J. Woodhouse, "Mechanics of the contact area between a violin bow and a string, part II: Simulating the bowed string", *Acta Acustica*, 84(3):744–757, 1998.
34. J. l. R. d'Alembert, "Investigation of the curve formed by a vibrating string, 1747", in *Acoustics: Historical and Philosophical Development*, R. B. Lindsay, Ed., pp. 119–123. Dowden, Hutchinson & Ross, Stroudsburg, 1973.
35. J. O. Smith, *Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications, 2nd edition*, W3K Publishing, http://w3k.org/books/, 2007, http://ccrma. stanford.edu/~jos/mdft/.
36. P. R. Cook, *Identification of Control Parameters in an Articulatory Vocal Tract Model, with Applications to the Synthesis of Singing*, PhD thesis, Elec. Engineering Dept., Stanford University (CCRMA), Dec. 1990, http: //www.cs.princeton.edu/~prc/.
37. J. L. Kelly and C. C. Lochbaum, "Speech synthesis", *Proc. Fourth Int. Congress on Acoustics, Copenhagen*, pp. 1–4, September 1962, Paper G42. Reprinted in [71, pp. 127–130].

38. H.-L. Lu, *Toward a High-Quality Singing Synthesizer with Vocal Texture Control*, PhD thesis, Elec. Engineering Dept., Stanford University (CCRMA), July 2002, http://ccrma.stanford.edu/~vickylu/thesis/.
39. J. O. Smith, "Principles of digital waveguide models of musical instruments", in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds, pp. 417–466. Kluwer Academic Publishers, Boston/Dordrecht/London, 1998.
40. J. O. Smith, "Efficient simulation of the reed-bore and bow-string mechanisms", *Proc. 1986 Int. Computer Music Conf., The Hague*, pp. 275–280, 1986, Computer Music Association, also available in [27].
41. J. O. Smith, *Physical Audio Signal Processing*, 2007, http://ccrma.stanford.edu/~jos/pasp/, Online book.
42. J. O. Smith, "Efficient synthesis of stringed musical instruments", *Proc. 1993 Int. Computer Music Conf., Tokyo*, pp. 64–71, 1993, Computer Music Association, also incorporated into [29].
43. J. Woodhouse, "Bowed string simulation using a thermal friction model", *Acta Acustica*, 89(2):355–368, 2003.
44. J. H. Smith and J. Woodhouse, "The tribology of rosin", *J. Mechanics and Physics of Solids*, 48(8):1633–1681, Aug. 2000.
45. S. Serafin, *The Sound of Friction: Real-Time Models, Playability and Musical Applications*, PhD thesis, Music Department, Stanford University, 2004.
46. S. Serafin, P. Huang, S. Ystad, C. Chafe, and J. O. Smith, "Analysis and synthesis of unusual friction-driven musical instruments", *Proc. 2002 Int. Computer Music Conf., Gothemborg, Sweden*, 2002.
47. F. Avanzini, S. Serafin, and D. Rocchesso, "Modeling interactions between rubbed dry surfaces using an elasto-plastic friction model", in *Proc. 2002 COST-G6 Conf. Digital Audio Effects (DAFx-02), Hamburg, Germany* pp. 111–116, September 2002, http://www.dafx.de/.
48. P. R. Cook, "Synthesis toolkit in C++, version 1.0", in *SIGGRAPH Proc.*, May 1996, Assoc. Comp. Mach.
49. P. Cook and G. Scavone, *Synthesis ToolKit in C++, Version 4.2.1*, Feb. 2006, http://ccrma.stanford.edu/CCRMA/Software/STK/.
50. J. Riionheimo and V. Välimäki, "Parameter estimation of a plucked string synthesis model using genetic algorithm with perceptual fitness calculation", *EURASIP J. Applied Signal Processing*, 3(8):791–805, July 2003, sound examples available at http://www.acoustics.hut.fi/publications/papers/jasp-ga/.
51. B. Bank, F. Avanzini, G. Borin, G. D. Poli, F. Fontana, and D. Rocchesso, "Physically informed signal processing methods for piano sound synthesis: A research overview", *EURASIP J. Applied Signal Processing*, 2003(10):941–952, Sept. 2003, http://www.mit.bme.hu/~bank/publist/jasp03.pdf.
52. V. Välimäki and T. Tolonen, "Development and calibration of a guitar synthesizer", *J. Audio Eng. Soc.*, 46(9):766–778, 1998.
53. M. Karjalainen, V. Välimäki, and T. Tolonen, "Plucked string models: From the Karplus–Strong algorithm to digital waveguides and beyond", *Computer Music Journal*, 22(3):17–32, Fall 1998, http://www.acoustics.hut.fi/~vpv/publications/ cmj98.htm.

54. J. Bensa, S. Bilbao, R. Kronland-Martinet, J. O. Smith, and T. Voinier, "Computational modeling of stiff piano strings using digital waveguides and finite differences", *Acta Acustica*, 91(2): 289–298, 2005.
55. D. Rocchesso and F. Scalcon, "Accurate dispersion simulation for piano strings", in *Proc. Nordic Acoustical Meeting (NAM'96), Helsinki, Finland*, June 12–14 1996, 9 pages.
56. M. Lang, "Allpass filter design and applications", *IEEE Transactions on Signal Processing*, 46(9):2505–2514, 1998.
57. J. Bensa, *Analysis and Synthesis of Piano Sounds Using Physical and Signal Models*, PhD thesis, Université de la Méditérranée, Marseille, France, 2003, http://www.lma.cnrs-mrs.fr/~bensa/.
58. M. Karjalainen and J. O. Smith, "Body modeling techniques for string instrument synthesis", *Proc. 1996 Int. Computer Music Conf., Hong Kong*, pp. 232–239, Aug. 1996, Computer Music Association.
59. C. R. Sullivan, "Extending the Karplus–Strong algorithm to synthesize electric guitar timbres with distortion and feedback", *Computer Music J.*, 14(3):26–37, Fall 1990.
60. E. Barbour, "The cool sound of tubes", *IEEE Spectrum*, pp. 24–35, Aug. 1998.
61. B. Santo, "Volume cranked up in amp debate", *Electronic Engineering Times*, no. 817, pp. 24–35, October 3, 1994, http:// www.trueaudio.com/at_eetjlm.htm.
62. M. Karjalainen, V. Välimäki, and Z. Jánosy, "Towards high-quality sound synthesis of the guitar and string instruments", *Proc. 1993 Int. Computer Music Conf., Tokyo*, pp. 56–63, Sept. 10–15 1993, Computer Music Association, http://www.acoustics.hut.fi/~vpv/publications/icmc93-guitar.htm.
63. V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical modeling of plucked string instruments with application to real-time sound synthesis", *J. Audio Eng. Soc.*, 44(5):331–353, May 1996.
64. J. O. Smith and S. A. Van Duyne, "Commuted piano synthesis", *Proc. 1995 Int. Computer Music Conf.*, Banff, pp. 319–326, 1995, Computer Music Association, http://ccrma. stanford.edu/~jos/pdf/svd95.pdf.
65. S. A. Van Duyne and J. O. Smith, "Developments for the commuted piano", *Proc. 1995 Int. Computer Music Conf.*, Banff, pp. 335–343, 1995, Computer Music Association, http://ccrma.stanford.edu/~jos/pdf/vds95.pdf.
66. G. P. Scavone, *An Acoustic Analysis of Single-Reed Woodwind Instruments with an Emphasis on Design and Performance Issues and Digital Waveguide Modeling Techniques*, PhD thesis, CCRMA, Music Dept., Stanford University, Mar. 1997, http://ccrma.stanford.edu/~gary/.
67. A. H. Benade, *Fundamentals of Musical Acoustics*, Dover, New York, 1990.
68. P. R. Cook, "A meta-wind-instrument physical model, and a meta-controller for real time performance control", *Proc. 1992 Int. Computer Music Conf., San Jose*, pp. 273–276, 1992, Computer Music Association.

69. A. D. Pierce, *Acoustics*, American Institute of Physics, for the Acoustical Society of America, 1989, http://asa.aip.org/ publications.html.
70. P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, McGraw-Hill, New York, 1968.
71. J. L. Flanagan and L. R. Rabiner, Eds, *Speech Synthesis*, Dowden, Hutchinson, and Ross, Stroudsburg, PA, 1973.
72. A. Askenfelt, Ed., *Five Lectures on the Acoustics of the Piano*, Royal Swedish Academy of Music, Stockholm, 1990, lectures by H. A. Conklin, Anders Askenfelt and E. Jansson, D. E. Hall, G. Weinreich, and K. Wogram. Sound example CD included. Publication No. 64, http://www.speech.kth.se/music/5_lectures/.

26

Modeling of Musical Instruments

Rolf Bader[1] *and Uwe Hansen*[2]

[1] *Institute of Musicology, University of Hamburg, Hamburg, Germany*
[2] *Department of Physics, Indiana State University, Terre Haute, IN, USA*

1 • Introduction

Signal processing techniques in acoustics address many concerns. Included are such things as wave propagation variables, amplitude considerations, spectral content, wavelength, and phase. Phase is primarily of concern when waves interact with each other, as well as with a medium, and the imposition of boundary conditions leads to normal mode vibrations. Such conditions are prevalent in all musical instruments, and thus relevant signal processing techniques are essential to both understanding and modeling the structure of musical instruments and the sound radiated. A coherent experimental technique for studying such normal mode vibrations is Modal Analysis, where the ratios of a series of FFT response spectra and associated excitations are recorded as transfer functions, and animated normal mode motion is subsequently extracted using appropriate coupled harmonic oscillator modeling. Finite Elements, a complementary theoretical technique is frequently also used to model such normal mode motion. This chapter is primarily concerned with such modeling and associated considerations.

1.1 Timbre

In musical acoustics, instruments can be classified by excitation mechanisms. On the one hand, drive mechanisms have an important influence on the spectra of the steady-state portion of the sound. On the other hand, the initial transients of the tones respond strongly to the nature of the drive mechanism (Bader 2002, Taylor 2001, Krumhansl 1989, Luce 1965). Musicians of all instrument families are aware of the importance of the attack and develop this important part of their artistic individuality. Thus, in terms of drive mechanisms, the following groups can be distinguished:

1. plucked instruments (guitars, lutes…)
2. bowed instruments (violins, double bass…)
3. reed-driven blown instruments (saxophone, oboe…)
4. air jet instruments (flute, flue organ pipes…)
5. lip-driven blown (brass) instruments (trumpet, horn…)
6. struck instruments (percussion, piano…).

Individual instruments of one given family sound different, of course, but the drive mechanisms cluster them into families. Thus, a western oboe is much closer in sound to a Persian zurna or a Thai saranai (all double-reed instruments) than to a western clarinet. Also, a western violin is closer in timbre to a Chinese erhu or an Indonesian rehab, (all bowed instruments) than to a western guitar. Individuals were asked to judge the similarity of pairs of instrument tones presented to them. Such hearing tests confirm instrumental clustering in families, depending on drive mechanisms (Grey 1977). While a true steady state is associated with a harmonic spectrum, departures from harmonicity in transients, such as attack and vibrato, reflect the nature of the mechanics of the transient and are thus instrument specific. Consequently, eliminating the initial transient from the hearing tests resulted in increased difficulty of separating and identifying the instrument type (Grey and Moorer 1977).

These perceptual experiments also found similarity judgments according to the acoustic properties of the sounds. The hierarchy is as follows

1. pitch
2. spectral centroid
3. spectral fluctuations
4. inharmonicities in the initial transient.

For the first parameter, subjects judged the similarity of sounds as far as pitch is concerned (Miller and Carterette 1975). This parameter is prominent in judging instrumental characteristics. If sounds are presented for which the pitch changes, listeners

neglect the other parameters when making similarity judgments. Second in importance is the spectral centroid, the "middle" frequency of the spectrum, which from a psychoacoustic standpoint relates to a bright sound for a high spectral centroid, or a dark sound for a low spectral centroid. Indeed, for constant pitch, the identification of musical instruments depends mostly on this spectral centroid. For example, violins have a high spectral centroid because of their resonances. Indeed, listeners will find artificial sounds with high spectral centroid much more similar to violins, than to plucked, or blown instruments.

The parameter third in importance for similarity judgment relates to spectral fluctuations exemplified by a vibrato, or amplitude fluctuations. Violins with their initial transients cut off are much better identified as violins as long as they are played with vibrato. Adding the vibrato, immediately leads to identification. The same holds for blown instruments, though not as strongly. It appears that a typical violin vibrato leads to identification for the instrument.

Finally, inharmonicities in the initial transients contribute to instrument identification. Inharmonicities appear in the attack for several, quite complicated reasons. In the guitar, for example, the resonance frequencies of the guitar body, are more prominent at the beginning of the tone. They are subsequently surpassed in loudness by the eigenfrequencies of the strings. For violin sounds, along with other causes, inharmonicities appear because of changes in bowing pressure, which lead the bow-string system to switch to another bifurcation regime. It has to be mentioned that the secondary role of inharmonicities in the initial attack of musical tones in terms of instrument similarity judgments stands in contrast to the fact that cutting the initial transient makes it hard to identify an instrument. So identification seems to be different from similarity judgments. And, as similarity tasks result in break-down when initial transients are eliminated, the influence on characterization and identification seems to increase in importance in comparison to the similarity feature.

All features mentioned above are associated with detailed structural features of musical instruments, thus the geometrical and physical reasons for the complex behavior of instruments during their initial transient phase need to be understood in order to discuss the character of instruments, including instrument quality, and to make suggestions to instrument builders for improvement or adjustments to the instrument geometry.

Thus modern musical acoustics is concerned with the geometrical features and with the detailed structure of musical

instruments. This fine structure is associated with fundamental aspects of instrument quality, to be described on an acoustical level.

1.2 Structure of Musical Instruments

Intricately detailed geometrical structures of musical instruments result in complex fine structures of the sounds produced. When classifying aspects of musical sounds in relation to specific musical instruments, four important features are noted:

1. timbre (sound quality)
2. attack behavior
3. overall loudness
4. degree of possible timbre variation.

Instrument builders and musicians are convinced that it is primarily these features which determine instrument quality and character. The importance of timbre is evident mostly within a given instrument family, rather than in comparison between families. A specific example is the classical guitar, where the timbre difference between Spanish and German instruments is quite small, but it is clearly discernable by experts. It is this difference which determines which type of instrument is to be played for a specific musical style, and furthermore, only that instrument type would be appropriate for that musical style.

As mentioned above, control of the attack behavior is very important. Generally, instruments with fast attack are judged higher in quality than instruments with slow or damped attack. One likely reason is that instruments with rapid response can be played faster and with greater ease. But external influences, such as humidity, may change this feature even for a given instrument. For example, violins played under conditions of high humidity react more slowly and are harder to play than under normal conditions. In the case of the violin, such considerations of individual instruments should, of course, include moisture absorption characteristics of bow hair and rosin. Furthermore, the attack transients of musical sounds contain characteristic instrument information, as well as information unique to the musical expression of the performing artist. This initial transient is short, in the range of about 50 ms, although a high-pitched saxophone sound may finish its attack after about 4 ms and a low pitched organ flue pipe may need 200 ms. In addition, the attack is very complex. Here good results are achieved by using modeling techniques to show the complexity as well as suggesting underlying reasons for the complexity

Builders of musical instruments generally attempt to maximize the overall loudness of the instrument. For example, the sound pressure radiated from a violin is relatively high, considering

the size of the instrument. An important element contributing to the radiation efficiency of the violin is the sound post. It forms a rigid connection between the front and back plates of the violin near the foot of the bridge on the treble side (Fletcher and Rossing 1999). In a violin without a sound post, the transverse bowing forces would preferentially excite normal plate modes which exhibit a nodal line along the length of the instrument in the middle. Because of their dipole nature, these modes exhibit rather low radiation efficiencies. The sound post limits the out-of-phase motion of the two sides of the violin thus rendering the overall radiation pattern more spherical, with increased radiation efficiency. Another example is found in some guitar designs, when the guitar body is designed in a fashion resembling a loudspeaker box, with the back plate and ribs very thick and rigid and a very thin top plate. Such a guitar is subjectively judged to be nearly twice as loud as a normal classical guitar, but the guitar timbre is changed and the degree of possible timbre variation by the player is decreased.

This degree of possible timbre variation produced by the player depends on the response of the instrument to loudness variations, as well as to playing position on the string with plucked and bowed instruments, or blowing pressure and mouth position with blown reed and brass instruments. Here, the player expects the instrument to not just behave in a linear fashion. For example, a change of lip pressure with a saxophone also changes the overtone content and thus the brightness of the sound. Moreover, the player may, expect an increase in lip pressure to have just a small influence on the sound in one pressure region and a dramatic effect in another region. This property of musical instruments also shows up in terms of abrupt phase changes. Tone production in reed instruments requires a threshold blowing pressure. Below this threshold pressure, no musical tone is produced, only a blowing noise. At a certain higher blowing pressure, the fundamental tone is no longer excited, and the instrument is overblown sounding in pitches of the harmonic overtone structure of the fundamental. Many instruments, like the horn or the trombone depend partly or completely on the possibility of sudden phase changes. But also small nonlinearities make instruments complex and individual. Some examples are the lip pressure mentioned above with saxophones, additional harmonic generation in strings with large displacement, or additional modes in the initial attack of the piano tone by reflection of the impulse wave at the piano hammer (Fletcher and Rossing 1999). It is this individuality, which among other things is responsible for the richness and quality of instruments.

Again, because many of these characteristics appear in the initial transients of sounds, musical acoustics, being concerned with musical problems, needs to take them into consideration.

1.3 Modeling using the Finite-Element Method (FEM) and the Finite-Difference Method (FDM)

On the one hand, this section is intended to illustrate typical problems and solutions of musical acoustics using geometric modeling of musical instruments. On the other hand, it provides an overview of the theoretical techniques of the Finite-Element Method **(FEM)** (Bathe 1996, Hughes 1987) and the Finite-Difference Method **FDM**. Thus, following the presentation of the FEM and FDM formulations, five basic analytical cases are presented:

Steady-state description: Turbulent $k-\epsilon$ model of a flute.
Eigenfrequency evaluation: Multiphonics in blown instruments as split-phase pressure description.
Time-dependent implementation: Complete geometry of a classical guitar.
Impulse-response technique: Guitar back plate timbre brightening by stress bending.
Contact problem: Church bell clapper time-dependent striking force.

For dynamic structures of complicated geometries, described by higher order differential equation which do not lend themselves to solutions in closed form, numerical methods such as *FDM* and *FEM* are used to obtain solutions. For *FEM*, the structure is subdivided into a mesh or grid, which, depending on the geometry, can consist of one-dimensional line elements, two-dimensional area elements, such as triangles of rectangles, or three-dimensional volume elements, such as regular or irregular tetrahedral or other polyhedra. Node points on the boundaries of these elements are used as discrete points on the structure under investigation. The dependent variables used in the equation systems are defined only at these node points. For a two-dimensional plate, for example, the dependent variables would give the displacement, which would only be known at the node points of the mesh. In addition to the structural and dimensional geometry and the mesh identification, mechanical properties such as density and elastic constants including Young's modulus, Poisson's ratio and shear modulus, as applicable, are needed along with appropriate boundary conditions. In *FDM*, the structure is not divided into a mesh of elements but rather discrete node points are identified on the geometry, and the equation system is solved for these points, considering differences with neighboring points. Thus, the *FEM* formulation is more complicated and involves more computational time. Its advantage

is the greater ease of handling boundaries. Furthermore, the choice of element geometry and element density in *FEM* is very flexible, making it very suitable for complicated geometries, in contrast to the usual fixed grid constants in *FDM*.

2 • Finite-Element Formulation

The elements used are most commonly defined as isoparametric elements. Here, an "ideal" element is defined with its own local coordinate system ranging normally from -1 to 1 as convenient integration and differentiation boundaries. Interpolation functions are defined in a way that when using the values of the dependent variable, i.e., the displacement of a plate, a continuous surface is achieved. So, for example, if a tetrahedral element were used, the element functions would be given as

$$N_1 = 1 - \xi - \eta - \zeta$$
$$N_2 = \xi$$
$$N_3 = \eta$$
$$N_4 = \zeta$$

Here, ξ, η, and ζ are local coordinate variables for the x-, y-, and z-direction and the element is defined for

$$-1 \leq \xi, \eta, \zeta \leq 1 \ . \tag{1}$$

This element has the advantage, that it can be precisely integrated analytically once for all elements used. Each element then must be transferred from the local to the global coordinate system using a Jacobian matrix

$$\mathbf{J} = \sum_{i=1}^{n} \mathbf{X}_i \ \frac{\partial N_i}{\partial \Xi} \ , \tag{2}$$

where $\Xi = \{\xi, \eta, \zeta\}$, $n = 4$ since there are four shape functions, and $\mathbf{X}_i$ are the coordinates in the global coordinate system, which is also called the reference frame. In addition to specifying the differential equation system, only the displacement at the node points needs to be known to obtain a complete interpolation solution for the geometry. This leads to a simple linear equation system which can be solved. For structural mechanics, the differential equation system is often based on a stress–strain formulation like

$$\sigma = E \ \epsilon \tag{3}$$

where σ is the stress, ϵ is the strain, and E is the Young's modulus. With higher dimensional problems, a matrix $\mathbf{C}$ is used which is the material matrix describing Young's modulus, the Poisson ratio, and the shear modulus. When using a three-dimensional model, this matrix is

$$\boldsymbol{C} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1 & \frac{\nu}{(1-\nu)^2} & \frac{\nu}{(1-\nu)^2} & 0 & 0 & 0 \\ \frac{\nu}{(1-\nu)^2} & 1 & \frac{\nu}{(1-\nu)^2} & 0 & 0 & 0 \\ \frac{\nu}{(1-\nu)^2} & \frac{\nu}{(1-\nu)^2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)^2} \end{bmatrix}. \quad (4)$$

Here E is the Young's modulus, which can also be formulated as different for all three directions, and ν is the Poisson's ratio. As the dependent variable is the displacement, the stress–strain equation can also be reformulated as

$$\mathbf{D}\ \mathbf{C}\ \mathbf{D}_{\epsilon}\mathbf{u} = \mathbf{F}\ , \quad (5)$$

where

$$\mathbf{D} = \mathbf{D}_{\epsilon}^{\mathrm{T}} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial z} \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial x} & \frac{\partial}{\partial z} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix} \quad (6)$$

is the differential matrix, $\mathbf{F}$ is an external force vector on each node point, and $\mathbf{u}$ is the discrete displacement vector

$$\mathbf{u} = \{u_1, v_1, w_1, \ldots u_m, v_m, w_m, u_n, v_n, w_n\}^{\mathrm{T}}. \quad (7)$$

Here, u is the displacement in the x-direction, v in the y-direction and w in the z-direction. If a matrix containing the shape functions of the isoparametric element is defined in the local frame as

$$\mathbf{H} = \begin{bmatrix} N_1 & N_2 & N_3 & N_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & N_1 & N_2 & N_3 & N_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_1 & N_2 & N_3 & N_4 \end{bmatrix}, \quad (8)$$

then a matrix of the derivatives can also be formulated as

$$\mathbf{B} = \mathbf{D}^{\mathrm{T}}\mathbf{H} \quad (9)$$

which allows differentiation of the shape functions analytically on a local level. This finally results in the following formulation

$$\sum_e \int_{\Omega_e} \mathbf{B}^{\mathrm{T}}\ \mathbf{C}\ \mathbf{B}\ d\Omega_e\ \mathbf{u} = \sum_e \int_{\Omega_e} \mathbf{H}\ \mathbf{f}^{\Omega_e} d\Omega_e + \sum_e \int_{B_e} \mathbf{H}\ \mathbf{f}^{B_e} \mathrm{d}B_e + \mathbf{f}^{P}\ . \quad (10)$$

Here, integration on the left side is over each element domain Ω_e and the summation is over all elements used in the global domain of the geometry. On the right side is the force vector $\mathbf{F}$ split up into three parts. There are forces $\mathbf{f}^{\Omega_e}$ on the level of the domain, forces $\mathbf{f}^{B_e}$ on the boundaries of the global domain and $\mathbf{f}^{P}$, forces on single node points. This can also be written in short like

$$\begin{aligned} &\mathbf{K} \quad \mathbf{u} = \mathbf{F}, \\ \text{with dimensions} \quad &[\mathrm{n} \times \mathrm{n}] \;\; [\mathrm{n}] = [\mathrm{n}], \end{aligned} \tag{11}$$

where $\mathbf{K}$ is called the stiffness matrix since it contains Young's modulus and the derivative matrices. In the static case, where, for example, an external force is acting on the domain, this linear equation system is solved for $\mathbf{u}$ when applying a force vector $\mathbf{F}$. This is still done most commonly with a Gauss elimination algorithm, although other algorithms have been proposed (Saad 2003). Still the Gauss algorithm seems to be the fastest with problems of structural mechanics. The above equation can also be transferred into a general eigenvalue system like

$$\mathbf{K}\,\mathbf{u} = \lambda\,\mathbf{M}\,\mathbf{u} \tag{12}$$

or in a time-dependent formulation like

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{D}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{F}\ . \tag{13}$$

Here, the mass matrix $\mathbf{M}$ is used and the dots denote derivatives with respect to time. Solving these equations can be quite complicated and iterative procedures may be used.

3 • Finite-Difference Formulation

The *FDM* is different from *FEM* as it is most commonly formulated as a time-dependent method. Nevertheless, *FDM* can also be put into a matrix form and solved like the *FEM*. On the other hand, time-dependent *FEM* problems are solved more like *FDM*. When having a time-dependent problem, the differential equation makes use of the definition of the derivative

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{\partial y}{\partial x}\ . \tag{14}$$

Here Δx and Δy are finite differences in the x- and y-direction of a coordinate system. So for the one-dimensional problem of a string, the differential equation of the string

$$c^2 \frac{\partial^2 y}{\partial x^2} = -\frac{\partial^2 y}{\partial t^2} \tag{15}$$

becomes

$$a_{n,t} = \frac{\partial^2 y}{\partial t^2} = -c^2 \frac{-2y_n + y_{n+1} + y_{n-1}}{\Delta x^2} \ . \tag{16}$$

This is the discrete version of the string problem with displacements y at node points n. The above equation holds for the node point n and calculates the mass weighted acceleration $a_{n,t}$ of node point n at time point t for a spacial grid constant of Δx. From this acceleration, the velocity of the next time step with width Δt is

$$v_{n,t_+} = v_{n,t} + a_{n,t} \, \Delta t \ . \tag{17}$$

Finally, the new displacement at this node point is

$$y_{n,t_+} = y_{n,t} + v_{n,t_+} \, \Delta t \ . \tag{18}$$

This is easy to implement. It is thus a very powerful tool for time stepping algorithms. It can be enlarged to incorporate complicated couplings between differential equations, and nonlinearities are quite easy to implement, too. Nevertheless, this method of solving is an explicit one. So overflow of the system may become a problem when more complicated equation systems are involved.

4 • Turbulent $k - \epsilon$ Model of Flute-Like Instruments

Flute-like instruments, such as the transverse flute, can be looked at in terms of impedance, which represents the resistance of the flute to the blowing of the player (Fletcher et al. 1982). The definition of the impedance Z involves the pressure p required to achieve a flow velocity v into

$$Z = \frac{p}{v} \ . \tag{19}$$

In experiments (Coltman 1968, 1973), it was shown, that just 3.5 % of the blowing energy of the player gets into the tube[1]. A ratio of energy going into the flute, to the energy radiated into the room outside the flute, can be obtained when modeling the tube of a flute, including the space outside the flute in the region

[1]Actually, from this 3.5% only 2.4% are radiated by the flute, so the overall efficiency of the flute is just 0.084%, a reason for its low sound level.

of the sound hole, using an inward flow. If this ratio is in the range of 3.5 %, our model is appropriate.

When modeling the flow entering the flute, it is appropriate to use a Navier–Stokes model of a momentum balance between flow û and pressure p like

$$\partial_t \hat{u}_i + \hat{u}_j \partial_j \hat{u}_i = -\frac{1}{\rho} \partial_i p + \nu \nabla^2 \hat{u}_i \tag{20}$$

where ϱ is the density of air, ν is its viscosity and the indexes i and j indicate spacial dimensions. In our two-dimensional model $i, j = 1, 2$, so u_1 and u_2 are the flow in x- and y-directions, respectively.

For a second equation, the incompressibility condition yields

$$\partial_t \hat{u}_i = 0 \ . \tag{21}$$

However, the flow into the flute is turbulent in nature. The oscillation at the blow hole makes the air flow change direction, which is the cause for additional turbulent viscous damping.

When modeling a turbulent field in FEM or FDM there are two ways to proceed. One way is to use a very fine mesh normally consisting of at least one million finite elements. Then all small eddies appearing in the flow are modeled in detail. This is very expensive in terms of memory and, what is more important, in terms of computing time. An alternative method is to use only a fairly rough mesh, but include turbulence in the mathematical equations modeling the system. Using the model with millions of elements would only require the use of the normal Navier–Stokes equation, since it is understood that this equation governs all flows provided it is modeled sufficiently precisely. When using the rough mesh, the most common model in use is Kolmogorov's k-ε model (Kolmogorov, 1941). Here it is assumed, that energy coming from the main flow of blowing is transferred into turbulent eddies. These eddies are no longer modeled in detail. Instead, a turbulent energy k is defined which takes over the energy of the main flow which is assumed to be taken as laminar. Kolmogorov's idea is that these turbulent eddies are getting smaller and smaller up to a smallest eddy size. This size is determined by the damping of the flow which allows no further splitting into ever smaller eddies. The turbulent energy then is thought of as being transported from the larger to the smaller eddies, resulting in turbulent energy dissipation e on the level of smallest eddy size, which then depends on normal viscosity. Since not all structural details of the turbulence are modeled, as mentioned above, a statistical time-averaged flow is used in the Navier–Stokes equation instead of the normal laminar flow.

The flow $\hat{u}$ of the Navier–Stokes model is split here into a mean flow U and a fluctuating flow u:

$$\hat{u} = U + u \ . \tag{22}$$

Inserting this in the Navier–Stokes equation results in the Reynolds-Averaged Navier–Stokes equation (RANS) (Durbin and Pettersson 2001)

$$\partial_t \ U_i + U_j \ \partial_j \ U_i = -\frac{1}{\rho} \ \partial_i \ p + \nu \nabla^2 \ U_i - \partial_j \ \overline{u_j \ u_i} \ . \tag{23}$$

The result is a normal Navier–Stokes equation with just one additional term on the right side, the turbulent damping term. A problem with this equation is due to averaging over the directions i and j because

$$\overline{u_j \ u_i} \neq \overline{u_j} \ \ \overline{u_i} \ . \tag{24}$$

This results in more variables than equations. Thus, the RANS is a so-called unclosed equation. To close it, several simplifications must be assumed and experimental values must be used (for details see Durbin and Pettersson 2001).

Comparing these two models, i.e., the Navier–Stokes and the Reynolds-Averaged Navier–Stokes model, conclusions can be drawn about the importance of turbulence and turbulent damping in the case of flute-like instruments.

Figures 1 and 2 show the results for the Navier–Stokes (NS) model and the Reynolds-Averaged Navier–Stokes (RANS) model, respectively. The NS model has a split of the blowing flow into the tube and outside the tube with 43% flowing inside the tube. The RANS model on the other hand just shows 0.55 % of the

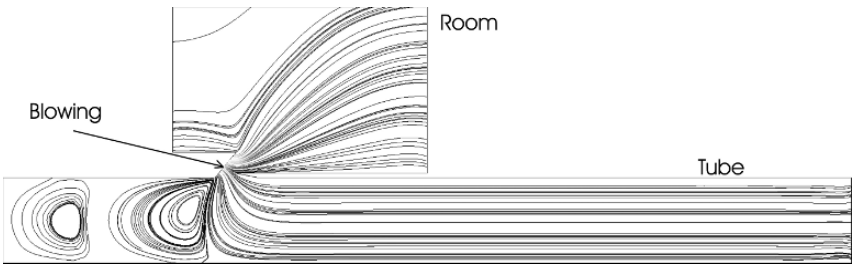


FIGURE 1 Two-dimensional Navier–Stokes flute model. The streamlines show the blowing flow starting at the embouchure hole, split into a tube part (*lower flow*) and a room part outside the flute tube (*upper part*). The split results in 43% of the flow occurring inside the tube.

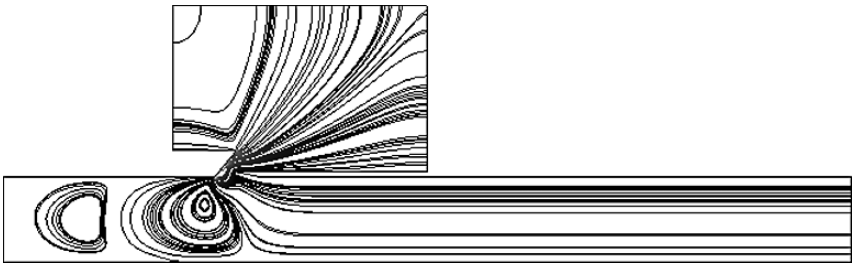


FIGURE 2 Two-dimensional Reynolds-Averaged Navier–Stokes (RANS) flute model. The streamlines again show a split flow as in Figure 1. But here, the split results in just 0.55% of flow inside the tube, which comes much closer to the experimental value of 2.4%.

flow making its way into the tube. This comes much closer to the experimental value of 2.4 %. In the Figures, this behavior can clearly be seen. Around the embouchure hole inside the tube, the NS model in Figure 1 shows less resistance (impedance) than the RANS model in Figure 2, where the flow into the tube is much more difficult. The reason for this is the additional turbulence with associated viscosity concerns, caused by the change of flow direction, which, unlike in the NS model, is not merely in one direction.

Thus, it is noted that turbulent viscous damping not only plays an important role in flute-like blowing behavior but also contributes to the reason for the flute's small amount of radiated sound energy. Additionally, by adjusting the geometry of the computer model, more effective flute impedance adjustments can be suggested to instrument builders.

5 • Multiphonics in Blown Instruments in Terms of a Split-Phase Pressure Description

Multiphonic sounds are mostly known from reed instruments like the saxophone or the clarinet (Gibiat 1988, 2000). Here, two or more tones are produced at the same time, a technique used in modern and free jazz, free improvisation, and contemporary art music. There are four ways of producing multiphonic sounds.

- overblowing
- underblowing
- complex sound hole closing patterns (fingerings)
- inhomogeneity in instrument geometry.

As mentioned above, blown instruments show abrupt phase changes for a linear change of a control parameter. Most prominently, the player's blowing pressure determines whether a "normal" tone is produced or whether the instrument is overblown (increased pressure) or underblown (decreased pressure). Normally, overblowing such an instrument causes it to sound with a tone of its overtone structure. But reed instruments have a second very important parameter of tone production, which is the pressure of the lips of the player on the reed. Since a change of this lip pressure with constant blowing pressure can change the pitch of a tone (up to three semitones) for each sound hole closing pattern (each played note), the lip pressure needs to be adjusted by the player in order not to detune the desired pitch. Now, when overblowing with increased blowing pressure, the lip pressure also has to be increased. If this is not done, and overblowing is done with normal lip pressure, the saxophone or clarinet will sound with more than one pitch. The same thing happens when the instrument is underblown.

Multiphonics may also occur with inhomogeneity of the instrument geometry. An example is the Chinese dizi flute. It has a membrane covering its highest sound hole. This membrane is not too tightly fixed, so it can vibrate to a certain extend. By coupling of this membrane to the air tube of the flute, a perfect bifurcation scenario occurs. The spectrum of the produced time series of radiated sounds by the flute consists of frequencies for which the following relationship holds

$$f = n\, f_1 \pm m\, f_2 \text{ with } n = 1, 2, 3 \ldots \text{ and } m = 0, 1, 2, 3 \ldots . \qquad (25)$$

This is the formula to which Backus refers for multiphonic sounds (Backus 1978). When m is zero, we have a normal harmonic overtone spectrum. With $m > 0$, two additional frequencies are obtained, differing from f_1 by $\pm m\ f_2$, which are reduced in amplitude. Since for each $n\ f_1$, these additional frequencies occur, not only additional inharmonic components are present, but harmonic overtone structures, consisting of frequencies like

$$f(k) = k\, f_1 \pm k\, f_2 \text{ with } k = 1, 2, 3 \ldots . \qquad (26)$$

For example, if $f_1 = 90$ Hz and $f_2 = 10$ Hz, then in the plus case the frequencies are $f(k) = 100,\ 200,\ 300$Hz….

The formula used by Backus holds perfectly in the case of inhomogeneous geometry, but it is not always appropriate in the case of complex sound hole closing patterns. Here, normally two tones are heard very prominently with additional inharmonic

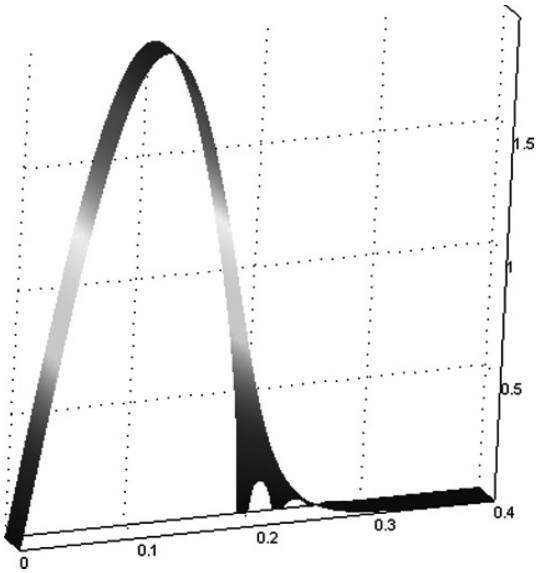


FIGURE 3 Pressure distribution of the first eigenfrequency of a tube with six sound holes open at the right end of the tube.

components around them (Keefe and Laden 1991). Inasmuch as most of the closing patters have many sound holes closed with an additional high-pitched open hole, the modeling of the tube in terms of eigenfrequencies gives insight into the pressure distribution along the tube for such complicated closing patterns.

Figure 3 shows the pressure distribution of the first eigenfrequency of a tube closed at both ends with six sound holes open at the right side of the tube. Note that the sound holes are just open at the upper side of the tube. Because of that, the pressure is not zero at the first open sound hole, but still has a finite value up to about the third open hole. This enlarges the size of the tube up to the sound hole a bit, so compared to the lowest possible frequency of the tube with all sound holes closed, the first overtone shown in Figure 3 is not in a perfect relation of 2:1 to the fundamental frequency.

Figures 4 and 5 now show the cases for tone hole closing patterns typical for multiphonic sounds. Of the six tone holes at the right side of the tube, the lower five are closed and just the highest tone hole is still open. First it is noted that the pressure is distributed over the entire tube. This is because the tone hole is just open at one side of the tube on the upper end. But this distribution can appear in two cases within just one eigenfrequency

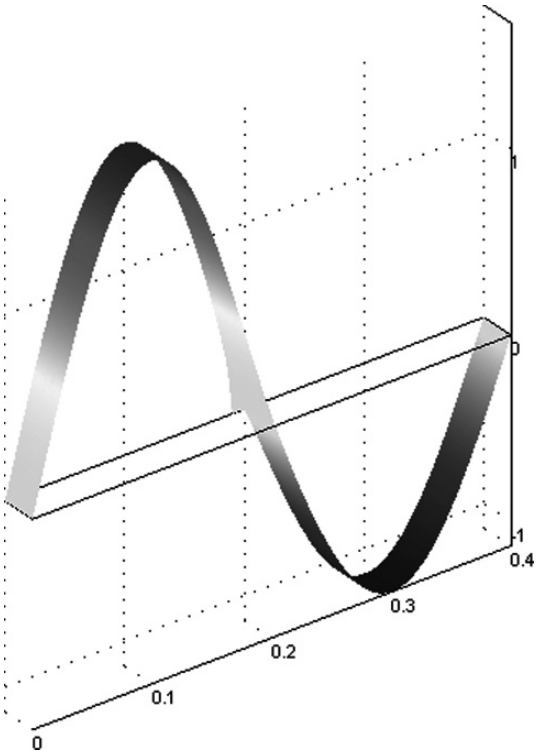


FIGURE 4 Pressure distribution of the anti-phase case of the first eigenfrequency of a tube with all sound holes closed except for one sound hole in the middle of the tube.

pressure distribution: one anti-phase, as shown in Figure 4, where the pressures of the two parts of the tube are of opposite sign; and one in-phase, as shown in Figure 5, where the pressures are of the same sign in the left and right part of the tube. Since this holds for every eigenfrequency of the tube, where at first there was only one frequency with each overtone, there now are two.

Table 1 shows the frequency ratios of the two cases of anti-phase and in-phase pressure distribution of the multiphonic playing technique for each partial. The ratios were calculated in relation to the fundamental frequency of each case. The anti-phase ratios correspond to the case of normal playing without complex sound hole closing patterns. It shows a harmonic overtone structure as expected. But astonishingly, the in-phase ratios of the additional frequencies caused by the multiphonic pattern form within themselves again a harmonic overtone structure. So here, two tones can be heard, which are in an

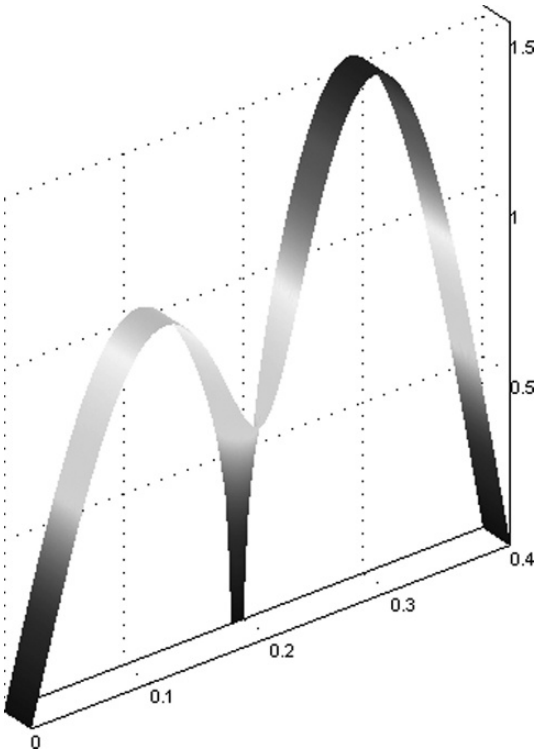


FIGURE 5 Pressure distribution of the in-phase case of the first eigenfrequency for a tube with all sound holes closed except for one sound hole in the middle of the tube.

TABLE 1 Frequency ratios for the anti-phase and in-phase cases, with each eigenfrequency of a tube played with a typical multiphonic sound hole closing pattern. The values are the relations to the fundamental frequency of each case. Not only the anti-phase relations form a harmonic overtone structure as in the normal playing condition but also the additional modes caused by the multiphonic playing technique form for themselves a harmonic structure and so an additional tone.

Partial	Anti-phase frequency ratio	In-phase frequency ratio
1	1.0	1.0
2	1.99966	2.02777
3	2.99984	3.08954
4	3.99972	4.15899

inharmonic relationship to one another. This is the case with multiphonic sounds caused by these complicated tone hole closing patterns, where two tones are heard most prominently.

6 • Time-Dependent Model of a Classical Guitar Body

Time-dependent solutions of complex geometries as shown here with a complete model of a classical guitar (Bader 2003, 2005) lead to a description of the complex development of the initial transient, (For the acoustics of the classical guitar see Fletcher and Rossing 1999, Elejabarrieta et al. 2002, Richardson and Roberts 1985, Meyer 1980). This complexity mainly comes from the inhomogeneous geometry of the guitar body which includes the top plate, the back plate, the ribs, the neck, the enclosed air, and the strings. Detailed modeling must of course include the bracing of the top and back plates, with particular attention to the top plate. However, considerable information about the role of each element can be obtained without that additional resolution. Interaction between these parts cannot be examined experimentally in their entirety. I.e., it is not possible to measure the in-plane waves of the plates. These waves are not radiated, but at discontinuities, like the boundaries of the plates, they are transformed into bending waves which are radiated. So a detailed knowledge of initial tone development leads to a variety of completely new insights into roles of the different guitar parts in the radiated sound. Furthermore, geometrical changes in terms of tone quality improvement can be examined and used by instrument builders.

In Figure 6 a classical guitar is shown as modeled by a Finite-Difference technique. The model distinguishes between bending waves with displacements in the direction normal to the guitar plates (top plate, back plate, ribs…), and so are radiated, and in-plane waves, which are nearly negligible as far as radiation is concerned. Inasmuch as these waves interact within the plates and at the boundaries, most prominently between top and back plate and ribs, and since, in-plane waves propagate about 8-12 times faster than bending waves, the overall development of the initial transient depends on these two coupled waves as described in Section 7. In addition, the ribs are looked at as a curved plate, which is shown in the coordinate system of Figure 6. This again leads to a steady transformation between bending and in-plane waves effecting the eigenfrequencies of the ribs and the temporal development of the sound radiated by the ribs.

Figure 7 shows the time series of the different parts of a classical guitar calculated by the Finite-Difference model. The

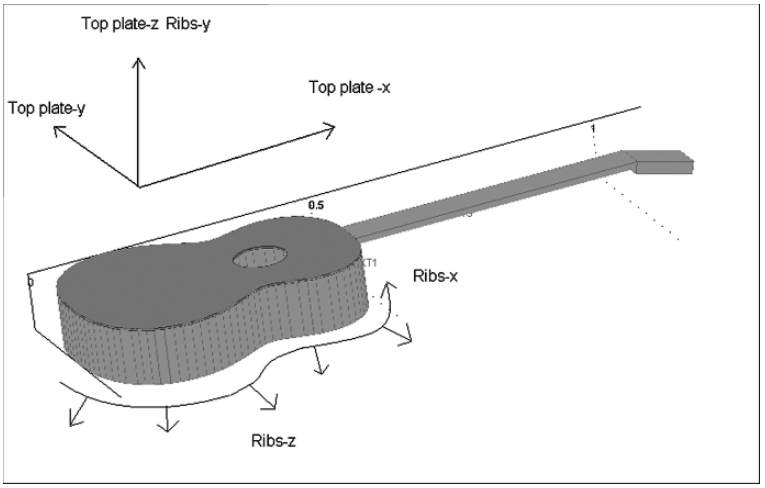


FIGURE 6 Finite-Difference model of a classical guitar with its coordinate system. The modeling of the ribs as a curved plate gives insight into the coupling between radiated bending and non-radiated in plane waves.

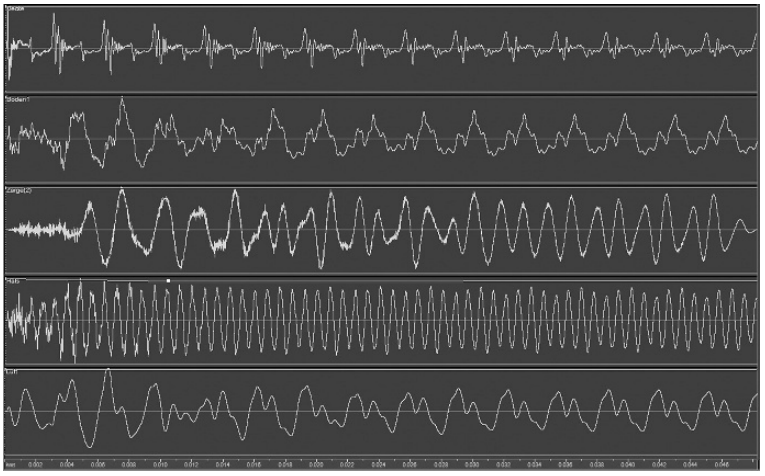


FIGURE 7 Time series calculated by the Finite-Difference model for the different guitar parts within the initial transient as radiated from these parts and recorded 1 m in front of the guitar. From top to *bottom*: top plate, back plate, ribs, neck, enclosed air. The roles of the different guitar parts in the overall sound show up clearly (see text for details), which can also be checked by ear.

open high e-string was virtually plucked (displaced in a trapezoidal shape) and the guitar vibrated "in slow motion" on the computer screen. This visualization aid is important for giving insight into the complex guitar motion since the number of degrees of freedom is too high to find the places and times of interest by pure numerical evaluation. At each iteration time point, the radiation from the different guitar parts was calculated by integrating over these parts with respect to a virtual microphone position 1 m in front of the guitar.

The different roles of the guitar parts in the initial transient show up clearly. The top plate is the only part which reacts nearly instantaneously. When listening to these calculated sounds, the top plate is the only one with a clear, hard attack, and since its sound is very bright, it is perceived as a hard and direct timbre. The back plate and the ribs need some time to come in. In the beginning, they still show some low amplitude vibrations caused by the coupling of the radiated bending waves to the in-plane waves. Although these initial vibrations are low in amplitude, they can still be heard because the frequencies fall in a very sensitive range of the human ear (around 1,000 Hz). The main attack of the back plate and the ribs is caused by the energy transport from the enclosed air, which is much slower in motion and so takes some time to displace the back plate and the ribs in the directions, which can cause radiation. The back plate and the ribs are prominent in a mid-frequency region and therefore responsible for what musicians call the tone color, the spiciness or taste of the sound. Next, the guitar neck is very high pitched as expected. But still this sound is needed in the complete guitar tone. Without it, some brilliance is missing. Finally, the enclosed air is responsible for a kind of boost or pressure impulse in the guitar sound, which can be heard by placing an ear (or microphone) next to the sound hole of the guitar. Some guitar recordings in the sixties or seventies (e.g., by Segovia) use this recording technique to achieve a very "earthy" or pumping sound.

Another aspect of the initial transient of a classical guitar is inharmonicities as discussed in Section 1.1 above about musical timbre. The pluck of a guitar string causes an impulse train which acts on the top plate. The first of these impulses is like a knocking on the plate, since no periodicity is established at this early point in time. This knocking, a kind of impulse, will drive many eigenfrequencies of the guitar body. Only because the impulse train has the constant periodicity of the played tone, the guitar body is forced to take over the string frequencies, and the eigenmodes of the body will disappear. Yet during the initial transient, some

eigenfrequencies, namely low ones, appear and interact with the frequencies of the string in an interesting manner.

Figure 8 shows a wavelet transformation (Haase et al. 2002) of a guitar tone (e_4) plucked on the high open E-string calculated by the Finite-Difference model. The temporal development of the frequencies is typical for an initial transient of a classical guitar. The fundamental and the first three overtones can clearly be seen. The typical part of a plucked instrument's initial transient is the disturbance of the string's fundamental frequency component by an eigenmode of the guitar body at the tone beginning. The component's amplitude drops down and comes up again. This is caused by a guitar body resonance frequency absorbing the energy of the string fundamental, as these two frequencies are close to each other. When the guitar body has taken the string mode energy, the guitar mode has to break down as well because it is driven by this string energy. As more than one eigenmode is driven during the initial phase, there is a kind of noise at the beginning of the tone, which can be perceived as "knocking". Indeed, if one knocks on the guitar top plate, the resulting sound

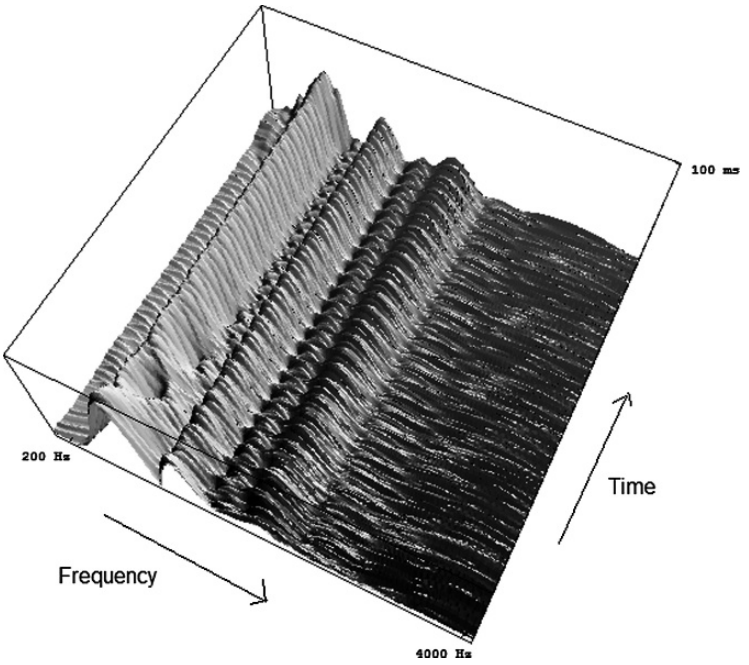


FIGURE 8 Wavelet-Transformation of a guitar tone (e_4) calculated by the Finite-Transient model.

is very similar to the noise during the guitars initial transient. So the eigenfrequencies of the guitar body are not only needed to amplify the string, but they also include a typical aspect of the sound, giving each guitar its own character. And, since we do not expect the frequencies driven by knocking to change, we do find that the development of the initial transient shown in Figure 8 for a particular classical guitar at a given pitch nearly always looks exactly the same, no matter which dynamic or articulation the player chooses.

So, in conclusion, since the initial transient is very important for the quality and character of the guitar, the exact development is of interest here. The complex structure of an instrument with ribs leads to an initial transient where different parts come in, one after another, and so enlarge the whole tone's beginning, causing the guitar to sound more rich and interesting. In addition, the first impulse of the string on the top plate drives many eigenfrequencies of the guitar body during the initial transient. As a rule of thumb, it can be said that the more complex the geometry of an instrument, the more complex the initial transient. This may be a reason for the development of plucked instruments with ribs. Lutes, widely used in ancient times and still in renaissance music, miss this complex initial phase and thus sound more light and free.

7 • Guitar Back-Plate Timbre Brightening by Stress Bending

Another technique for studying musical instruments is the use of impulse response. This corresponds to the knocking technique used by instrument builders, especially with the violin and the guitar, where a knock on the different plates produces their eigenfrequencies. Here, an example of a guitar back plate shall be shown. This flat wooden plate is built a little larger than the geometry of the ribs allow. When added to the ribs, it has to buckle out to fit in and so is under stress. It is claimed by guitar builders that this technique causes an increase of brightness of the guitar sound.

A possibility to model such buckling is the coupling between bending and in-plane waves in the back plate. This was discussed in the previous section. As the plate is too large in the in-plane directions, it accommodates the dimensional discrepancy by deformation in the transverse direction. The stresses applied to the in-plane direction, on the one hand, causes a strain in this in-plane direction, and, on the other hand, a strain in the bending direction.

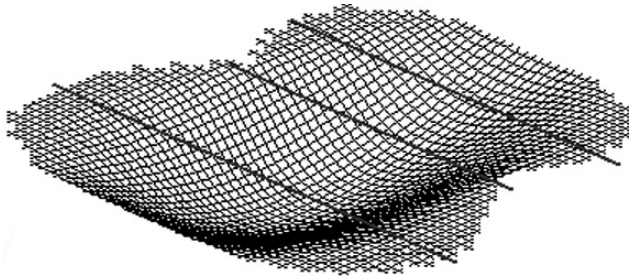


FIGURE 9 Curved back plate of the classical guitar Finite-Difference model. For this geometry, the coupling between bending and in-plane waves leads to an increase of the eigenfrequencies and also to an increased brightness as measured by an impulse response calculation.

If this model is appropriate, an increase of both the eigenfrequencies of the system and its brightness should result.

In Figure 9, the back plate is shown in its equilibrium position as curved inward by an applied force in the in-plane direction at the boundaries. The applied coupling results in the desired curvature. As the model is time-dependent, an impulse can be applied to the back plate. This impulse has a continuous spectrum with all frequencies present. The radiation from the back plate will then produce the eigenfrequencies of the plate. The amplitudes of these frequencies depend on the place of knocking, but the frequencies themselves do not change. However, when comparing the flat back plate to the curved, the place of knocking was held constant.

Table 2 shows the frequencies of the first five modes of the flat and buckled back plate. The buckled values are higher, as expected. Also the maximum spectral centroid value, shortly after the impulse was applied, was increased from 905 Hz for the flat case to 1,430 Hz with the buckled case. This finding was also expected from the experience of instrument builders and the difference in brightness and frequency can clearly be heard when listening to the calculated impulse responses.

TABLE 2 Frequencies of the first five modes of the back plate flat and buckled. The buckled values are higher

Mode	Back plate flat	Back plate buckled
1	229.44	239.04
2	420.48	439.68
3	480.96	500.16
4	652.8	681.6
5	912.0	931.2

8 • Church Bell Clapper Time-Dependent Striking Force

A bell clapper striking a church bell is a contact problem in terms of a Finite-Element Formulation (Wriggers 2002). As this is a time-dependent problem, the contact area between the clapper and the bell needs to be changed for each time step. At first, the clapper is not in contact with the bell. After the first contact, both objects are deformed and so the contact area grows larger. As soon as the velocity of the clapper changes direction and the clapper starts to move from the bell again, the contact area grows smaller until the clapper leaves the bell.

So, for each geometry, there is an external force F_{external}, which depends on the deformation and therefore on the distance A between the objects. Thus the contact condition can be formulated as:

$$F_{\text{external}} \begin{cases} f(\mathbf{u}), & \text{with contact } A \leq 0; \\ 0, & \text{without contact } A > 0; \end{cases} . \tag{27}$$

Here, during contact, the force generally is a function $f(\mathbf{u})$ of the displacement $\mathbf{u}$.

It has been shown that the contact of a church bell and its clapper is not quite a Hertzian contact and that the contact time is about 1 ms, which is surprisingly small (Fletcher 2002). The resulting impulse function in time of the clapper on the bell is of importance because it determines the amplitudes of the eigenfrequencies of the church bell being driven. An ideal Dirac delta impulse would drive all possible frequencies with the same amplitude. The more the impulse function flattens, the fewer higher frequencies are present.

In Figure 10 a schematic clapper, together with a part of a bell wall is shown. The clapper hits the bell with a speed of 100 m/s moving toward the left. The impulsive time function of the impact, integrated over the contact area of the clapper–bell, is shown in Figure 11, along with its spectrum in Figure 12. It can be seen, that the contact time is just a bit longer than 1 ms as measured in experiments.

In Figure 11, the contact time sequence is divided into two parts, left and right of the maximum. Inserted into the drawing are the values obtained when integrating over time for each segment. This clearly shows a greater value for the approach time than for the separation time. This seems to be because when the impulse function has passed its maximum, the bell rim moves with the clapper causing the impulse to be weaker. Figure 12 shows the spectrum of the impulse function. It has a periodicity of $1/T$ if T is the contact time of about 1 ms. Such a periodic spectrum is known from string instruments, where the periodicity

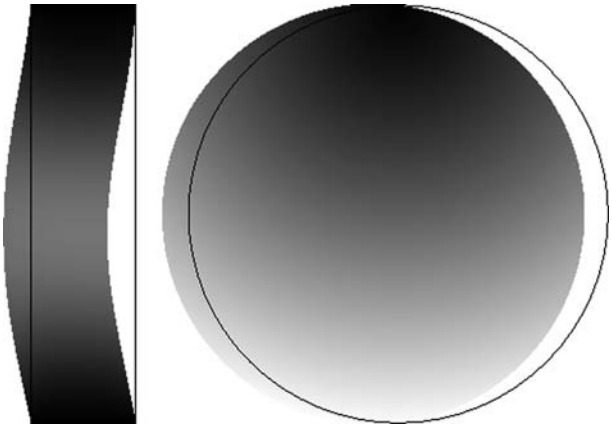


FIGURE 10 Schematic view of a clapper (right) acting on a bell rim (left). The clapper is shown here with a certain distance for clarity. The clapper moves left with an initial velocity and deforms the bell rim until the clapper slows down and moves back again. The overall process takes about 1 ms. The frame shown here corresponds to 0.7 ms after the beginning of the strike.

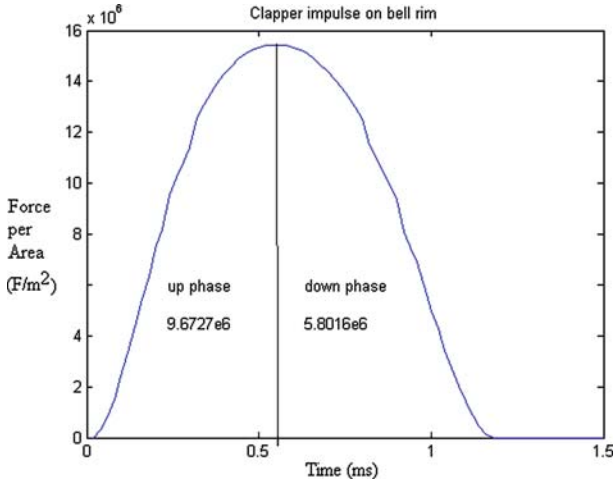


FIGURE 11 Clapper impulse time function on the bell rim integrated over the bell rim height. The impulse length is about 1 ms as expected. The impulse itself is stronger on the left side than on the right side, which can be seen when integrating the force over the area of the two sides as shown in the figure.

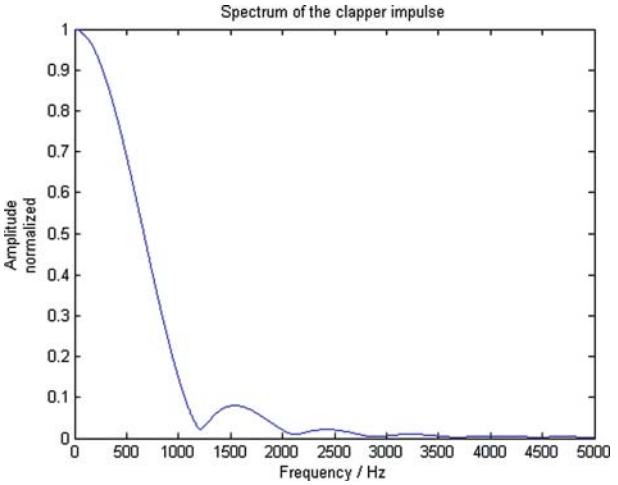


FIGURE 12 FFT spectrum of the clapper–bell rim contact of Figure 11. The spectrum shows a periodicity which comes from the finite length of the impulse forbidding frequencies with positive and negative amplitude parts being distributed equally over the whole impulse function time length. When performing the FFT, these frequency time series are multiplied (or weighted) by the sinusodial. As the impulse function is all positive in amplitude and the sinusodial has equally large positive and negative amplitude parts, the resulting amplitude of the frequency is zero.

depends on the point of plucking (Fletcher and Rossing 1999, p. 42). This seems to be because frequencies having equal positive and negative amplitude values over the impulse time cannot contribute to the final shape of the FFT plot, as their convolution integral over the impulse time, the Fourier integral over the impulse function is zero. As the bell can be thought of as a resonator in terms of impulse response, just the frequencies of the impulse can be excited. But this does not mean that the bell spectrum has to be of the same periodic shape as the clapper impulse FFT function as coupling between modes can occur at places of discontinuities and so can enlarge the spectrum.

References

Backus, J.: Multiphonic tones in the woodwind instruments. J. Acoust. Soc. Am. 63, 591–599, 1978.

Bader, R.: Fractal correlation dimensions and discrete-pseudo-phase-plots of percussion instruments in relation to cultural world view. In: Ingenierias, Octubre-Diciembre 2002, Vol. V. No. 17, 1–11.

Bader, R.: Physical model of a complete classical guitar body. Proc. Stockholm Musical Acoustics Conf. 2003, R. Bresin (ed.) Vol. 1, 121–124, 2003.
Bader, R.: Computational Mechanics of the Classical Guitar. Springer 2005.
Bathe, K.-J.: Finite Element Procedures. Prentice Hall, 1996.
Coltman, J.W.: Acoustics of the flute. Physics Today, 21, 11, 25–32, 1968.
Coltman, J.W.: Mouth resonance effects in the flute. J. Acoust. Soc. Am. 54, 417–420, 1973.
Durbin, P.A. & Pettersson, R.: Statistical Theory and Modeling for Turbulent Flows. Wiley, 2001.
Elejabarrieta, M.J., Ezcurra, A. & Santamaria, C.: Coupled modes of the resonance box of the guitar. J. Acoust. Soc. Am. 111, 2283–2292, 2002.
Fletcher, N.H.: Bell clapper impact dynamics and the voicing of a carillon. J. Acoust. Soc. Am. 111, 1437–1444, 2002.
Fletcher, N.H. & Rossing, Th.D.: The Physics of Musical Instruments. Second Ed. Springer 1999.
Fletcher, N.H., Strong, W.J. & Silk, R.K.: Acoustical characterization of flute head joints. J. Acoust. Soc. Am. 71, 1255–1260, 1982.
Gibiat, V.: Phase space representations of acoustical musical signals. J. Sound and Vibration, 123/3, 529–536, 1988.
Gibiat, V. & Castellengo, M.: Period doubling occurences in wind instruments musical performance. Acustica, 86, 746–754, 2000.
Grey, J.M.: Multidimensional perceptual scaling of musical timbres. J. Acoust. Soc. Am. 61(5), 1977, 1270–1277, 1977.
Grey, J.M. & Moorer, J.A.: Perceptual evaluations of synthesized musical tones. J. Acoust. Soc. Am. 62(2), 454–462, 1977.
Haase, M., Widjajakusuma, J. & Bader, R.: Scaling laws and frequency decomposition from wavelet transform maxima lines and ridges. In: Emergent Nature, M.M. Novak, (ed.): World Scientific 2002, 365–374.
Hughes, J.R.: The Finite Element Method. Linear Static and Dynamic Finite Element Analysis. Dover Publications, 1987.
Keefe, D.H. & Laden, B.: Correlation dimension of woodwind multiphonic tones. J. Acoust. Soc. Am. 90(4), 1754–1765, 1991.
Kolmogovov, A.N.: The local structure of turbulence in incompressible viscous fluid for vary large Reynolds number. Doklady Akademi Nauk SSSR, 30, 301–305, 1941.
Krumhansl, C.L.: Why is musical timbre so hard to understand? In: Structure and Perception of Electroacoustic Sound and Music. Nielzn, S. & Olsson, O. (ed.). Amsterdam 1989, 43–53, 1989.
Luce, D. & Clark, M.: Durations of attack transients of nonpercussive orchestral instruments. J. Audio Eng. Soc. 13, 194–199, 1965.
Meyer, J.: Verbesserung der Klangqualitt von Gitarren aufgrund systematischer Untersuchungen ihres Schwingungsverhaltens *Improvement of the sound quality of guitars by systematic investigations of its vibrating behaviour*. Physikalisch-Technische Bundesanstalt Braunschweig, Forschungsvorhaben Nr. 4490, 1980.

Miller, J.R. & Carterette, E.C.: Perceptual space for musical structures. J. Acoust. Soc. Am. 58(3), 711–720, 1975.

Richardson, B.E. & Roberts, G.W.: The adjustment of mode frequencies in guitars: A study by means of holographic interferometry and finite element analysis. In: Proc. Stockholm Musical Acoustics Conf. 1983, 285–302, 1985.

Saad, Yousef: Iterative Methods for Sparse Linear Systems, Cambridge University Press, 2003.

Taylor, Ch.: Sound. In: The New Groove Dictionary of Music and Musicians, S. Sadie (ed.). Second ed. 23, 759–776, 2001.

Wriggers, P.: Computational Contact Mechanics. Wiley, 2002.

PART V
SPEECH

Douglas D. O'Shaughnessy

INRS-Telecommunications, Montreal, Canada

27 Display and Analysis of Speech

Ben Milner 449

Introduction • Speech Production, Sampling and Quantisation • Display of Speech Waveforms • Analysis of Speech Waveforms • Conclusion

28 Estimation of Speech Intelligibility and Quality

Stephen Voran 483

Introduction • Estimators of Speech Intelligibility • Estimators of Speech Quality in Telecommunications • Conclusion

29 Gaussian Models in Automatic Speech Recognition

Jeff Bilmes 521

Introduction • Background • Training Regularized Gaussian Decompositions • Training Shared Regularized Gaussian Decompositions • A Generalized EM Procedure • Conclusions

30 Speech Synthesis

Thierry Dutoit and Baris Bozkurt 557

Concatenative Text-to-Speech Synthesis • Prosody Modification • Smoothing • Estimation of Unit Selection Costs • Voice Quality/Speaker Modification and Other Open Issues • Further Reading and Relevant Resources

31 Speech Coders

Roch Lefebvre and Philippe Gournay 587

Introduction • Speech Coding Concepts • Issues Related to Input Conditions • Issues Related to Channel Conditions • Current Developments and Future Challenges

27

Display and Analysis of Speech

Ben Milner

School of Computing Sciences, University of East Anglia, Norwich, Norfolk, UK

1 • Introduction

The three most important speech processing applications are speech recognition, speech synthesis and speech coding. The display and analysis of speech signals provide knowledge of the important characteristics of speech, facilitating the design of speech processing algorithms for these applications. For example, examination of speech signals allows important properties to be identified which can subsequently be used to discriminate between different speech sounds, which is necessary in both speech recognition and speech synthesis applications. Similarly, structures in speech allow parametric models of speech production to be formulated that are fundamental to recognition, synthesis and coding.

The primary acoustic correlates of speech are the voicing (whether the speech is voiced or unvoiced), the fundamental frequency (for voiced speech) and the frequency and shape of the formants. The extraction of these features has been the subject of research for many years and has led to the development of many forms of speech analysis. Of these, the three most commonly used are linear predictive coding (LPC) analysis, frequency domain analysis and cepstral analysis. These methods all enable voicing, fundamental frequency and formant analysis

although it is important to note that none of them are guaranteed to make error-free estimates of these features.

Some speech processing applications make use of these acoustic correlates directly. For example, both speech synthesisers and very low bitrate speech codecs require accurate fundamental frequency estimates in order to produce realistic sounding speech. In contrast to this, many other speech processing applications use derivatives of these fundamental acoustic features. For example, LPC analysis is frequently used in low-bitrate speech coding algorithms to provide vocal tract information rather than specific formant parameters. Similarly, speech recognition systems often extract vocal tract information using cepstral analysis but with a frequency scaling in line with the frequency response of human hearing.

This chapter examines three domains for displaying speech signals: time domain, frequency domain and time–frequency domain. Speech signals are shown in each of these domains and important features identified in each. Three common methods of speech analysis are then discussed. The discussion begins with LPC analysis and its application to both fundamental frequency and formant estimation. Frequency-based analysis is then reviewed and finally cepstral analysis. Finally, this chapter briefly discusses other features that can be extracted from speech signals that are applicable to speech processing applications such as recognition and coding.

However, before the discussion of the display and analysis of speech signals, the next section describes the procedure for converting an acoustic speech signal, as generated by the human speech production mechanism, into a digital signal. This is the necessary first stage of speech processing that a signal must go through to facilitate analysis by computer.

2 • Speech Production, Sampling and Quantisation

A speech signal is generated by the vocal organs in the human speech production mechanism which causes pressure fluctuations in the surrounding air. The speech begins as compressed air produced by the lungs, which in turn passes through the glottis where it can be altered in several different ways. If the vocal cords are under tension, the airflow will cause them to vibrate (at the fundamental frequency) and air will be released in the form of a quasi-periodic sequence of glottal pulses. If the vocal cords are in a relaxed state, this will result in turbulent airflow which generates wideband noise. The distinction between these two modes of operation is not always clear and can result in a

mixture of the two forms of signal being produced. A final mode of operation is when the vocal cords close completely and temporarily shut off the airflow. This is followed by a sudden release of air and leads to glottal stop sounds such as the phonemes /g/ and /k/.

The airflow emitted from the glottis is termed the excitation signal and passes through the vocal tract before being expelled as speech from the mouth, nose or both. Cavities in the vocal tract modulate the excitation signal and give rise to resonances, or formants, which define the particular sound being produced. Our ears respond to these pressure fluctuations through the movement of the ear drum. This in turn stimulates hair cells on the basilar membrane of the inner ear which causes neural impulses which we perceive as sound.

For computer-based processing of speech, a microphone is employed to receive the acoustic waveform and convert it to a time-varying voltage. This signal is continuous in both time and amplitude which is suitable for input into analogue devices, such as a tape recorder, but not for processing by computer. To convert this analogue signal into a digital form it must be sampled in time and quantised in amplitude. The remainder of this section discusses the sampling and quantisation processes with application to speech signals.

2.1 Sampling

The sampling operation converts the continuous-time analogue signal into a series of regularly spaced amplitude samples. From an analogue signal, $x_a(t)$, the sampling operation can be represented as a multiplication with an impulse train:

$$x_s(t) = \sum_{n=-\infty}^{\infty} x_a(t)\,\delta(t - nT). \tag{1}$$

The resulting sampled signal, $x_s(t)$, can then be converted into a discrete-time signal, $x(n)$, as

$$x(n) = x_s(nT). \tag{2}$$

An important consideration is the selection of the sampling period, T, which relates to the sampling frequency as $f_s = 1/T$. The sampling frequency places an upper limit on the bandwidth of the sampled, or discrete-time, signal. If a bandwidth of f_b Hz is required of the sampled signal, then the sampling frequency must be at least double that, i.e.

$$f_s \geq 2f_b. \tag{3}$$

The frequency $2f_b$ is also referred to as the Nyquist frequency and indicates the minimum sampling rate necessary to be able

to completely reconstruct the signal up to f_b Hz. The choice of sampling frequency is largely governed by the application. For example, telephony speech is sampled at 8 kHz to give a bandwidth of 4 kHz. Higher quality speech applications often increase the sampling frequency to 11 or 16 kHz. For the high-fidelity audio encoded on CD, the sampling rate is 44.1 kHz to give a bandwidth of 22.05 kHz. The downside of a high sampling frequency is the additional storage or processing necessary for the samples.

2.2 Quantisation

After sampling, the continuous-time signal has been converted into a discrete-time signal in the form a series of regularly occurring pulses of continuously valued amplitudes. To complete the digitisation process, these amplitude values must be quantised, which involves dividing the amplitude range of the signal into a set of M discrete levels. Assuming the dynamic range of the signal is from $-V$ to $+V$, then for a linear quantiser, the quantisation step size will be $\Delta = 2V/M$. To perform the quantisation, a set of $M+1$ decision boundaries, $\{b_0, b_1, \ldots, b_M\}$, is needed together with a set of M reconstruction levels, $\{y_1, y_2, \ldots, y_M\}$. For linear quantisation, the lowest decision boundary is set to the minimum amplitude of the signal, $b_0 = -V$. The remaining boundaries are separated by the quantisation step size to give the final boundary at the maximum amplitude level, $b_M = +V$. The reconstruction levels are normally set to the mid-point between two decision boundaries:

$$y_i = \frac{b_i - b_{i-1}}{2} \quad 1 \leq i \leq M. \tag{4}$$

Finally, the quantised amplitude, $x_q(n)$, is obtained:

$$x_q(n) = y_i \quad \text{iff} \quad b_{i-1} < x(n) < b_i. \tag{5}$$

The difference between the quantised and real amplitude values is termed the quantisation error, $q(n)$, and is related to the number of quantisation levels,

$$q(n) = x(n) - x_q(n), \tag{6}$$

and lies in the interval $[-\Delta/2, +\Delta/2]$. If it is assumed that the signal $x(n)$ is uniformly distributed, then the quantisation error is also uniform over this interval with probability density function

$$f_q(q(n)) = \frac{1}{\Delta}, \tag{7}$$

and the quantisation noise power, σ_q^2 can be computed:

$$\begin{aligned}\sigma_q^2 &= \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} q^2(n)\,\mathrm{d}q(n) \\ &= \frac{\Delta^2}{12}.\end{aligned} \tag{8}$$

Given that the signal lies in the interval [−V, +V] and again assuming a uniform PDF, the signal power, σ_x^2, can be computed in a similar way:

$$\begin{aligned}\sigma_x^2 &= \frac{1}{2V}\int_{-V}^{V} x^2(n)\,\mathrm{d}x(n) \\ &= \frac{V^2}{3}.\end{aligned} \tag{9}$$

From the quantisation noise power and signal power, the signal-to-quantisation noise ratio (SQNR) can be calculated:

$$\mathrm{SQNR(dB)} = 10\,\log\left(\frac{\sigma_x^2}{\sigma_q^2}\right) = 10\,\log\left(\frac{V^2}{3}\frac{12}{\Delta^2}\right). \tag{10}$$

Substituting into this the relationship between the number of bits used in quantisation, b, and quantisation step size, $\Delta = 2V/2^b$, leads to the following expression for the SQNR:

$$\mathrm{SQNR(dB)} = 6.02\mathrm{b\,dB}. \tag{11}$$

This shows that the SQNR is raised by 6.02 dB for each additional bit used in quantisation. However, it should be noted that in the derivation of this result the assumption of the speech amplitudes having a uniform distribution is far from realistic. This will be demonstrated later in this section. For speech processing applications, 8-bit quantisation is about the lowest acceptable level without introducing too much quantisation noise. For higher quality applications, such as CD encoding, 16-bit quantisation is generally used. An interesting experiment is to quantise a speech signal using only a single bit to give two quantisation levels (set at, for example, −V/2 and +V/2). Surprisingly, the resulting speech signal is still intelligible although of very poor quality due to the high levels of quantisation noise.

Several improvements can be made to the basic uniform quantiser which exploit the nature of the speech signal to improve

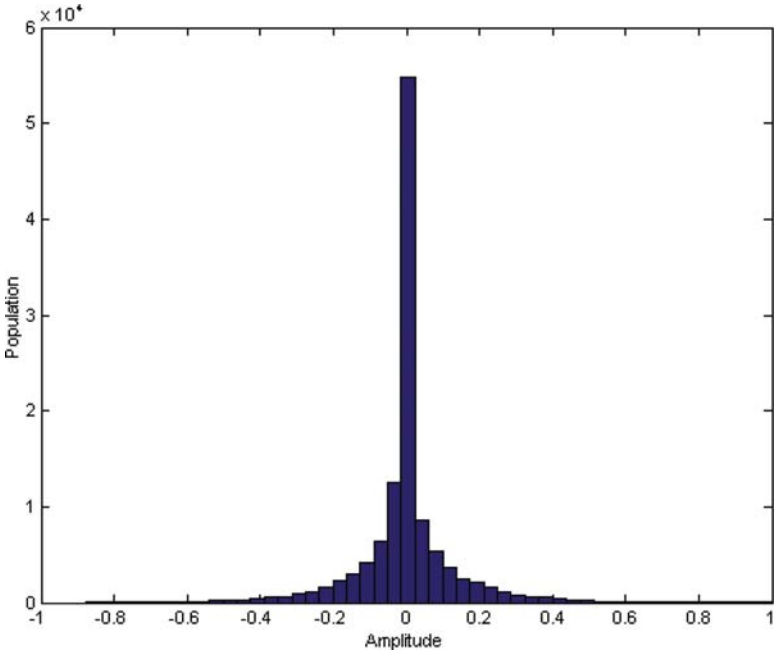


FIGURE 1 Histogram of speech sample amplitudes.

the bitrate to SQNR trade-off. Figure 1 shows a histogram of speech amplitudes taken from several sentences of conversational speech. The distribution of amplitude levels takes a Gaussian-like shape with considerably more small amplitude samples than larger amplitude samples. The large number of samples with close to zero amplitude arises from silence periods in the speech signal.

This implies that a uniform spacing of quantisation levels is not optimal for encoding a speech signal. A better quantiser would therefore use a non-linear spacing of levels which have a closer distribution at low-amplitude regions and coarser distribution at larger amplitude levels. Such quantisers are termed non-linear quantisers and can be designed by methods such as the Lloyd-Max algorithm which iteratively determines the optimal position for the decision boundaries and reconstruction levels based on an estimate of the probability density function (PDF) of the signal.

An alternative to having a set of non-linearly spaced quantisation levels is to apply a logarithmic compression to the amplitude samples. This transforms the amplitude levels into a more uniform distribution which allows standard linear quantisation to be applied without introducing large quantisation errors for larger amplitude signals. Upon de-quantisation a logarithmic

expansion is applied to return the amplitude values to the correct range. This is known as *companding* (compression-expanding) and gives a significant increase in SQNR. Several standard companders exist, such as A-law and μ-law, and these are employed in telephony-quality speech to give a SQNR using only 8 bits which is comparable to that obtained using 12-bit linear quantisation.

3 • Display of Speech Waveforms

This section examines three methods of displaying a speech signal, namely time domain, frequency domain and time–frequency domain. In each domain, particular attributes of the speech signal are identified and relationships shown between the different representations.

3.1 Time Domain Representation

The time domain representation is the most straightforward method of displaying a speech signal and shows the amplitude of the signal against time or sample number. Figure 2 illustrates the time domain representation of the speech waveform "*I think that you air your ideas very challengingly*", which has a duration of about 3 s.

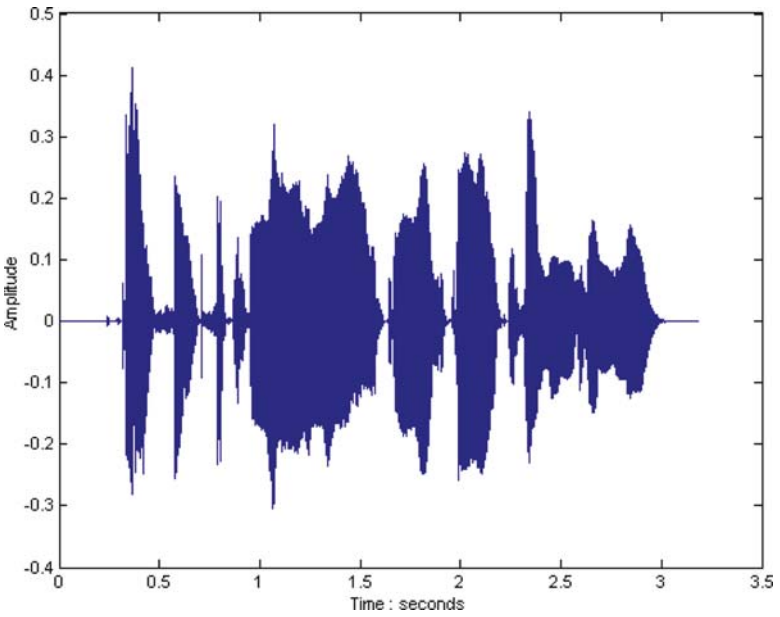


FIGURE 2 Three-second time domain representation of the utterance "*I think that you air your ideas very challengingly*".

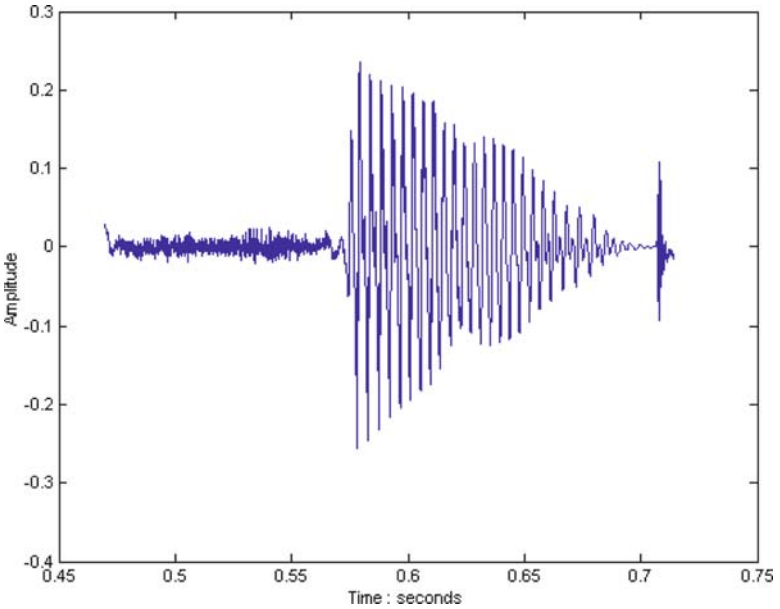


FIGURE 3 0.25-second duration time domain representation of the word "*think*".

The long time duration of this signal means that fine details of the speech waveform cannot be seen, but instead a good representation of the energy contour is present which provides a useful indication of silence and speech periods. For example at the beginning and end of the utterance can be observed silence periods of about 0.25 s each. Very short silences can also be seen between some of the words in the utterance. To contrast this, Figure 3 shows a much shorter duration segment, taken from the utterance shown in Figure 2, of the word "*think*". This begins at 0.45 s and lasts for about 0.25 s.

More detail is now visible and in particular the different regions of the speech waveform which correspond to the phonemes which make up the word. The unvoiced phoneme /th/ is recognisable at the start of the word as a noise-like signal. This is followed by the two voiced vowels, /ih/ and /ng/. Their voiced nature is identified by the pitch pulses which dominate that region of the signal. In fact an examination of the time domain waveform in this region would allow measurement of the pitch period. However, the time domain representation does not allow the boundary between the two phonemes to be easily located (this actually occurs at 0.63 s). At the end of the waveform, the structure of the unvoiced stop phoneme /k/ is evident. The

beginning of the phoneme /k/ is identified by no speech energy resulting from the closure of the glottis in order to build up subglottal pressure. The sudden release of air pressure as the glottis opens is also clearly evident at the end of the phoneme.

In general, the time domain waveform is useful for identifying different classes of speech sound (voiced, unvoiced, silence) but shows insufficient detail to discriminate between phonemes within a particular class. Measurement of the pitch period is also possible from the time domain representation through the location of pitch pulses in the waveform.

3.2 Spectral Domain Representation

The time domain representation of speech enables certain aspects of the signal to be identified, but it is more common to transform the signal into the frequency (or spectral) domain for examination. Historically, this conversion to the frequency domain has been implemented using a series of band-pass filters. The subsequent amplitude outputs of these filters provided a measure of energy in the various frequency regions of the signal. With digital signal processing, an easier method is to transform the signal into the frequency domain using a discrete Fourier transform (DFT). In applying the DFT to a signal, an assumption is made that the signal is stationary. As Figures 2 and 3 have shown, speech is highly non-stationary, and therefore to satisfy this stationarity assumption of the DFT, only a short-time window of the signal should be transformed. The length of time the speech signal remains stationary varies and depends on the particular speech sound under analysis. In general, window widths in the range 10–32 ms are used in speech processing applications.

The shape of the windowing function can take many different forms, and this has an impact on the resulting spectral representation of the signal. The simplest window is the rectangular window, but this introduces severe discontinuities at the beginning and end of the frame of speech under analysis which affects the spectral representation. More commonly used windows are the Hamming and Hanning windows which have much smoother profiles than the rectangular window and are based on cosine functions.

Assuming a time domain signal, $x(n)$, and windowing function, $w(n)$, the complex frequency spectrum, $X(k)$, is computed using the DFT as

$$X(k) = \sum_{n=0}^{N-1} w(n)\, x(n)\, \mathrm{e}^{-\mathrm{j}2\pi kn/N}. \tag{12}$$

Given that the window contains N time domain samples which were sampled at a rate of f_s samples per second, the DFT

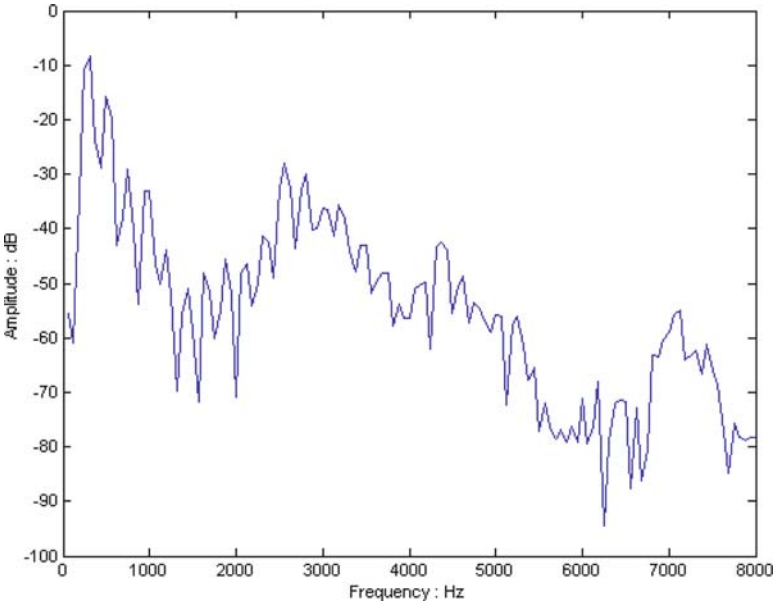


FIGURE 4 Log magnitude spectrum of vowel sound /ih/ taken from a 32-ms window.

coefficients, or DFT bins, $X(k)$, are spaced at f_s/N Hz intervals. For display purposes, it is more useful to convert the complex frequency spectrum into the magnitude spectrum, $|X(f)|$, and phase spectrum, $\angle X(f)$. It is also common to display the log of the magnitude spectrum as this reduces the amplitude of dominant regions of the spectrum which allows examination of lower energy regions. For illustration, Figure 4 shows the log magnitude spectrum of a 32-ms Hamming windowed frame of speech taken from the mid-point of the vowel /ih/ shown in Figure 3.

The figure reveals an important property of speech signals in that the magnitude spectrum can be described in terms of two components: the broad spectral envelope and the finer spectral detail (see Figure 18 for more discussion of this). The spectral envelope shows four spectral peaks (centred at 500, 3,000, 4,200 and 7,000 Hz) which are produced from modulation effects occurring within the vocal tract. These are termed formants and are characterised by their frequency, amplitude and bandwidth. The rapid fluctuations in the magnitude spectrum arise from the fundamental frequency and its harmonics. Changing the fundamental frequency of the speech signal would alter the positions of the harmonics but would leave the spectral envelope broadly unaltered.

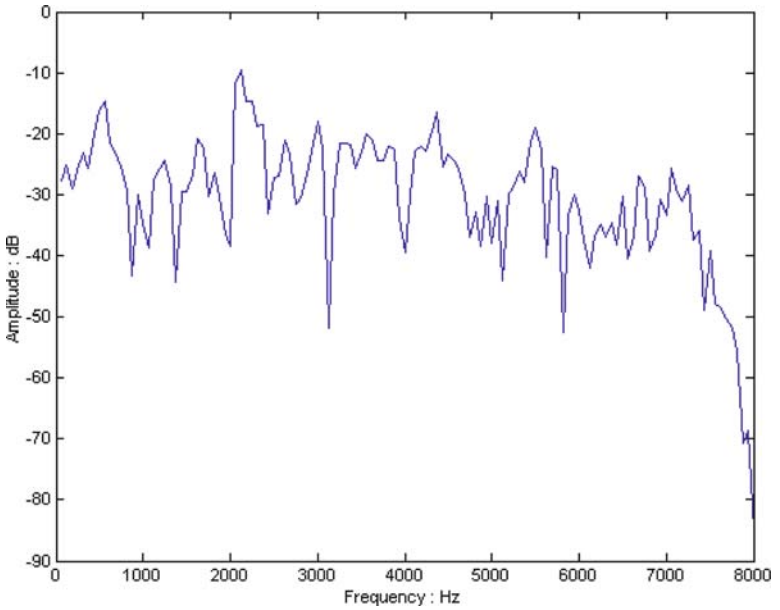


FIGURE 5 Log magnitude spectrum of unvoiced sound /th/ taken from a 32-ms window.

The time domain signal in Figure 3 shows that the unvoiced phoneme /th/ had very different structure to the two voiced phonemes /ih/ and /ng/. These differences can also be examined in the frequency domain, and to illustrate this, Figure 5 shows the log magnitude spectrum of the phoneme /th/ which has been computed using a DFT with a 32-ms window centred at 0.54 s.

The spectrum shows that much less structure is present in the signal. No formants are present and the harmonics no longer exist due to the excitation signal being wideband noise rather than a series of glottal pulses.

In the time domain waveform in Figure 3, it was not obvious that two different voiced phonemes were present and it was certainly difficult to analyse their structure. In the frequency domain, it is much easier to see these differences. For example, Figure 6 shows the log magnitude spectrum of the voiced phoneme /ng/ taken from a 32-ms window centred at time 0.65 s.

Comparing this spectrum with that of the vowel /ih/ shown in Figure 4 shows significant differences in the spectral envelope. The spectral envelope of the vowel /ng/ only contains one formant which has a frequency of 500 Hz in comparison to the four

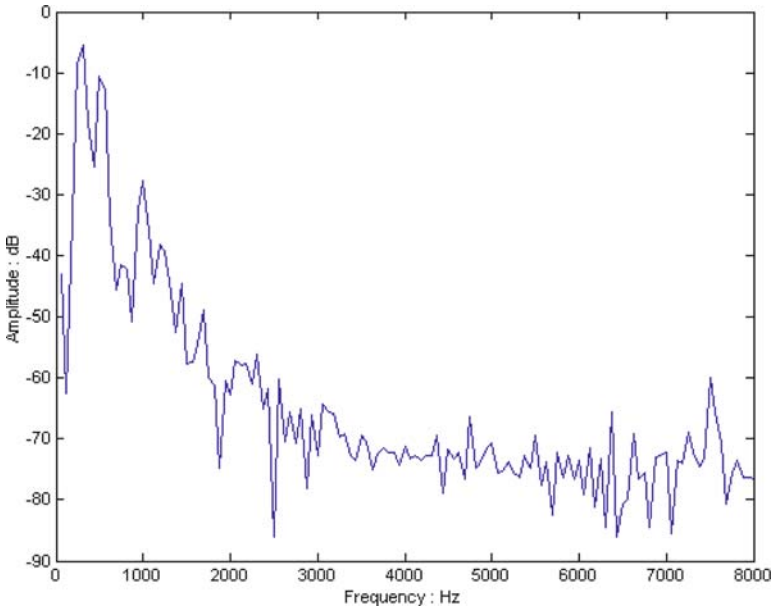


FIGURE 6 Log magnitude spectrum of vowel sound /ng/ taken from a 32-ms window.

formants of the vowel /ih/. In the time domain, these waveforms looked similar, but in the frequency domain very obvious differences in formant structure can be observed. It is these differences in formant structure which give the different sounds to the phonemes.

When using the DFT, it is important to consider the effect of different window widths on the spectral representation of speech. Selecting a large window width of time domain samples gives a very detailed spectral representation of the signal as the frequencies represented by the DFT bins will be very close together. However, the large number of time domain samples may mean that the signal is not stationary over this duration. Certainly, if the speech has changing characteristics within the window, then this will lead to a smearing of the resulting spectrum. This can be avoided by reducing the window width, but the downside to this is a reduction in spectral resolution. For example, halving the window width will double the frequency spacing between the spectral bins. A trade-off therefore exists between the window width and the desired spectral resolution of the signal that the application or analysis requires.

3.3 Time–Frequency Representation

The previous sections have shown that the time domain representation is useful for showing the time-varying nature of a signal, while the frequency domain representation enables the structure of a short-time stationary window of speech to be examined. These two representations can be combined into a time–frequency representation termed a spectrogram. Figures 8 shows an example spectrogram of the waveform *"I think that you air"* which was shown in Figure 2. Time is shown along the *x*-axis of the spectrogram and frequency along the *y*-axis. The darkness indicates the energy of the signal at a particular time and frequency. Examination of the spectrogram shows the structure of the speech signal in terms of formants, fundamental frequency and voicing.

The spectrogram is constructed from a series of short-time Fourier transforms which are successively computed along the time domain waveform. As the analysis window slides along the signal, the log energy within each DFT bin is converted into a shade of colour which is then positioned at the appropriate time and frequency. This operation is illustrated in Figure 7.

The resolution of the spectral analysis leads to two different forms of spectrogram. Spectrograms showing the broad spectral envelope are termed wideband spectrograms and those showing the finer harmonic structure are termed narrowband.

3.3.1 Narrowband Spectrogram

Figure 8 shows the narrowband spectrogram of the utterance *"I think that you air"*. For the narrowband spectrogram, the Fourier transform is computed from a relatively long-duration

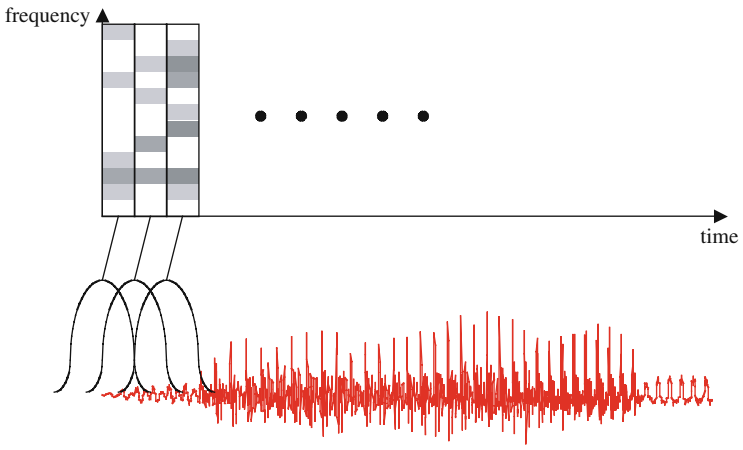


FIGURE 7 Construction of a spectrogram from a time domain waveform.

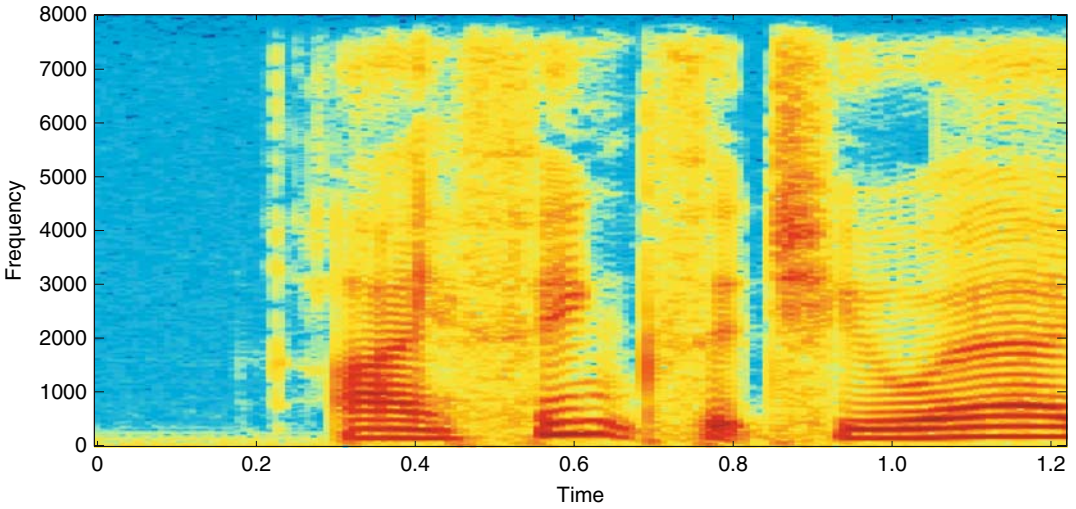


FIGURE 8 Narrowband spectrogram of utterance *"I think that you air"* together with the positions of the phonemes making up the utterance.

time window which gives a detailed spectral representation at the expense of less precise temporal localisation. The spectral bandwidth of the spectrogram shown in Figure 8 is 31.25 Hz. This allows individual frequency harmonics of the glottal pulses to be resolved leading to the characteristic narrow horizontal bands which make up the narrowband spectrogram. Examining a vertical slice of the spectrogram is equivalent to the magnitude spectrum plots shown in Section 3.2.

Voiced regions can be identified by the presence of pitch harmonics such as in the phonemes /ih/ and /ng/ which are shown in Figures 4 and 6. Unvoiced phonemes are characterised by the lack of pitch harmonics such as the phoneme /th/ which exhibits the noise-like structure shown in Figure 5. Silence regions are characterised by a lack of energy as is seen before the speech begins.

3.3.2 Wideband Spectrogram

Figure 9 shows the wideband spectrogram of same utterance shown in Figure 8. The wideband spectrogram is computed in a similar manner to the narrowband spectrogram except that now a much shorter time frame of speech is analysed by the Fourier transform. This results in a smoother spectral representation which shows the spectral envelope but not the individual pitch harmonics. In fact individual glottal pulses can be observed as vertical bands in the spectrogram. The short duration of the analysis window also improves the temporal

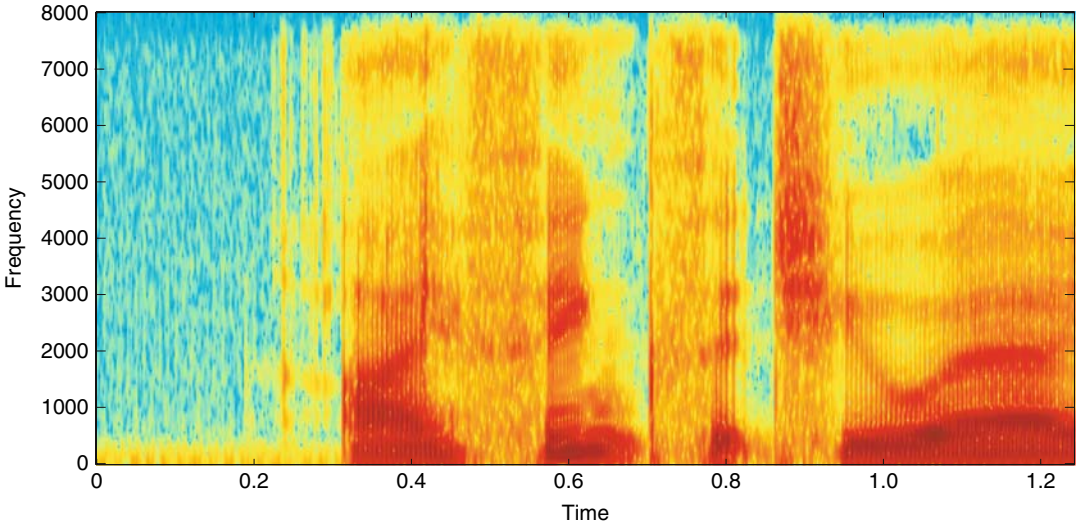


FIGURE 9 Wideband spectrogram of utterance *"I think that you air"* together with the positions of the phonemes making up the utterance.

resolution of the wideband spectrogram in comparison with the narrowband spectrogram. The spectral bandwidth of the wideband spectrogram is around 300 Hz, which means that several harmonics of the glottal pulses fall in each frequency band to give the spectral averaging leading to the spectral envelope.

The use of the wideband spectrogram is preferable when performing phonetic analysis of speech as the effect of the fundamental frequency and its harmonics is removed leaving only formant structure. More accurate temporal localisation is also possible as a result of the shorter time frames. This is demonstrated by the very narrow representation of the glottal stop phoneme /k/ shown in the wideband spectrogram in comparison to the more blurred representation in the narrowband spectrogram.

4 • Analysis of Speech Waveforms

Speech analysis takes many different forms and calls upon a variety of different signal processing methods. The classification of speech as voiced or unvoiced, the estimation of the fundamental frequency and the identification of formants are commonly required parameters which can be obtained through speech analysis. More specific speech features can also be extracted that are related to a particular speech processing application. For example, there are a wide variety of feature

extraction algorithms that can be used in the initial stages of speech recognition systems. These extraction methods use various forms of speech analysis to extract features that can discriminate between different speech sounds. In this section, three important methods of speech analysis are discussed and also their application to fundamental frequency and formant estimation. The section is completed with a brief examination of feature extraction methods used in speech recognition, coding and synthesis.

4.1 Linear Predictive Coding (LPC) Analysis

A commonly used model of speech production is the source–filter model which finds applications in low-bitrate speech coding, speech recognition and speech synthesis. The model comprises two parts – source and filter. The source models the excitation signal from the lungs and vocal cords and can take two forms. For voiced speech, the excitation is modelled as a series of impulses, each separated by the pitch period, which represent the glottal pulses produced as the vocal cords open and close. For unvoiced speech, the excitation takes the form of white noise which simulates the airflow when the vocal chords are relaxed leading to turbulent airflow. A separate gain term is used to amplify the excitation signal, and this reflects variations in speech amplitude that can be produced principally by varying the pressure generated in the lungs. The filter component of the model represents the frequency response of the vocal tract. This is usually modelled by an all-pole filter which can reproduce the formant-based resonances in the vocal tract. The all-pole structure is effective for most speech sounds, although to model the anti-resonances found in nasal sounds zeros are also required in the filter. The need for a filter with both poles and zeros increases the complexity of the model and can generally be avoided by increasing the order of the all-pole filter.

The examination of waveforms in the previous section has shown that speech is highly non-stationary. This means that the parameters of the vocal tract model must themselves be time varying. To allow for this the LPC analysis is performed on sequences of short-time frames of speech which are extracted progressively from the utterance. A typical analysis frame rate is 10–30 ms. Figure 10 illustrates the source–filter model of speech production.

The output speech signal, $x(n)$, can be expressed in terms of the excitation signal, $u(n)$, the gain, G, and the vocal tract filter coefficients a_k:

$$x(n) = \sum_{k=1}^{p} a_k x(n-k) + Gu(n). \tag{13}$$

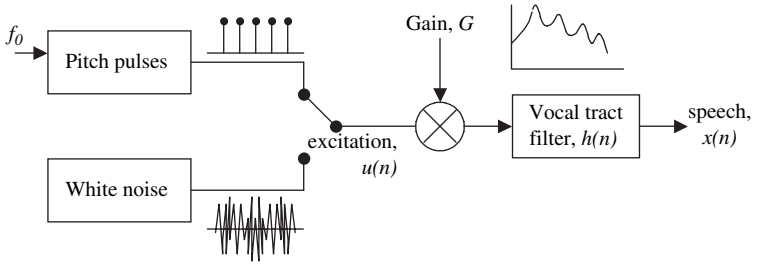


FIGURE 10 Source–filter model of speech production.

The task now remains of how to calculate the parameters of the model, namely the set of filter coefficients, $\boldsymbol{a} = [a_1, a_2, \ldots, a_P]$, the gain, G, the fundamental frequency, f_0, and the voicing decision.

Taking the Z-transform of Eq. (13) followed by re-arrangement leads to

$$H(z) = \frac{X(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}, \tag{14}$$

where $H(z)$ is the Z-transform of the vocal tract filter. The similarity between the source–filter model of speech and linear prediction leads to a straightforward method for estimating the model parameters. A linear predictor makes a prediction of a speech sample, $\tilde{x}(n)$, from a weighted summation of the p previous speech samples using a set of predictor coefficients α_k:

$$\tilde{x}(n) = \sum_{k=1}^{p} \alpha_k x(n-k). \tag{15}$$

The prediction error, $e(n)$, can then be defined as

$$e(n) = x(n) - \tilde{x}(n) = x(n) - \sum_{k=1}^{p} \alpha_k x(n-k). \tag{16}$$

Taking the Z-transform of Eq. (16) gives an expression for the filter, $A(z)$, which generates the prediction error samples:

$$A(z) = 1 - \sum_{k=1}^{p} \alpha_k z^{-k}. \tag{17}$$

Comparing Eq. (14) and (17) and setting $\alpha_k = a_k$ and $e(n) = Gu(n)$ allow the vocal tract filter, $H(z)$, to be expressed in terms of the prediction filter $A(z)$:

$$H(z) = \frac{G}{A(z)}. \tag{18}$$

Therefore, the vocal tract filter coefficients of the source–filter model are equal to the linear predictor coefficients of which computation is straightforward. The standard method for computing the predictor coefficients is to minimise the mean squared prediction error over a short-time analysis window. This has been very well documented and is generally solved using either the autocorrelation method or the covariance method. Considering the autocorrelation method of solution results in a set of linear equations:

$$\sum_{k=1}^{p} \alpha_k R(|i-k|) = R(i) \qquad 1 \le i \le p. \tag{19}$$

$R(k)$ is the *kth* autocorrelation coefficient of the frame of N speech samples under analysis which is defined as

$$R(k) = \sum_{n=0}^{N-1-k} x(n)\, x(n+k). \tag{20}$$

Equation (19) can be rewritten in matrix form, where the solution for the set of predictor coefficients, **a**, takes the form

$$\mathbf{a} = \mathbf{R}^{-1}\mathbf{p}. \tag{21}$$

The Toeplitz structure of matrix **R** in Eq. (21) allows several efficient methods of solving the equation to be applied such as Durbin's recursive solution. Knowledge of the vocal tract filter coefficients enables estimation of both the fundamental frequency and formants and forms the remainder of this Section.

4.1.1 LPC-Based Formant Analysis

Equation (13) has shown that the frequency response of the vocal tract can be modelled by an all-pole filter with a set of filter coefficients, $\boldsymbol{a} = [a_1, a_2, \ldots, a_p]$. The challenge for formant analysis is to use these filter coefficients to estimate formant parameters such as frequency, amplitude and bandwidth. This can be achieved by factoring the denominator of the vocal tract filter to give the poles. To illustrate this, Figure 11 shows a 32-ms short-time frame of speech taken from the phoneme /ih/ in the word "*think*" illustrated in Figure 3. Figure 12 shows a pole–zero diagram which indicates the poles of the vocal tract filter extracted from the frame of speech using 10th-order LPC analysis.

The pole–zero diagram shows the position of the 10 poles of the vocal tract filter where it can be seen that the poles exist as complex conjugate pairs. Each pair of poles contributes a resonance in the filter frequency response, the frequency of which

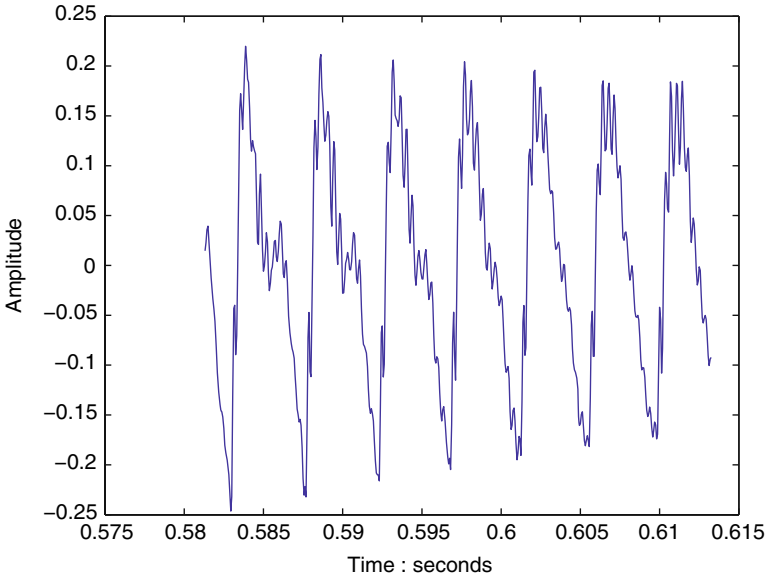


FIGURE 11 32-ms frame of speech from the phoneme /ih/.

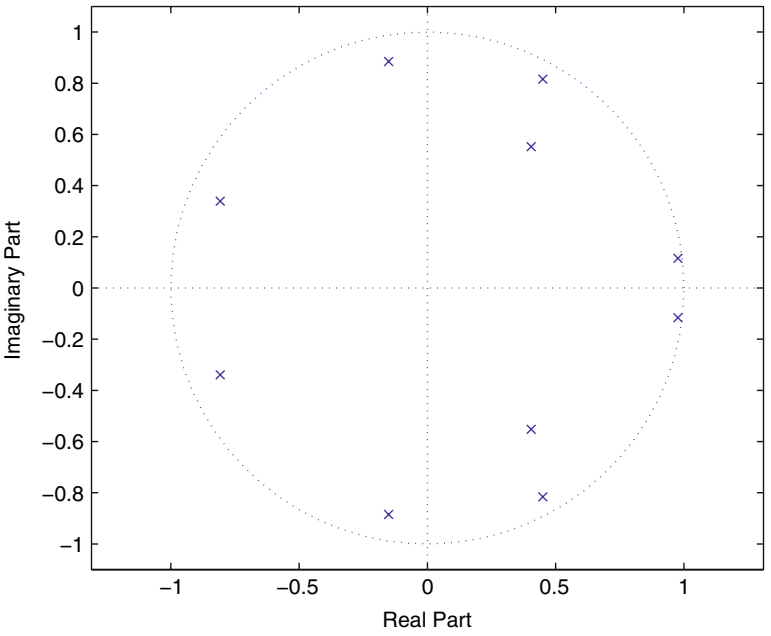


FIGURE 12 Pole–zero diagram showing the poles of the vocal tract filter.

is determined from the angle of the pole and the amplitude by its distance out from the origin. For the frame of speech shown in Figure 11, the pole–zero diagram shows that four poles pairs – 1, 3, 4 and 5 – are very close to the unit circle and will therefore contribute to significant spectral peaks in the vocal tract filter frequency response. The second pole pair has much less amplitude and will have significantly less effect on the frequency response of the vocal tract filter. The effect of these poles can be observed by examining the frequency response of the vocal tract filter. This is illustrated in Figure 13 and has been obtained by taking the Fourier transform of the vocal tract filter's impulse response.

The frequency response shows clearly the four formant resonances which have been modelled by the vocal tract filter. The frequency, F_i, of each formant is determined from the angle of the pole, θ_i, i.e. $F_i = \theta_i f_s/2\pi$ where f_s is the sampling frequency. The contribution of the second pole pair occurs just before the second formant and its effect is negligible due to its much smaller magnitude. It is interesting to compare the log magnitude spectrum of the vocal tract filter to the log magnitude spectrum, shown in Figure 4, which was obtained from the Fourier transform of the same frame of speech samples. The fine spectral detail shown in Figure 4 arises from the

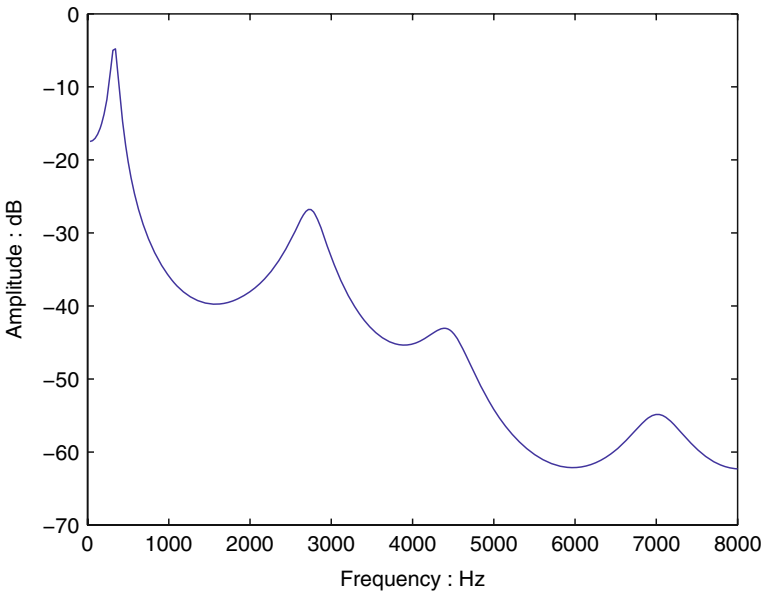


FIGURE 13 Frequency response of the vocal tract filter.

fundamental frequency and its harmonics which have been removed in the LPC analysis leading to the vocal tract filter. Therefore, the LPC analysis provides a useful method of obtaining the spectral envelope of the signal.

4.1.2 LPC-Based Fundamental Frequency Analysis

The fundamental period of the frame of speech shown in Figure 11 can be estimated by measuring the time between the peaks caused by the glottal pulses. Examination of the figure reveals the fundamental period to be approximately 4.5 ms which corresponds to a fundamental frequency of 222 Hz. For the speech shown in Figure 11, pitch estimation is relatively straightforward as the glottal pulses are clearly evident. However, for some voiced speech signals, the glottal pulses are not so clear, which makes their accurate identification and measurement prone to error leading to incorrect fundamental frequency estimates. This situation is made even worse when acoustic noise has contaminated the speech and has the effect of masking the glottal pulses.

A better signal from which to estimate the fundamental frequency is the excitation signal itself. The excitation signal has not been modulated by the vocal tract and has more clearly defined glottal pulses. An estimate of the excitation signal can be obtained by filtering the speech signal with a filter corresponding to the inverse of the vocal tract filter as is illustrated in Figure 14.

Examination of Eq. (18) shows that the inverse of the vocal tract filter, $H(z)$, is the predictor filter, $A(z)$. Passing the speech signal through this inverse filter removes the effect of the vocal tract filter frequency response. This leaves the resulting excitation signal with a flat frequency response, and hence the process is also termed spectral whitening. Figure 15 illustrates the excitation signal which has been generated by passing the frame of time domain samples in Figure 11 through the inverse vocal tract filter obtained from the filter shown in Figure 12.

Comparing the excitation signal to the original time domain signal shows that the glottal pulses are now much more evident.

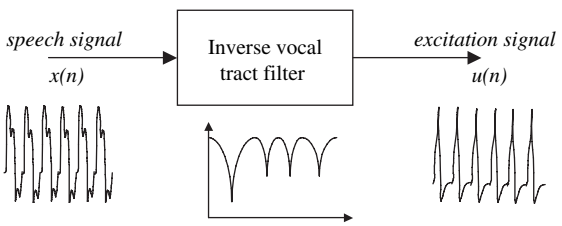


FIGURE 14 Inverse filtering the speech signal to recover the excitation signal.

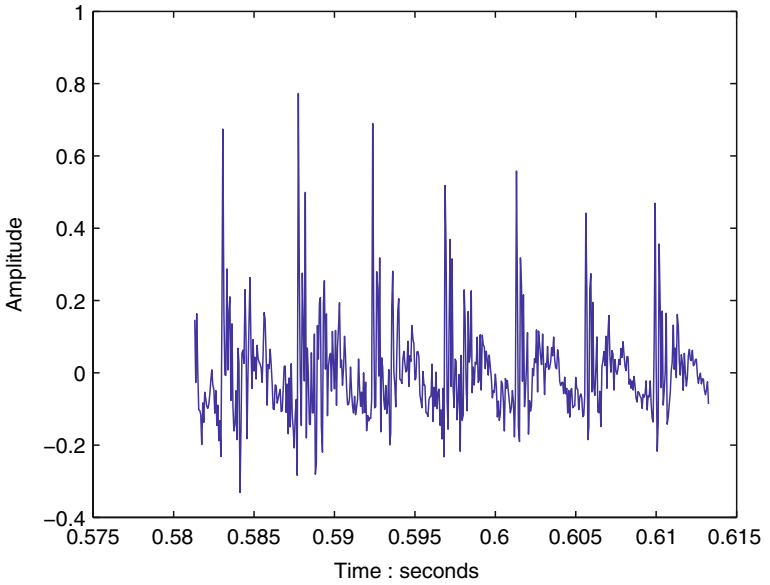


FIGURE 15 Excitation signal produced by inverse filtering.

Identification of the glottal pulses can be made more accurately than from the original speech signal, and measuring the precise time period between pulses is more accurate.

4.1.3 Autocorrelation-Based Fundamental Frequency Analysis

The fundamental frequency measurement in the previous section relies on correctly identifying pitch pulses and accurately measuring the time period between them. A more robust method of estimating the pitch period is to employ the autocorrelation analysis which was used in the calculation of the vocal tract filter coefficients. Autocorrelation analysis applied to a time domain signal is simple to implement and is used to identify patterns and structures in the signal. One such pattern that autocorrelation analysis can be used to identify is the periodic repetition of pulses in a speech waveform which results from the opening and closing of the glottis in voiced speech. The definition for autocorrelation coefficients, $R(k)$, is given in Eq. (20) where the first $p+1$ values are used in LPC analysis to determine the vocal tract filter coefficients. The vocal tract can be estimated from only a small number of autocorrelation coefficients due to its short-term nature. However, to estimate the fundamental frequency, a larger number of autocorrelation coefficients are necessary.

To illustrate the autocorrelation analysis, Figure 16 shows a 32-ms frame of speech taken from the phoneme /eh/ in the word "*air*" shown in Figure 2.

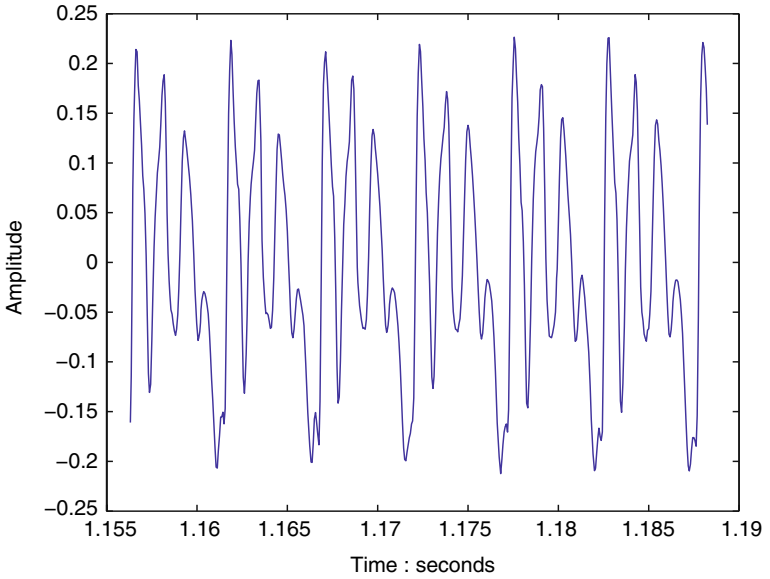


FIGURE 16 32-ms short-time window of speech taken from the phoneme /eh/.

Analysis of the time domain waveform to estimate the pitch period is not as straightforward as for the speech frame shown in Figure 11, and designing robust algorithms to identify glottal pulses would be more challenging. Figure 17 shows the first 512 autocorrelation coefficients, $R(0)$–$R(511)$, computed from the frame of time domain samples using Eq. (20), which corresponds to time lags up to 32 ms. As is normally the case, the autocorrelation coefficients have been normalised by $R(0)$.

The autocorrelation coefficients exhibit similar periodicity to the time domain signal shown in Figure 16 and have an identical period of repetition. Ignoring the peak at time lag zero, the largest peak occurs at a time of $t = 5.4$ ms which corresponds to the pitch period observed in Figure 16. However, the autocorrelation analysis makes identification of the peak corresponding to the pitch period much easier to locate. The smaller peaks found in the autocorrelation analysis occur from resonances in the vocal tract and from the shape of glottal pulses. Applying the autocorrelation-based fundamental frequency estimation to the excitation signal obtained from inverse filtering, discussed in the previous section, leads to even more robust fundamental frequency estimation.

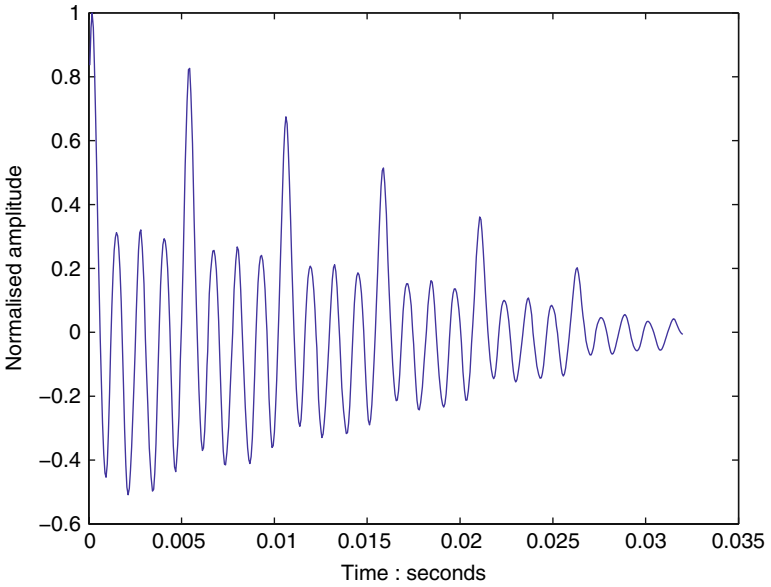


FIGURE 17 Autocorrelation analysis of a frame of speech from phoneme /eh/.

It should be emphasised, however, that robust fundamental frequency estimation across all speech sounds is a very difficult problem to solve and is still the subject of much research even today.

4.2 Frequency Analysis

The frequency domain representation of speech discussed in Section 3.2 shows that the magnitude spectrum comprises a spectral envelope component and a harmonic structure resulting from the glottal pulses. In the time domain, the speech signal is generated by the convolution of the excitation signal with the vocal tract filter. Therefore, in the frequency domain, the spectrum of the speech is given as a multiplication of the excitation signal spectrum and vocal tract frequency response. Figure 18 illustrates this process for a typical voiced vowel segment of speech.

Changing the fundamental frequency will alter the spacing of harmonics but will leave the spectral envelope, or phoneme being spoken, essentially unchanged. In a similar way, changing the spectral envelope will leave the fundamental frequency constant. From the speech spectrum, sufficient information exists to allow estimation of both the fundamental frequency and the formants.

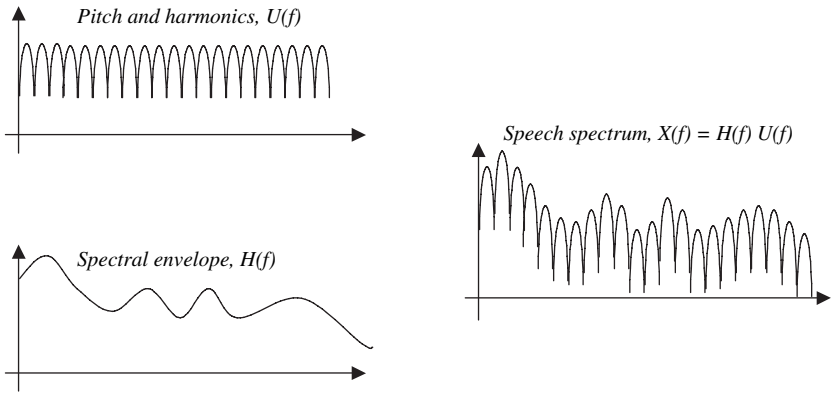


FIGURE 18 Fundamental frequency and spectral envelope components of the speech spectrum.

4.2.1 Frequency-Based Fundamental Frequency Analysis

Examination of the magnitude spectra of voiced speech in Figures 4 and 6 shows that the glottal pulses in the time domain signal appear as closely spaced harmonics in the spectral domain. The first harmonic occurs at the fundamental frequency, f_0, while subsequent harmonics occur at approximately integer multiples of the fundamental frequency, nf_0. The harmonics are in turn modulated by the frequency response of the vocal tract which gives the spectral envelope.

Identification of the fundamental frequency and its harmonics in the frequency domain therefore forms a method of estimating the fundamental frequency. The location of a particular harmonic may not always be reliable as its amplitude will have been modulated by the frequency response of the vocal tract. In fact it is common for the first few harmonics not to be present in situations where the speech has been received across a telephone channel which employs bandlimiting of the signal. To avoid these problems, and to increase robustness, most frequency domain methods of fundamental frequency estimation attempt to identify as many harmonics as possible.

One successful method of frequency domain fundamental frequency estimation is the comb function. This operates by first selecting a set of candidate fundamental frequencies, c_1 to c_N, to compare against the real fundamental frequency of the signal. These candidates may, for example, begin at 100 Hz and end at 400 Hz, incrementing in 10-Hz steps. For each of these candidate frequencies, an estimate of the magnitude spectrum of the glottal pulse train is synthesised by positioning a series of impulses at harmonics of

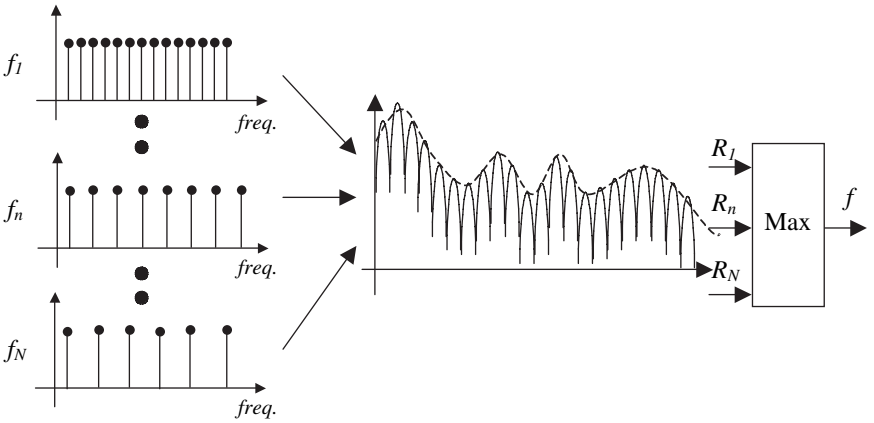


FIGURE 19 Illustration of the comb function for frequency-domain fundamental frequency estimation.

the candidate fundamental frequency. An example of a set of candidate fundamental frequencies and their synthesised magnitude spectra is shown in Figure 19. For each candidate fundamental frequency, the cross-correlation between the synthesised magnitude spectrum and the speech magnitude spectrum is measured. When the candidate fundamental frequency is close to the real fundamental frequency the harmonics will line up leading to high cross-correlation. The estimate of the fundamental frequency is selected as the candidate fundamental frequency which gave the highest cross-correlation with the spectrum of speech. The process is illustrated in Figure 19.

4.2.2 Frequency-Based Formant Analysis

The magnitude spectrum of a frame of speech not only shows the positions of formants but also contains excitation information in the form of harmonics. More reliable estimates of the formants can be obtained from the spectral envelope which does not show the finer spectral details. The spectral envelope can be obtained by applying the Fourier transform to a frame of speech which is sufficiently short so as not to contain a sequence of glottal pulses. However, according to discussions into time-frequency trade-offs in Section 3.2, this means that the frequency resolution will be poor. This can be avoided by zero padding the frame of speech to create a longer duration frame from which a smooth spectral envelope can be computed. Figure 20 shows a smoothed magnitude spectrum which has been computed from a frame of speech comprising just 20 time domain samples taken from the phoneme /ih/. This equates to a duration of 1.25 ms. Before applying the Fourier transform, the frame of speech has been

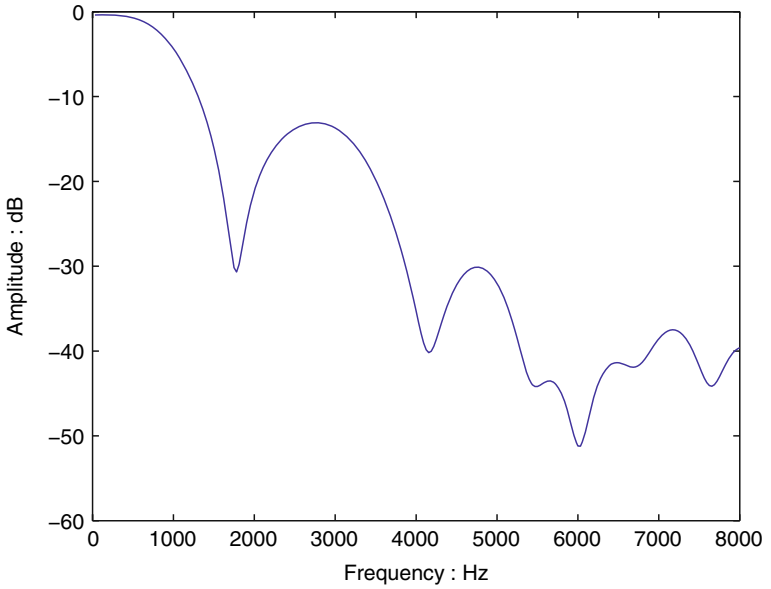


FIGURE 20 Spectral envelope obtained from 1.25-ms frame of speech and zero padding.

zero padded up to a duration of 32 ms through the augmentation of 492 zeros.

The smoothed spectral envelope makes identification of the formants straightforward through peak-picking algorithms. It should be noted that Figure 20 shows the formants very clearly. In practice, when analysing many different frames of speech the formant structure may not be so clear, leading to estimation errors.

4.3 Cepstral Analysis

Analysis of the magnitude spectrum of a speech signal showed that it comprises both spectral envelope information, relating to the vocal tract frequency response, and fine spectral detail from the fundamental frequency and its harmonics. Formant analysis can be applied effectively to the spectral envelope, while fundamental frequency estimation requires the fine spectral detail. Cepstral analysis is a form of homomorphic processing which enables these two components to be separated from each other. The cepstral transform is defined as the inverse discrete Fourier transform of the log power spectrum of a short-time frame of the signal as is illustrated in Figure 21.

The effect of the cepstral transform on a speech signal can be considered as follows. The input speech signal, $x(n)$, comprises

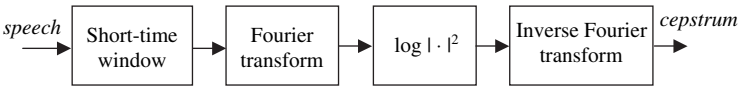


FIGURE 21 Cepstral transformation of an input speech signal.

the convolution of two components: the vocal tract impulse response, $h(n)$, and the excitation signal $u(n)$:

$$x(n) = h(n)^* u(n). \tag{22}$$

Transforming the signal $x(n)$ into the log power spectrum, $\log |X(f)|^2$, converts the convolutional operation into an addition operation as (ignoring the scaling by 2)

$$\log |X(f)| = \log |H(f)| + \log |U(f)|. \tag{23}$$

The log power spectrum of the vocal tract component, $\log |H(f)|$, is observed as the spectral envelope which varies slowly along the frequency axis. The component $\log |U(f)|$ gives the fine spectral detail from the fundamental frequency and its harmonics and has much faster variation along frequency than the vocal tract component. To complete the cepstral transform, the inverse discrete Fourier transform is taken of the log magnitude spectrum, $\log |X(f)|$. This serves to separate the vocal tract and excitation components of the signal. The slowly varying vocal tract information is located in the lower quefrency region of the cepstrum (quefrency is the name given to the domain following the cepstral transform and has units measured in seconds). The faster varying excitation component is located in the higher quefrency regions of the cepstrum. Applying a lowpass windowing function to the cepstrum retains the vocal tract information, while applying a highpass window retains the excitation information.

Figure 22 illustrates the log magnitude spectrum of the short-time frame of speech shown in Figure 11 which is taken from the phoneme /ih/.

In this figure, the fast moving spectral detail from the fundamental frequency and its harmonics can be seen together with the slowly varying spectral envelope resulting from the vocal tract. Figure 23 shows the cepstrum of this signal up to a quefrency of 16 ms.

Examining the cepstrum shows significant energy to be present at low quefrencies up to about 1 ms. This corresponds to the slowly time-varying vocal tract information. A peak in the cepstrum can be observed at a quefrency of about 4.5 ms. This

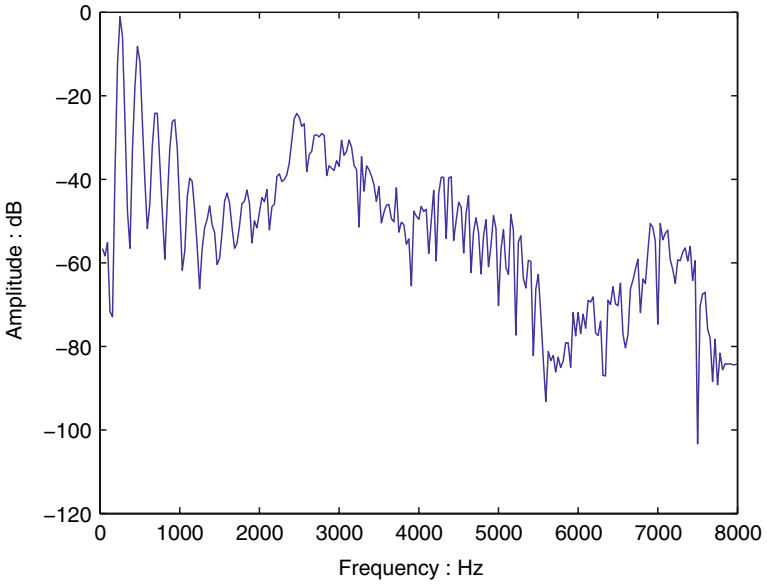


FIGURE 22 Log magnitude spectrum of a short-time frame of speech from phoneme /ih/.

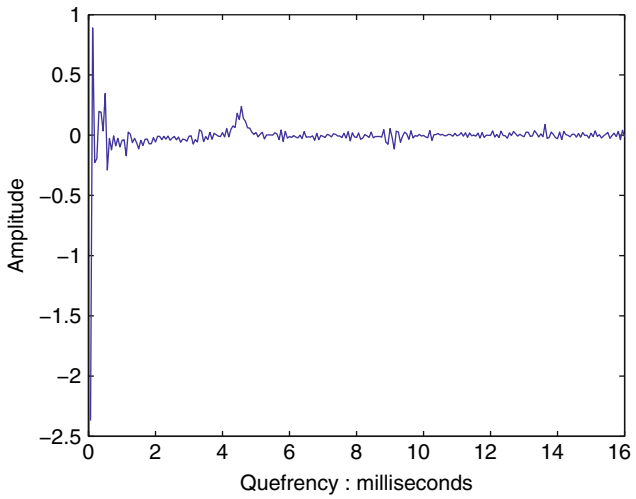


FIGURE 23 Cepstrum of frame of speech shown in Figure 11.

peak corresponds to the pitch period and shows that the cepstrum is an effective method for estimating the fundamental frequency of voiced frames of speech. Above this quefrency the information comprises the faster spectral variations arising from the excitation signal.

Using the fact that the cepstral transform is a homomorphic transform means that the vocal tract and excitation can be separated from one another by either lowpass or highpass filtering the cepstrum (filtering is actually termed liftering in the cepstral domain). For example, Figure 24a shows the resulting log magnitude spectrum when the discrete Fourier transform is applied to a lowpass-filtered cepstrum containing only the first 10 cepstral coefficients (up to 0.625 ms). Similarly Figure 24b shows the resulting log magnitude spectrum when the discrete Fourier transform is applied to a highpass-filtered cepstrum containing all cepstral coefficients except the first 10.

The figure shows the ability of the cepstral transform to separate the vocal tract and excitation components is very effective. From the lowpass-filtered cepstrum, the log magnitude spectrum in Figure 24a shows very clearly the spectral envelope of the vocal tract which makes identification of the formants obvious. The highpass-filtered cepstrum has the vocal tract modulation effects removed as shown by the comparatively flat log magnitude spectrum. This signal comprises mainly excitation signal components.

The position of the cepstral filter cut-off determines the information retained in the two log magnitude spectra. For example,

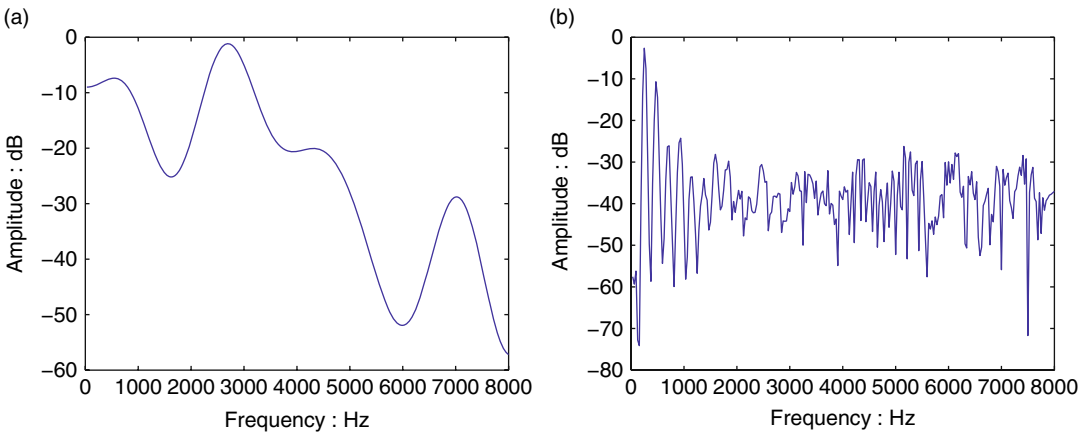


FIGURE 24 Log magnitude spectrum of (a) lowpass-filtered cepstrum and (b) highpass-filtered cepstrum at a quefrency cut-off of 0.625 ms.

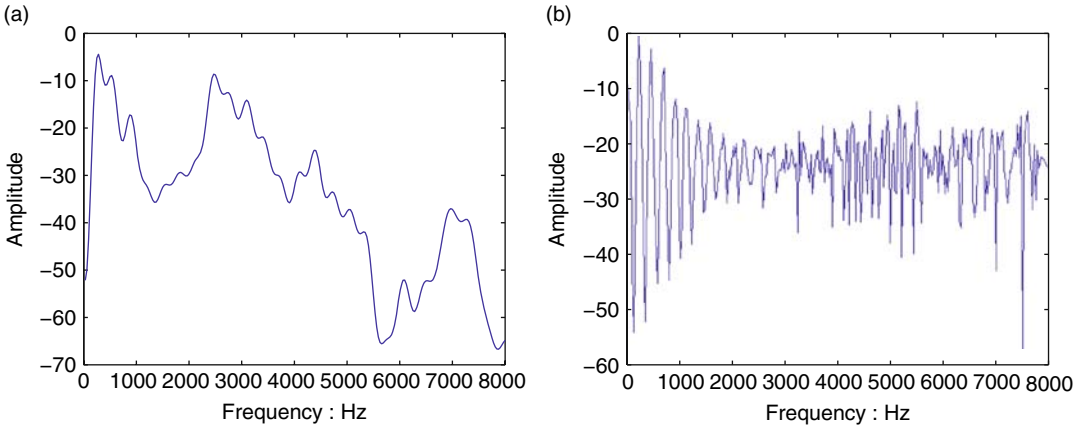


FIGURE 25 Log magnitude spectrum of (**a**) Lowpass-filtered cepstrum and (**b**) highpass-filtered cepstrum at a quefrency cut-off of 3.125 ms.

moving the cut-off from 0.625 to 3.125 ms increases the amount of faster moving information present in the lowpass-filtered cepstrum and removes some of the slower moving information from the highpass-filtered cepstrum. Figures 25a and b shows the resultant log magnitude spectra from lowpass- and highpass-filtered cepstra at a cut-off quefrency of 3.125 ms.

Comparing Figure 25a with Figure 24a shows that the increase in quefrency cut-off allows more faster moving spectral information to be retained in the lowpass-filter cepstrum.

Figures 24 and 25 have shown that the cepstrum is effective at separating vocal tract and excitation information which leads to effective fundamental frequency and formant estimation. To conclude the cepstral analysis of speech signals, Figure 26a shows the time domain representation of a frame of unvoiced speech taken from the phoneme /th/. The log magnitude spectrum of this signal has been shown in Figure 5 where the relatively flat frequency response and lack of harmonics indicate the unvoiced nature of the signal. Figure 26b shows the cepstrum of the signal.

The cepstrum of the unvoiced speech does not contain a peak as there are no harmonics present in the speech spectrum. There is also less energy in the lower quefrency coefficients in comparison to that observed for the voiced speech cepstrum shown in Figures 24 and 25. This is due to the reduction in spectral shape that the relatively flat frequency spectrum of the unvoiced speech has.

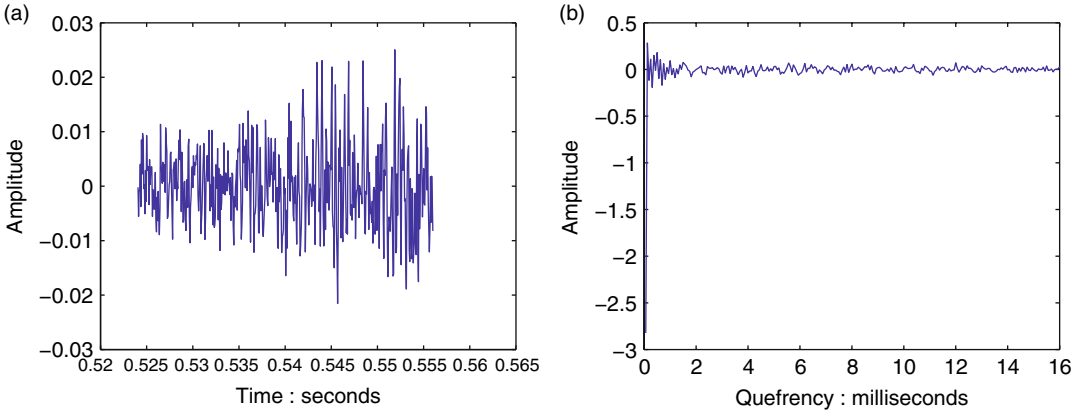


FIGURE 26 (**a**) Time domain representation of a frame of unvoiced speech; (**b**) cepstrum of the frame of speech.

4.4 Analysis for Speech Processing Applications

This section concludes with a brief discussion of feature analysis methods that have been designed specifically for use with particular speech processing applications such as recognition, coding and synthesis. The fundamental acoustic features of fundamental frequency and formants discussed earlier in this section are very important for analysis and understanding purposes but generally do not find direct use in speech processing applications. This is due in part to insufficient understanding of how such features can be successfully integrated into speech processing applications and also because of the difficulty in accurately estimating them across a range of different speech sounds, speakers and noise levels. Certainly, formant and fundamental frequency features have been used in formant-based speech synthesis applications, but for high-quality applications, these have been superseded by concatenative-based systems which are based on very large speech databases.

The majority of speech recognition systems use a variant of cepstral features termed mel-frequency cepstral coefficients (MFCCs). MFCC analysis aims to extract only the vocal tract component of the speech signal but does so using a frequency resolution based on the mel scale. This mimics the frequency response of the human ear by having fine spectral resolution at low frequencies and coarse resolution at high frequencies. Another form of analysis used in speech recognition is perceptual linear predictive (PLP) cepstra. This combines the vocal tract modelling used in LPC analysis with the non-linear frequency response of human hearing using a bark scaled warping. As with

the analysis methods described earlier in this section, both MFCC analysis and PLP analysis are based on short-time frames of speech with durations between 10 and 32 ms.

Some very low bitrate speech codecs, such as the vocoder, make use of the fundamental frequency directly but generally have poor quality. Instead, low-bitrate codecs generally use some form of vocal tract filter which is excited by an encoded representation of the excitation signal. Standard LPC analysis provides the vocal tract filter coefficients. In practice, these are converted to features such as line spectral pairs (LSP) which are more robust to the quantisation necessary to reduce the bitrate in low-bitrate speech codecs. The encoding of the excitation signal can take many different forms such as found in code-excited linear prediction (CELP) and mixed excitation linear prediction (MELP).

5 • Conclusion

This chapter has shown that a speech signal must be digitised before it can be displayed and analysed by computers. Various factors need to be considered in the sampling and quantisation processes such as the sampling frequency and number of quantisation levels. The display of speech signals has shown the time domain is useful for showing broad speech classes such as voiced and unvoiced. With sufficient detail in the time domain representation, measurement of the fundamental period is also possible. The frequency domain allows much more structure of the speech signal to be seen. Importantly, this shows how a speech signal can be decomposed into its vocal tract and excitation components. Combining the time and frequency domains in the form of a spectrogram allows the time-varying nature of the fundamental frequency and formants to be seen very clearly.

LPC analysis, frequency analysis and cepstral analysis have been shown as effective methods for recovering vocal tract and excitation information from the speech signal. Demonstrations have shown how these analysis methods can be used to extract the fundamental acoustic correlates of fundamental frequency and formants from a speech signal. It has also been mentioned that these methods are subject to errors, particularly in the presence of acoustic noise, and are still the subject of much research even today.

Finally, some speech analysis methods have been discussed that are applicable to speech processing applications such as recognition, coding and synthesis. The parameters that these systems require are often based on standard analysis methods but with subtle differences.

28

Estimation of Speech Intelligibility and Quality

Stephen Voran

Institute for Telecommunication Sciences, Boulder, CO, USA

1 • Introduction

Speech communication requires a talker and a listener. Acoustical and in some cases electrical representations of the speech are carried from the talker to the listener by some system. This system might consist of the air in a room, or it might involve electro-acoustic transducers and sound reinforcement or telecommunications equipment. Interfering noises (including reverberation of speech) may be present and these may impinge upon and affect the talker, the system, and the listener. A schematic representation of this basic unidirectional speech communication scenario is given in Figure 1.

In general, a talker can be a human or a machine (speech synthesizer) and a listener can also be a human or a machine (automatic speech recognizer). The quality or intelligibility of speech synthesis and the effectiveness of automatic speech recognition are topics unto themselves and are outside the scope of this chapter. Quality and intelligibility issues in bidirectional communications are also outside the scope of this chapter, but these issues are addressed from a telecommunications perspective in [1–3] for example.

Given the situation in Figure 1, there are numerous interesting questions that one might ask. These include: How much of the speech communication is correctly heard and how much effort

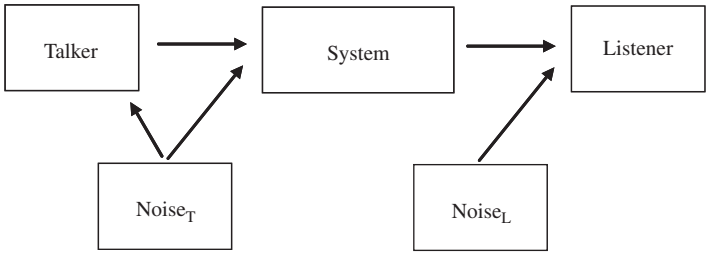


FIGURE 1 A unidirectional speech communication scenario.

does this take? What is the listener's opinion of the quality of the speech that he/she hears? How significant are the impairments to the speech that the listener hears? How do the answers to these questions depend on the talker, the listener, the system, and the noises?

These questions can be considered from multiple perspectives. For example, the question of speech quality could be pursued with respect to acoustic integrity, talker pronunciation and enunciation, or content. These are very different perspectives indeed. Throughout this chapter, our perspective is driven by consideration of how the system and the noises have altered the acoustic integrity of the talker's speech, and what those alterations are likely to mean to a listener with hearing characteristics that fall within normal ranges. Within this context, we use the term "intelligibility" to describe how understandable a speech message is and the term "quality" to describe how pleasing a speech message sounds. Indeed we are only scratching the surface of the myriad of issues in speech communication. One could argue, though, that notions of acoustic integrity are fundamental prerequisites for the more general discussion of the integrity of speech communications.

Each of the questions above ultimately concerns a human listener. Human factors experiments, called subjective tests or subjective experiments, commonly can provide the most direct answers. Further information on subjective speech intelligibility tests and subjective speech quality tests can be found in [4–7]. These tests generally require specialized equipment (e.g., sound-isolated rooms, high-quality transducers, and playback equipment) and they demand significant time associated with careful test design, conduct, and analysis. These costs have led numerous investigators to develop signal processing algorithms that seek to provide approximate answers to some of these questions. In effect, such algorithms need to emulate the

behaviors of the human subjects in the experiments that they seek to replace.

In general, the design of a signal processing algorithm that emulates a human behavior can become as complex as one will allow. Some simplification can be gained by assuming that both the talker and the listener are adults with speech production and hearing faculties that fall within the "normal" range. It is generally prudent to focus the algorithm on a single distinct question such as intelligibility, quality, listening effort, or impairment level. One might seek to develop an algorithm that uses speech signals or some specially constructed non-speech test signals. The algorithm may process the system input and output, or it may be driven by the system output alone. Perhaps most challenging is the range of systems that an algorithm seeks to address. These could include rooms with reflections and/or noises, possibly combined with electronic sound reinforcement, telecommunications, or speech playback systems. The electronic portions of the system may be approximately linear and time-invariant (e.g., amplifiers) or they may be highly non-linear, dynamic, signal-adaptive devices (e.g., non-waveform digital speech coders), or they may be mixtures of the two.

Given the many dimensions of the problem, it is understandable that investigators have proposed numerous specific solutions for specific cases. Due to the richness of the field, a comprehensive survey is not possible here. Instead a selection of examples from the range of proposed solutions is provided. These examples were selected to demonstrate the range of signal processing techniques that have proven applicable to this problem.

Section 2 introduces some techniques for estimating speech intelligibility and Section 3 addresses estimation of perceived speech quality in telecommunications. Section 4 offers further observations and references to additional work in the field.

2 • Estimators of Speech Intelligibility

Background noises and reverberation can reduce the intelligibility of the speech that arrives at the listener. Several algorithms have been developed to estimate the relationship between these factors and speech intelligibility. These estimation algorithms are principally applicable to room acoustics problems where any sound reinforcement system that may be present applies only linear processing. They also find application in some basic telecommunications problems where analog or

waveform-preserving digital elements are used and no non-linear processing comes into play.

2.1 Articulation Index

About 1921, Harvey Fletcher began a research program at Bell Labs to determine, among other things, how the different portions of the speech band (approximately 200–6,000 Hz) contribute to successful speech communications [8]. His experiments with filtered speech and human subjects led to mathematical models that largely answered this question, at least in the relevant context of the day. Fletcher's work in the 1920s stands today as the fundamental basis for understanding speech intelligibility. His development, called the Articulation Index (AI), was largely held internal to Bell Labs until the publication of [9] and [10] in 1947 and 1953, respectively. (Note that in the case of the AI, articulation refers to effects in speech transmission, not speech production.)

Kryter did additional work to formulate [11], validate [12], and standardize [13] two specific, relatively accessible AI calculations. Both are based on speech and noise levels in multiple frequency bands. One calculation uses 20 bands, each of equal importance to speech intelligibility as determined by Fletcher. The other uses 5 or 6 octave bands or 15 third-octave bands, perhaps making it more practical given the filters available at that time.

The 20 equal importance bands cover the range from 200 to 6,100 Hz. They are specified in [11,14] and are depicted graphically on both the Hertz frequency scale and psychoacoustic Bark (or critical band) frequency scale in Figure 2. This figure also shows the commonly used Hertz to Bark relationship:

$$f_{\text{Bark}} = 6 \cdot \operatorname{arcsinh}\left(\frac{f_{\text{Hz}}}{600}\right). \tag{1}$$

Below 400 Hz this relationship is approximately linear and above 1 Hz it is approximately logarithmic. Critical band frequency scales, such as the Bark scale, have been empirically derived to account for a basic property of human hearing, and additional details are provided in "Bark Spectral Distortion". On the Hertz scale the 20 bands generally increase in width as the center frequency increases, and these widths span about a decade: from 100 to 1,050 Hz. On the Bark scale the bands are wider at the low- and high-frequency ends of the scale and narrower in the center of the scale. The widths span just a factor of 2: from 0.6 Bark near the center of the scale to about 1.2 Bark at either end of the scale.

Thus the widths of the AI bands are significantly more uniform on the Bark scale than on the Hertz scale. Since the Bark scale

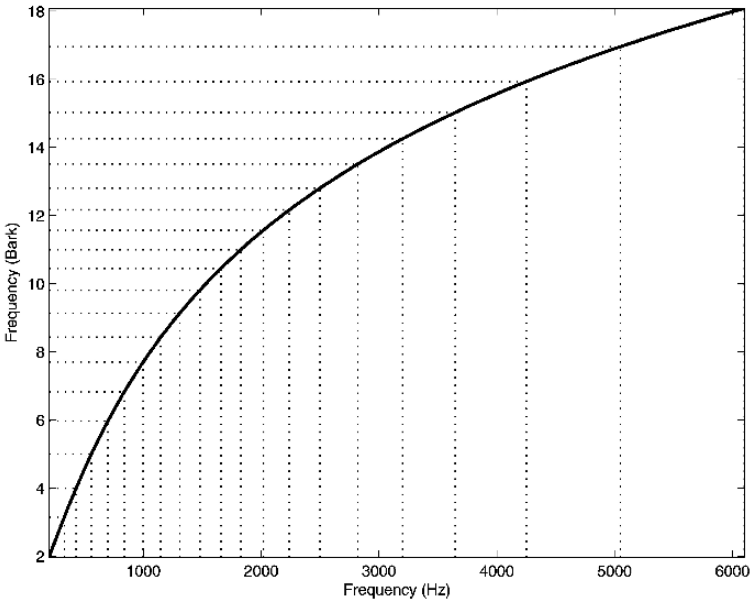


FIGURE 2 The 20 bands used in the calculation of AI, shown on both the Hertz and Bark frequency scales, along with the Hertz to Bark relationship of equation (1).

models an attribute of human hearing, this suggests that human hearing accounts for a portion of the band importances that drive the AI principle. The fact that the AI bands are not completely uniform suggests that human hearing does not account for all of the band importances, and thus the way that information is carried in the human speech spectrum must account for the remainder of the observed effect.

Fletcher determined that overall speech intelligibility can be modeled as the sum of individual contributions from multiple bands. Further, for the cases of band limiting and background noise at least, the contribution of each band is sufficiently described by the long-term signal-to-noise ratio (SNR) in each band. Thus the AI calculations given in [11] include the following steps:

1. Measure or calculate the average power spectrum of the speech and the background noise at the location of interest.
2. Modify the noise spectrum to account for the non-linear growth of masking in the human auditory system for bands where the noise is more than 80 dB above the hearing threshold. Further modify the noise spectrum to account

for the spread of masking across frequencies in the human auditory system.

3. In each band, use the signal power S_j and the noise power N_j to calculate the SNR. Limit that SNR to the range 0–30 dB, then normalize it to the range 0–1 to find the per band contribution to articulation:

$$\begin{aligned} \mathrm{SNR}_j &= 10 \cdot \log_{10}\left(\frac{S_j}{N_j}\right), \\ a_j &= \frac{\min\left(\max\left(\mathrm{SNR}_j, 0\right), 30\right)}{30}. \end{aligned} \tag{2}$$

4. Now combine the contributions of each of the bands to attain AI:

$$\mathrm{AI} = \sum_{j=1}^{N} \alpha_j \cdot a_j, \text{ where } 1 = \sum_{j=1}^{N} \alpha_j. \tag{3}$$

For the $N = 20$ band calculation, each band contributes equally to speech intelligibility and thus $\alpha_j = 0.05$, $j = 1$–20. When the one-octave or third-octave bands are used, the weights α_j reflect the differing contribution of each band. These band definitions and weights can be found in [11]. In either case, since $0 \le a_j \le 1$ and $\sum_{j=1}^{N} \alpha_j = 1$, it follows that $0 \le \mathrm{AI} \le 1$.

Graphs demonstrating the monotonic relationship between AI and various types of subjective speech intelligibility test results can be found in [11,12,14].

2.2 Speech Transmission Index

In many situations the undesired signals that interfere with speech communication include reverberations of the desired signal. AI was not intended to address this situation. Further, there are practical issues making application of AI in this environment difficult or impossible. Since the desired speech signal produces at least a portion of the undesired interfering signal, separate measurements of the two signals' power spectra may not always be physically possible. The accurate calculation of such power spectra is generally of prohibitive complexity as well.

Steeneken and Houtgast built on the AI principle to develop algorithms that could properly account for effects of reverberations, and possibly some non-linear speech distortions as well. Amplitude-modulated noise signals and modulation transfer functions are key to this work. In early work a noise signal (with power spectrum matched to the long-term speech power

spectrum) was modulated by a 3-Hz square wave and this evolved to square waves of multiple frequencies. This approach progressed to a more sophisticated algorithm where the system under test (SUT) is excited with test signals that employ sinusoidal amplitude modulation of the power of seven one-octave bands of noise. Fourteen different modulating frequencies are used. This algorithm is called the Speech Transmission Index (STI) and first appeared in [15]. A recent revision of STI is called STIr. The development of STIr is described in [16], the verification is given in [17], and a complete STIr measurement procedure is standardized in [18]. Various implementations are possible, and one is summarized here:

1. Generate a noise signal and filter it into seven different one-octave-wide bands, centered at 125, 250, 500, 1,000, 2,000, 4,000, and 8,000 Hz. Adjust the relative levels of the seven signals to match the given long-term average speech spectra for female or male talkers. In the case of female talkers, the band centered at 125 Hz is not used. The use of these seven bands allows for the independent characterization and weighting of the SUT behavior at each octave of importance in speech transmission. The time domain samples of these seven signals are $n_j(i)$ where i is the sample time index and $j = 1$–7 is the band index.
2. Amplitude-modulate each of these seven signals at 14 different modulation frequencies f_k, $k = 1$–14, to create 98 modulated noise signals with sample rate f_s:

$$n_j^k(i) = n_j(i)\sqrt{1+\cos(2\pi f_k i/f_s)}. \tag{4}$$

   The modulation frequencies f_k range from 0.625 to 12.5 Hz and are spaced in third-octave steps. Equation (4) results in a sinusoidal modulation of signal power. Note that these signals are fully modulated (modulation index or modulation depth = 1) in the sense that the modulating function goes to zero on each modulation cycle. The presence of this periodic null in the test signal aids in the evaluation of the noise and reverberation power that may be present in the signal after passing through the SUT. The range of modulation frequencies is consistent with those found in actual human speech signals.
3. Sequentially pass each of these 98 signals $n_j^k(i)$ through the SUT to create 98 new signals $\tilde{n}_j^k(i)$.
4. Filter each of the 98 received signals $\tilde{n}_j^k(i)$ through the jth octave-wide bandpass filter. Square and low-pass filter the result using a cut-off frequency that removes the noise carrier

associated with band j, but that preserves the modulation frequency f_k. Model the resulting extracted envelope as

$$\mathrm{LPF}\left(\tilde{n}_j^k(i)^2\right) \approx g \cdot \left(1 + M_j^k \cos\left(2\pi f_k \cdot (i - i_0)/f_s\right)\right), \tag{5}$$

where M_j^k is the observed modulation index for band j and modulation frequency f_k, g is a gain factor, and i_0 is a time delay associated with the SUT. The observed modulation indices M_j^k can be reduced from their original value of 1 by the presence of noise and reverberations induced by the SUT.

5. Adjust each modulation index M_j^k, to account for the spread of masking from the octave band below as given in [18]. The new modulation indices are designated $\tilde{M}_j^k$. This allows the subsequent calculations to account for how a listener would perceive the test signals that have passed through the SUT.
6. Convert each modulation index to an effective SNR, limit each SNR to the range −15 to +15 dB, then normalize to the range 0–1 to find the transmission index for each octave band and modulating frequency:

$$\begin{aligned} \mathrm{SNR}_j^k &= 10 \cdot \log_{10}\left(\frac{\tilde{M}_j^k}{1 - \tilde{M}_j^k}\right) \\ \mathrm{TI}_j^k &= \frac{15 + \min\left(\max\left(\mathrm{SNR}_j^k, -15\right), 15\right)}{30}. \end{aligned} \tag{6}$$

7. Average across the 14 modulating frequencies in each octave band to find the modulation transfer index for each band:

$$\mathrm{MTI}_j = \frac{1}{14}\sum_{k=1}^{14} \mathrm{TI}_j^k. \tag{7}$$

8. Then calculate STIr through a weighted summation of the modulation transfer indices for the seven octave bands along with the corresponding inter-band redundancy correction:

$$\mathrm{STI}_r = \sum_{j=1}^{7} \alpha_j \cdot \mathrm{MTI}_j - \sum_{j=1}^{6} \beta_j \sqrt{\mathrm{MTI}_j \cdot \mathrm{MTI}_{j+1}}, \tag{8}$$

where

$$\sum_{j=1}^{7} \alpha_j - \sum_{j=1}^{6} \beta_j = 1. \tag{9}$$

These weights result from the fitting of experimental data and their values can be found in [16] and [18]. The weights α_j reflect the relative contribution of each octave band to speech transmission. The inter-band redundancy correction (consisting of a weighted sum of the geometric means of adjacent band modulation transfer indices) is the primary difference between STIr and STI. This correction compensates for the apparently redundant contributions made by adjacent octave bands in some cases [16], and their inclusion does improve the agreement between STIr and subjective speech intelligibility tests [17]. Graphs demonstrating the monotonic relationship between STI or STIr and various types of subjective speech intelligibility test results can be found in [15,17,18].

Because the STIr test signals are based on noise, reliable STIr results can require significant measurement intervals for each of the 98 test signals. A typical choice for the measurement interval is 10 s, and this then requires over 16 min to perform the full STIr measurement. For increased measurement efficiency, three abbreviated STI measurement procedures have been developed for specific applications and the associated set of speech impairments: STITEL for telecommunications systems [18], STIPA for room acoustics and public address systems [18], and RASTI [18,19] for room acoustics alone. Each of these uses only a specific subset of the 98 test signals, and in some cases simultaneous rather than sequential excitation of the SUT leads to further efficiency. For example, one implementation of the RASTI algorithm applies only 12 s of test signal to the SUT.

In addition to the modulated noise signal-based measurements, the original STI publication [15] proposed that STI could be made more applicable to non-linear distortions of speech through the use of a more speech-like noise signal. Such a signal would have a pseudo-random modulation (modeled after real speech) in all octave bands other that the current band under test. The noise signal in the band under test would have a sinusoidal modulation. The desired response to non-linear distortions has remained elusive, and several authors have proposed techniques based on actual speech signals. Reference [20] summarizes some of these proposals and offers some new proposals as well.

2.3 Speech Intelligibility Index

The Speech Intelligibility Index (SII) [21] can be described as an updated and expanded version of AI. Expressions for the spread of masking, the notion of standard speech spectrum levels, and the relative importance of the individual bands have been updated. SII places the AI concept into a more general framework that allows one to account for hearing thresholds that deviate

from the average and measurements made in a free field or at the eardrum. As a direct descendant of the AI family, SII is driven at its core by per frequency band SNRs. Two main techniques for arriving at the applicable signal and noise level values are described in [21]. One involves direct measurements or calculations, similar to the procedure for AI. The second method extracts the signal and noise levels from modulation transfer functions as STI does, but with different modulation frequencies and signal bands. The first method will generally be faster, but of more limited applicability. The second method is better suited to reverberant environments, but will generally require longer measurement intervals.

3 • Estimators of Speech Quality in Telecommunications

Modern telecommunications and speech storage systems are enabled by digital coding of speech waveforms. Speech coders that operate at higher bit rates can approximately preserve speech waveforms. Lower rate speech coders generally invoke constrained models for speech and extract and transmit model parameters that allow for the approximate reproduction of the speech sound, but with limited or no attempt to preserve the speech waveform. Because of this, these speech coders can exhibit highly non-linear, time-varying, signal-dependent behaviors. The distortions they introduce cannot be modeled as additive noise or reverberations. These speech coders, combined with other elements in telecommunications networks, severely limit the applicability of the speech intelligibility estimators presented above.

Many telecommunications systems incur virtually no impairment to the ability to conduct speech communications. While the intelligibility of the speech delivered by such systems may not be an issue, such systems can still deliver a wide range of perceived speech qualities. For example, one experiment compared the perceived quality of speech constrained to the nominal passband used in narrowband telephony (300–3,400 Hz) with speech constrained to the nominal passband used in wideband telephony (50–7,000 Hz) [22]. A clear difference in perceived speech quality was found and subjects clearly preferred the wideband passband. On the other hand, the AI correctly estimates that speech intelligibility is increased by a miniscule amount when wideband telephony replaces narrowband telephony. Thus speech quality, not speech intelligibility, has been the main recent concern in telecommunications.

It is worth noting, however, that speech intelligibility can be a very real issue in some wireless telephony, Voice over Internet Protocol (VoIP), and land mobile radio systems, yet existing intelligibility estimators generally are not applicable to these situations. One recent effort in this area can be found in [23].

Most estimators of perceived speech quality in telecommunications are focused on the effects of speech coders, transmission channels, and possibly some other network elements. The issues associated with noise at the talker and listener locations are presently under study.

3.1 General Principles

Figure 3 is a high-level block diagram for many speech-based, signal processing-based algorithms that estimate perceived telecommunications speech quality.

These algorithms make use of both the speech signal that goes into the SUT and the speech signal output from the SUT. We denote the time domain samples of these two signals $x(i)$ and

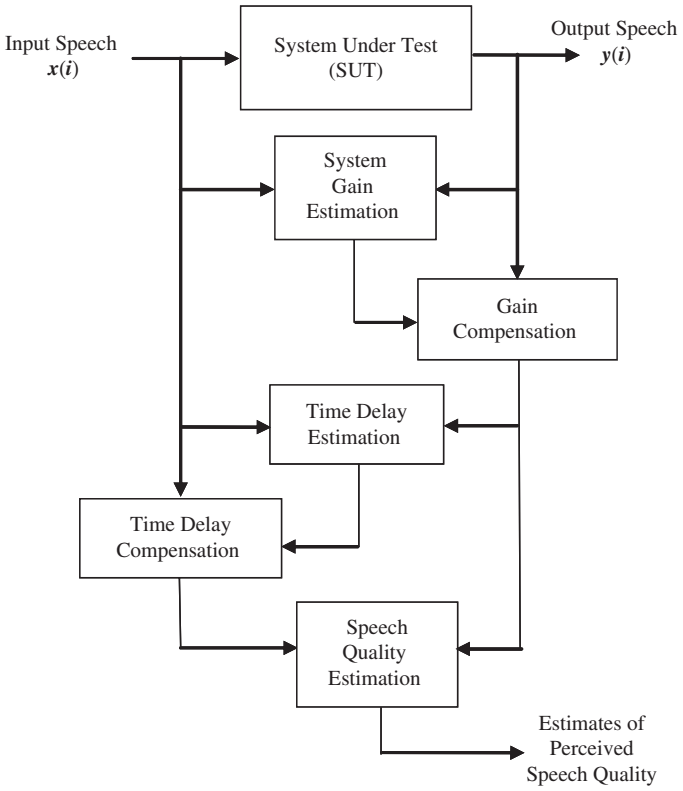


FIGURE 3 High-level steps in speech quality estimation.

$y(i)$, respectively, and they are also known as the reference and test signals. These signals are typically acquired in the analog or digital electrical domains, though it is also possible to acquire one or both of them in the acoustic domain. It is important that speech quality estimation be performed using real or artificial [24,25] speech signals rather than test signals because speech coders and other network elements can respond very differently to different classes of signals.

There is also a signal-free approach to telecommunications speech quality estimation. No signals are required, but complete knowledge of the system elements and operating parameters is required. Previously determined relationships between system elements and speech quality are combined in a mathematical model that produces an estimate of the speech quality that will be delivered by the SUT. This can be very efficient, but is limited to system elements and interactions that have already been completely characterized. Examples of this approach can be found in [1,2,26,27]. We do not consider this approach further in this chapter.

A constant gain in the telecommunications SUT is not typically considered to be a relevant part of the speech quality issue. This gain is often estimated and removed as indicated in Figure 3. The gain estimation can be performed using power ratios or by selecting a gain that minimizes the noise or maximizes the SNR associated with the SUT [28]. It is subsequently assumed that the speech is delivered to the listener at some preferred, standardized listening level.

3.2 Estimation of Constant Delay

As shown in Figure 3, any constant delay in the SUT must be estimated and removed so that the speech quality estimation block receives properly synchronized input signals. If unidirectional speech quality is of interest, this fixed delay is of no consequence. If it is large enough it can affect the quality of a bidirectional conversation. If this is of interest, then the effect of the delay must be combined with the estimated speech quality to fully understand the quality one might experience in the bidirectional case.

Estimation of SUT delay can be a significant problem. If the SUT approximately preserves waveforms, then conventional waveform cross-correlation delay estimation techniques can suffice. However, when waveforms are not preserved, these techniques can fail. In addition, when waveforms are not preserved one must look beyond waveforms to even define what delay means. Since speech envelopes and power spectra are approximately preserved, these are useful representations. The

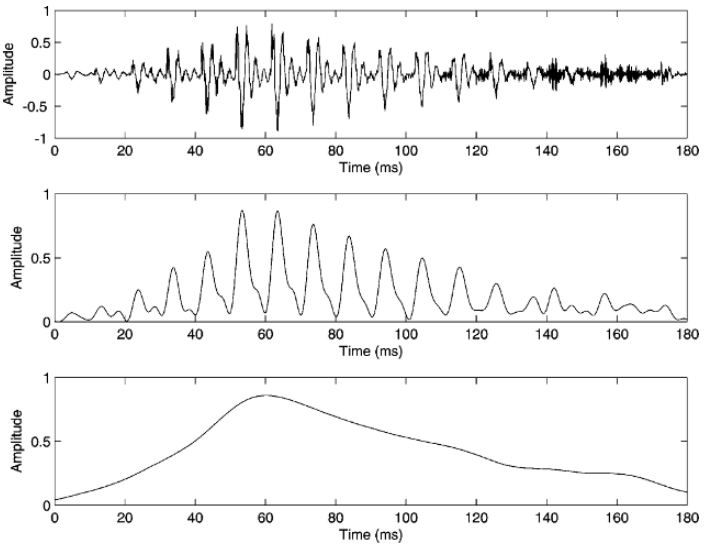


FIGURE 4 Speech waveform, 250-Hz bandwidth speech envelope and 63-Hz bandwidth speech envelope (*top to bottom*), for a male speaking the word "is".

delay estimation technique summarized in [29] and standardized in [30] uses two stages. The first stage performs a wide, low-resolution search to generate an initial delay estimate. This is accomplished by cross-correlating subsampled speech envelopes. In this case, speech envelopes are calculated by rectifying and low-pass filtering speech waveform samples, and examples of such envelopes are given in Figure 4. Subsampling consistent with the envelope bandwidth can then be performed. The peak in the cross-correlation function between the envelopes of the test and reference signals provides the initial estimate of the SUT delay, which is then removed.

The second stage performs cross-correlations between estimated power spectral densities (PSDs). A single reference signal PSD is correlated against a family of test signal PSDs that are calculated from a range of adjacent time windows in the test signal. If these correlations consistently indicate an additional time shift between the test and reference signals, then this second stage result is combined with the first stage result. If the second stage results do not pass the consistency test, then the first stage delay estimate is used alone, and the associated uncertainty in delay is deemed to be an inherent property of the SUT.

3.3 Estimation of Time-Varying Delay

Some SUTs can introduce time-varying delays into speech signals. When speech is transmitted as a series of packets over a shared transmission channel (as in VoIP), transmission errors and congestion can cause the received stream of packets to be non-uniformly spaced in time, and some packets may be lost completely. Packet buffers at the receiving end can mitigate these effects to some extent, but in practice the received speech may still contain temporal discontinuities where the end-to-end delay increases or decreases abruptly. Depending on the duration and location of these temporal discontinuities, they can affect perceived speech quality [31].

Time-varying delays must be estimated and removed so that the quality estimation block receives properly synchronized input signals. Ideally, the perceptual effects of the accompanying temporal discontinuities would be accounted for in the quality estimation algorithm as well. In other words, when time-varying delays are present, the delay estimation and compensation problem and the speech quality estimation problem are no longer truly separable.

3.3.1 A Top-Down Time Delay Estimation Algorithm

One algorithm for estimating time-varying delays is given in [32] and is included as part of a standardized algorithm for estimating perceived speech quality [33]. This algorithm includes a coarse alignment step that performs cross-correlation on logarithmically compressed speech envelopes. The fine alignment step uses 64-ms frames of speech samples. These frames are weighted by the Hann window and are overlapped by 75%. Time domain cross-correlation is applied to these frames and the index and value of the correlation maximum are used to calculate an estimated fine time delay and a confidence in that estimate, respectively.

The coarse alignment step is first applied to the entire available recordings of the reference and test signals to find a single best global time delay. To refine this result each individual contiguous segment of active speech (each utterance) is then considered separately. The coarse and fine steps are applied to each utterance and a measure of confidence is calculated. Multiple two-way splits of each utterance are then explored, and the split that gives the highest confidence is compared with the confidence before splitting. The situation (split or no split) with the higher confidence is considered to be the correct situation. If a split is identified, that utterance is marked as consisting of two different segments, each with its own separate delay value. If an utterance is identified to have two different delays, then the trial splitting is again applied to each segment identified. This recursive trial splitting continues until no segment can be successfully split into

a pair of segments. This can be described as a top-down approach, starting with the full recordings and successively working down to smaller and smaller segments of constant delay.

3.3.2 A Bottom-Up Time Delay Estimation Algorithm

A different solution for estimating time-varying delays in speech is described in [34] and is standardized in [35]. This algorithm starts with cross-correlation of subsampled speech envelopes with 63-Hz bandwidth to find the single best global time delay. This delay value is then further refined through additional steps. The next step uses speech envelopes with 250-Hz bandwidth. Unlike the 63-Hz bandwidth, the 250-Hz bandwidth preserves much of the pitch information in the speech signal (see Figure 4), thus allowing for its potential use in delay estimation. A 150-ms window of the test signal speech envelope is cross-correlated with 150-ms windows of the reference signal envelope to find the best delay estimate for each such window in the test signal. This window is advanced in 40-ms steps, and the sequence of results is then further processed and median filtered to generate an initial delay history. For each segment of constant delay identified in this initial delay history, additional cross-correlation is performed using rectified speech. The use of rectified speech makes the algorithm completely robust to phase reversal in the SUT, and it exploits the full bandwidth of the speech signal. If the cross-correlation results pass a set of reliability tests, the associated time shift is used to further refine the initial delay history. This algorithm can be described as a bottom-up algorithm, since its core stages move small processing windows over longer speech signals. When a real-time implementation is desired, data-flow issues can make a bottom-up solution more practical than a top-down solution.

In the following summaries of algorithms that estimate telecommunications speech quality, we use $x(i)$ and $y(i)$ to denote the reference and test signals presented to the speech quality estimation algorithm, and we assume that gain compensation and delay compensation have already been performed using one of the methods outlined above.

3.4 Frequency-Domain Estimators

Digital speech coders may exploit the fact that the preservation of speech waveforms is sufficient, but not necessary for the preservation of speech quality. Estimating speech quality directly from the sampled time domain signals $x(i)$ and $y(i)$, or the error signal $y(i) - x(i)$, can be effective for devices that preserve speech waveforms, but it cannot be effective across the broader class of speech coders, network elements, and transmission channel conditions. Since speech power spectra are

approximately preserved, estimators that work in the frequency domain provide a good starting place. Two popular frequency-domain estimators are summarized below. Additional details and relationships between them are given in [36].

3.4.1 Log Spectral Error

Perhaps the most basic frequency-domain estimators of perceived speech quality are log spectral error (LSE) based estimators. Here the input and output speech sample streams are typically grouped into frames of $2N$ samples, windowed with a smooth window, and transformed via an FFT to create N complex frequency-domain samples X_j for each frame of the reference signal, and likewise (Y_j) for the test signal. Frame lengths from 8 to 32 ms are typical. Frames shorter than 8 ms typically do not contain enough samples to yield the desired frequency resolution, and frames longer than 32 ms are very unlikely to contain a stationary segment of speech. Frames are often allowed to overlap by 50% so that any portion of a speech signal that is interrupted by a frame boundary will also be processed uninterrupted in the center of a frame as well. When combined with the choice of an appropriate window, overlap can also ensure that every speech sample receives the same total weight in the calculation. For each such frame, the spectral representations can be used to calculate a log spectral error as

$$\mathrm{LSE} = \left(\frac{1}{N} \sum_{j=1}^{N} \left| \alpha_j \cdot 20 \cdot \log_{10} \left(\frac{|Y_j|}{|X_j|} \right) \right|^p \right)^{\frac{1}{p}}. \tag{10}$$

The weights α_j allow for spectral weighting in the LSE calculation. With $p = 1$, the resulting LSE is the mean absolute spectral error in dB. With $p = 2$, it becomes the RMS spectral error in dB. LSE provides a simple scalar indication of how well an SUT preserves speech power spectra on a per-frame basis. It has been commonly used as a figure of merit in the quantization of linear prediction coefficient, reflection coefficient, or line-spectral frequency representations of smoothed speech spectra. It can be averaged or otherwise processed across multiple frames to generate an estimate of perceived speech quality for a speech signal. With the addition of simple normalization steps, LSE can form the basis of a relatively effective estimator of perceived speech quality [37].

3.4.2 Cepstral Distance

An alternate way of comparing log speech power spectra is through cepstral analysis. The discrete time Fourier transform of the logarithm of a sampled speech PSD yields a set of cepstral

coefficients. In a typical cepstral coefficient calculation the speech signal is processed in frames of 8–32 ms in length, often with 50% overlap of frames. Each speech frame is multiplied by a smooth window and an FFT is then applied, followed by the magnitude and then the logarithm operators, yielding a sampled estimated log PSD for the given frame of speech. An FFT is then applied to this result to generate a set of cepstral coefficients. Alternately, cepstral coefficients can be calculated from linear prediction coefficients or reflection coefficients. Cepstral coefficients can then be used in the calculation of cepstral distance (CD):

$$\mathrm{CD} = \left(\frac{1}{N} \sum_{i=1}^{N} \left(\alpha_i \cdot \left| c_x(i) - c_y(i) \right| \right)^p \right)^{\frac{1}{p}}, \tag{11}$$

where $c_x(i)$ and $c_y(i)$ represent the cepstral coefficients associated with a given frame of the reference and test speech signals. The cepstral-domain quantity, CD, can then be averaged or otherwise processed across multiple frames to generate an estimate of perceived speech quality for an entire test speech signal. While (11) gives a generalized form of CD, most published works have used $N = 8$–12, the norm $p = 2$, and weights $\alpha_i = 1$.

Optimization of N, p, and α_i in (11) may yield improved speech quality estimates for specific applications. The cepstral coefficients provide a spectral representation for the log PSD. Thus the lower-order cepstral coefficients describe the amount of smoother, wider variation in reference and test speech frame PSDs with respect to frequency (e.g., spectral tilt) while the higher-order cepstral coefficients indicate the amount of more abrupt, localized variation in the PSDs with respect to frequency (e.g., speech formant structure). The observation that changes in these two types of variation (e.g., fixed SUT frequency response vs. SUT formant distortion) are perceived very differently by listeners motivates a search for an optimal set of weights. More generally though, these and other observations motivate a search for perception-based estimators of perceived speech quality.

3.5 Perception-Based Estimators

When listeners report their perceptions of speech quality, they are necessarily making judgments based on acoustic stimuli that have been processed by their human auditory systems. When seeking to develop a signal processing algorithm that estimates perceived speech quality, it is thus advantageous to account for both the human auditory system and human judgment in that algorithm. Equipped with complete understanding of these processes and accurate mathematical models for them, one might expect to

develop a signal processing algorithm that can indicate perceived speech quality in a way that is very close to the indications coming from human listeners. This notion has led to the development of perception-based algorithms that estimate perceived speech quality. These perception-based algorithms often work in the frequency domain, but it is not the Hertz scale frequency domain. Rather, it is a psychoacoustic frequency domain, developed to be consistent with basic properties of human hearing.

Figure 5 shows the common approach to perception-based estimation of perceived speech quality. It can be helpful to think of the perceptual transformation as an emulation of human hearing and the distance measure as an emulation of human judgment. Note, however, that hearing and judgment are complex interdependent processes. Treating them as separable sequential processes is pragmatic simplification. Figure 5 shows a distance measure (judgment) that is fed simultaneously by two perceptually transformed speech signals. This differs from the situation with humans in subjective tests. They are never exposed to the two speech signals simultaneously, and in the most common subjective speech quality tests, called absolute category rating (ACR) tests, they never hear the reference signal at all, yet they reliably judge the quality of the test signal. It may be that human subjects facilitate this task by forming an internal, idealized version of the test signal, against which they can judge the test signal. In the less common degradation category rating (DCR) subjective tests, human subjects hear the reference and test signals sequentially, a situation that would be closer to but different from that described in Figure 5.

It is frequently the case that a perception-based algorithm will contain both a priori models for hearing and judgment and empirically determined components. The a priori models are typically based on hearing and judgment results found in the literature and the empirically determined components are often selected to maximize the agreement between the algorithm

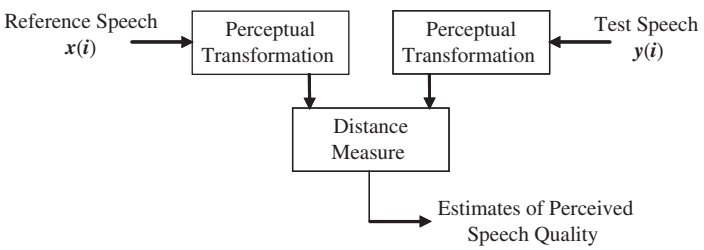


FIGURE 5 Perception-based speech quality estimation.

output and some body of speech quality truth data. Each of the estimators summarized below was designed to estimate speech quality in the 300 to 3,400-Hz telephone bandwidth, but straightforward extensions to wider bandwidths are possible.

3.5.1 Bark Spectral Distortion

Bark spectral distortion (BSD) [38] is an algorithm for estimating perceived speech quality that incorporates three basic properties of human auditory perception. The first of these is simultaneous masking, whereby a spectral component at a given frequency can raise the threshold of hearing at neighboring frequencies, making it harder to detect other spectral components that may be present at that same instant in time. This property of hearing can be conveniently conceptualized and modeled through the notion of critical bands and critical band filters [39]. The shape of a critical band filter serves as a simple model for simultaneous masking at a particular frequency. A set of appropriately spaced critical band filters forms a critical band filterbank, and the filterbank output can exhibit a first-order approximation to the simultaneous masking effects on complex signals. The filterbank output can be described as a critical band power spectrum, or a Bark power spectrum, since Bark is one psychoacoustic unit of measure for frequency. The filters are typically spaced at uniform intervals on the Bark scale and the filterbank output thus provides a power spectrum that is uniformly sampled on the Bark scale.

Figure 6 displays the magnitude response of four example critical band filters on the Hertz scale. These filters are centered at 310, 870, 1,830, and 3,630 Hz, corresponding to 3, 7, 11, and 15 Bark, respectively (see equation (1) and Figure 2). The −3 dB filter bandwidths are 110, 180, 320, and 610 Hz, corresponding to a 1.0-Bark width in all four cases. When viewed on the Bark scale these four filters are uniformly spaced (4.0 Bark), have uniform −3 dB bandwidth (1.0 Bark) and identical shape. Each filter shows asymptotic slopes of +10 dB/Bark below center and −25 dB/Bark above center and these slopes reflect the nominal first-order simultaneous masking properties of human hearing.

A second basic property in human hearing is its frequency-dependent sensitivity. For example, a sinusoidal signal at 3 kHz is perceived as significantly louder than a sinusoidal signal at 100 Hz with the same power. In the BSD algorithm, these first two properties are modeled by a weighted critical band filterbank. Each filter in the bank is weighted to reflect the relative sensitivity of human hearing in the corresponding frequency range. In the BSD algorithm the filters are spaced at 1-Bark intervals and the first filter is centered at 1.5 Bark. This arrangement is a design choice and results in a total of 15 filters across the 1.5–15.5 Bark

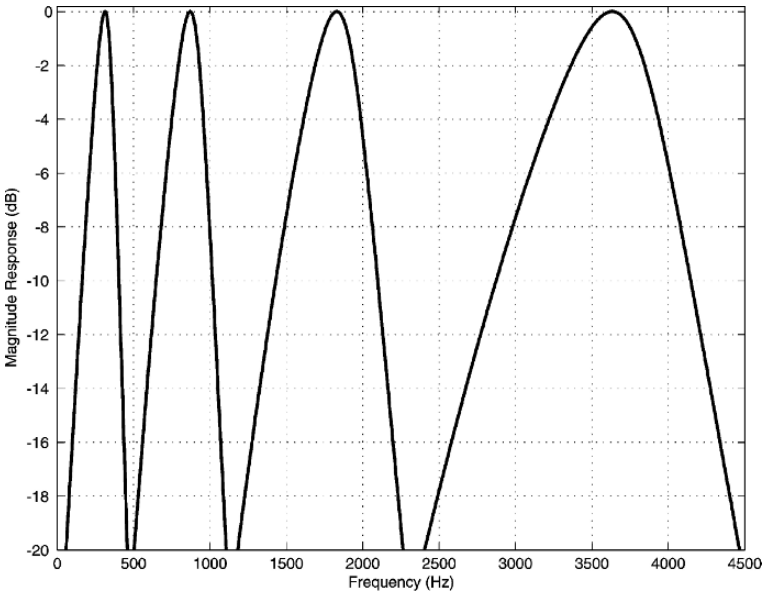


FIGURE 6 Magnitude response of four example critical band filters.

(150–4,000 Hz) band. An example of such a filterbank is given in Figure 7.

In the BSD algorithm, the filtering described above is implemented in the frequency domain. The speech signals are processed in 10-ms frames with 50% overlap. Each frame is then Hamming windowed and an FFT and magnitude function is applied. This results in 40 frequency-domain magnitude samples between 0 and 4,000 Hz for each 10-ms frame of speech that is processed. For a given frame, these 40 samples can be organized into the column vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ for the reference and test signals, respectively. A 15-by-40 matrix, F, can then be used to properly accumulate these samples as the 15 filters (see Figure 7 for example) would, resulting in the length 15 Bark spectrum column vectors $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$:

$$\tilde{\boldsymbol{x}} = F\boldsymbol{x}, \tilde{\boldsymbol{y}} = F\boldsymbol{y}. \tag{12}$$

Each row of F contains the power response of one filter.

The third hearing property treated in the BSD algorithm is the perception of loudness. Even at a single fixed frequency, the perception of loudness is not linearly related to signal power. The transformation applied to each sample of $\tilde{x}$ and $\tilde{y}$ to account for this is

$$Lx_j = \alpha 2^{20\log_{10}(\tilde{x}_j)(\tilde{x}_j)/10} = \alpha\tilde{x}_j^{0.60},\ Ly_j = \alpha\tilde{y}_j^{0.60}, \tag{13}$$

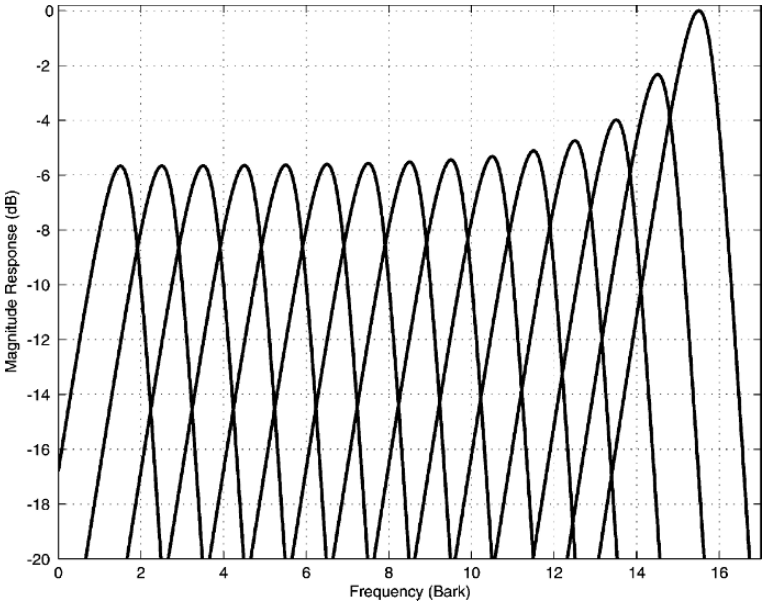


FIGURE 7 Magnitude response of example filterbank with 15 critical band filters centered at 1.5, 2.5,..., 15.5 Bark and weighted to reflect the approximate frequency sensitivity of hearing.

where α is a constant. Reference [38] attributes this transformation to [40]. The samples Lx_j and Ly_j form the Bark loudness spectra of the corresponding frames of the reference and test signals. They reflect the ear's non-uniform integration across frequencies, its non-uniform sensitivity to different frequencies, and the non-linear growth of perceived loudness as a function of signal power. The samples Lx_j and Ly_j are the final output of the perceptual transformations used in BSD.

The distance measure used in the BSD algorithm is squared Euclidean distance:

$$\mathrm{BSD} = \sum_{j=1}^{15} \left(Lx_j - Ly_j\right)^2 . \tag{14}$$

The BSD value given by (14) corresponds to one pair of 10-ms frames of reference and test speech. These are then averaged across all frames of interest and the resulting value is normalized by the time average of $\sum_{j=1}^{15} Lx_j^2$, thus creating a final BSD value which might also be described as perceptual noise-loudness-to-signal-loudness ratio.

Other researchers have noted that some smaller loudness differences might be masked by the speech signal and therefore would not be noticed by listeners. Thus they extended the BSD to model this effect [41]. The relationships between BSD values and subjective speech quality test results can be found in [38].

3.5.2 Perceptual Speech Quality Measure

The perceptual speech quality measure (PSQM) algorithm introduced in [42] and standardized in [43] further extends the notion of perception-based speech quality estimation. PSQM transforms signals into a perceptual domain where they are then compared by a distance measure. The PSQM does not use critical band filtering because experiments indicated that modeling simultaneous masking was not advantageous in this application [42]. It does use a critical band or Bark scale spectral representation with 56 samples across the 0–4,000 Hz band, resulting in sample spacing of about 0.31 Bark.

The two speech signals are processed in 32-ms frames with 50% overlap. Each frame is then Hanning windowed and an FFT and magnitude squared function is applied. To transform this Hertz domain representation (with 128 samples of PSD) to a Bark scale representation (with 56 samples), appropriately sized contiguous groups of Hertz domain samples are averaged and then normalized by the local sample density correction factor $\frac{\Delta \mathrm{Hz}(f)}{\Delta \mathrm{Bark}}$. This can be implemented in the same manner as the matrix filterbank implementation used in the BSD, but in this case the filters are rectangular and non-overlapping.

The PSQM algorithm then performs a power normalization step. Next it models the power frequency response of a typical telephone earpiece and the background noise spectrum that is sometimes used in subjective speech quality tests.

The frequency-dependent sensitivity of hearing and the perception of loudness are modeled together as

$$Lx_j = \alpha \left(\frac{T_j}{0.5}\right)^{\gamma} \cdot \left[\left(0.5 + 0.5\frac{Px_j}{T_j}\right)^{\gamma} - 1\right],$$
$$Ly_j = \alpha \left(\frac{T_j}{0.5}\right)^{\gamma} \cdot \left[\left(0.5 + 0.5\frac{Py_j}{T_j}\right)^{\gamma} - 1\right], \tag{15}$$

where T_j is the threshold of hearing for spectral sample j, Px_j and Py_j are the reference and test signal powers at spectral sample j, and α is a constant. Reference [43] attributes this loudness model to [44]. The loudness growth exponent γ was empirically optimized to maximize agreement between the PSQM output and subjective speech quality test results, resulting in the value

$\gamma = 0.001$. The net effect of this very small exponent is that at each frequency j, the calculated loudness value Lx_j or Ly_j is nearly logarithmically related to the corresponding power Px_j or Py_j. These logarithmic loudness relationships are individually shifted at each frequency, consistent with the hearing threshold at that frequency [45]. A loudness normalization step completes the operations of the perceptual transformation. For each pair of speech frames processed (reference and test), the output of the perceptual transformation is two sets of 56 compressed loudness density values Lx_j and Ly_j.

The distance measure used in PSQM is based on absolute differences of compressed loudness densities $|Ly_j - Lx_j|$. However, it was recognized that added spectral components (which cause $Ly_j > Lx_j$) are likely to be more annoying than missing spectral components (which cause $Ly_j < Lx_j$). Thus the information lost when taking the absolute difference is then restored with additional processing that effectively weights added components about twice as heavily as subtractive components [45].

This is followed by averaging over frequency and time, with a provision to assign different weights to results from frames that contain active speech and results from frames that contain only background noise. The final result is a non-negative value called noise disturbance. Higher values of noise disturbance indicate lower estimated speech quality. Relationships between noise disturbance values and subjective speech quality test results are given in [42].

3.5.3 Measuring Normalizing Block Algorithms

When seeking to emulate human responses, it is logical to use both perceptual transformations and a distance measure, as shown in Figure 5. Further, it is logical that these two components should be equally well developed so that both hearing and judgment receive sufficient consideration. There are an abundance of experimental results available to guide modeling of human hearing and thus guide the development of perceptual transformations. There is significantly less information available to guide the modeling of auditory judgment and thus the development of distance measures. This has understandably resulted in fairly sophisticated perceptual transformations that feed relatively simplistic distance measures. Intuitively, this situation seems less than optimal. Indeed, one study has concluded that distance measures are at least as important as perceptual transformations, and the benefits of sophisticated perceptual transformations tend to be lost when they are followed with simplistic distance measures [46]. Simplistic distance measures may show

unrealistic invariance to sign (added vs. missing signal components) or unrealistic invariance to the temporal or spectral scale of impairments (constant vs. rapidly fluctuating gain, low-pass vs. comb filtering).

In the measuring normalizing block (MNB) speech quality estimation algorithms, attention is given to bringing the contributions of the perceptual transformation and the distance measure into balance and eliminating the undesirable invariances cited above. These algorithms are described in [29], verified in [47], and standardized in [48] and [49]. In the MNB algorithms the perceptual transformation models only two basic properties of human hearing. It works in the frequency domain using bands organized according to the critical band concept and uses a simple logarithmic non-linearity to create an approximation of the relationship between signal power and perceived loudness.

The distance measure in the MNB algorithm consists of a hierarchy of processing blocks that can measure and then normalize spectral differences out of the test signal. This can be done on different time and frequency scales, and proceeds from larger scales to smaller scales, following the hypothesis that this would be most likely to emulate listeners' patterns of reaction and adaptation to spectral differences. In addition, the measurements produced by these blocks separate positive and negative spectral differences at the various scales.

To measure spectral differences over a given time-frequency scale, the perceived loudness time-frequency representations $X(i,j)$ and $Y(i,j)$ for the reference and test signals are each individually summed over that time-frequency region. The difference between the two results is the measurement for that scale, and this difference is subtracted from the appropriate region of $Y(i,j)$ to form $\tilde{Y}(i,j)$. Because of this normalization, $X(i,j)$ and $\tilde{Y}(i,j)$ now "agree" at that scale. That is, repetition of this measurement would now yield a measurement of zero. However, further measurements at smaller scales (sub-regions) will continue to reveal additional, smaller scale differences between $X(i,j)$ and $\tilde{Y}(i,j)$. As the measurements and normalizations to the test signal proceed, the normalized test signal becomes more and more similar to the reference signal, first from a large-scale view and then from progressively smaller scale views.

An example of this process is given in Figure 8. For visual simplicity the time scale used is just the time associated with a single 16-ms frame. Thus the figure shows Bark loudness spectra for a single pair of reference (dotted) and test (solid) speech frames using arbitrary loudness units. The frequency scales used

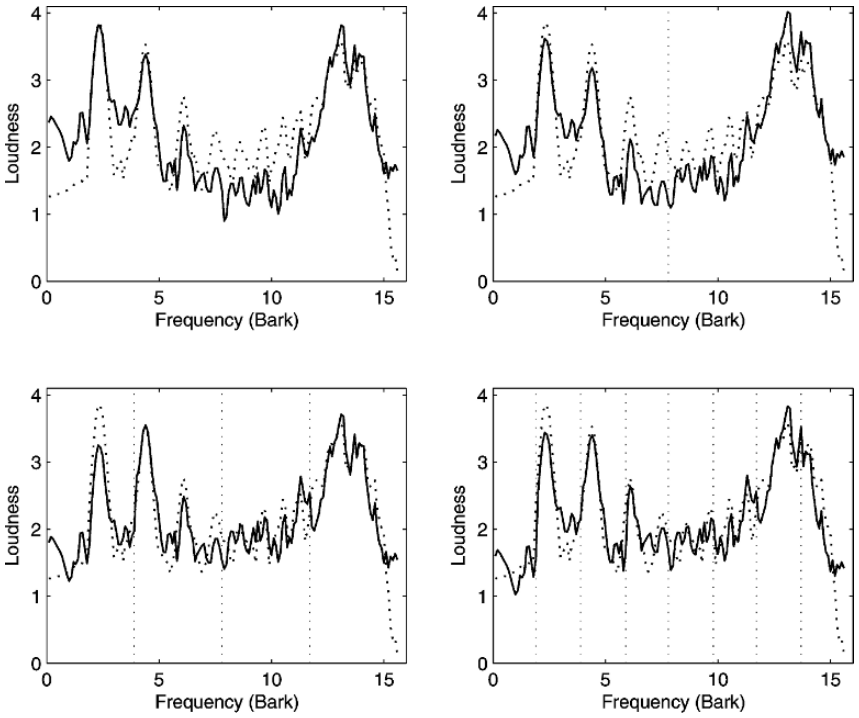


FIGURE 8 Example reference (*dotted*) and test (*solid*) signal loudness spectra for a sequence of MNB steps at decreasing frequency scales (*top left to lower right*). Loudness units are arbitrary.

range from 15.5 to 1.94 Bark and these decrease as one moves through the four panels of the figure. The upper left panel shows the two spectra after the test spectrum has been normalized across the full 15.5-Bark band. Because of this normalization, the areas under the test and reference spectra are the same. The upper right panel shows the two spectra after two additional measurements and normalizations. In this case the measurements and normalizations were conducted at the half-band (7.75 Bark) frequency scale and the consequence is that for both the lower and upper halves of the frequency range, the areas under the reference and test spectra agree. The lower left panel shows the two spectra after four additional measurements and normalizations and these were conducted at the quarter-band (3.88 Bark) frequency scale. Thus for each quarter of frequency range, the area under the reference and test spectra agree. Finally, the lower right panel depicts the result after eight additional measurements and normalizations at the eighth-band (1.94 Bark) scale. As one moves through the four panels in the figure, the

test spectrum is evolving toward the reference spectrum and is showing agreement at increasingly finer frequency scales.

In an MNB algorithm a record of the measurements that drove this evolution is preserved and these measurements are organized by sign, time scale, and frequency scale. For example, a measurement associated with a larger frequency scale may track SUT spectral tilt, while a measurement associated with a smaller frequency scale may indicate SUT distortion of speech formants, independent of that spectral tilt.

The MNB distance measure process can also be viewed as a dynamic decomposition of the test signal where the decomposition proceeds in a space that is defined partly by human hearing and judgment (via the MNB structure) and partly by the reference signal. In this view the measurements become the coefficients of the decomposition.

A linear combination of the measurements then forms a single auditory distance (AD) value that describes the perceived distance between the input reference and test signals. Since it is a distance, $0 \leq \mathrm{AD} < \infty$ and $\mathrm{AD} = 0$ would indicate perfect transmission and thus perfect quality. Subjective speech quality tests generally produce results on a finite scale with increasing values indicating higher speech quality. Thus a logistic function

$$\mathrm{L(AD)} = \frac{1}{1 + \exp(\alpha \cdot \mathrm{AD} + \beta)} \tag{16}$$

with $0 < \alpha$ is used to map AD to the interval, $0 < \mathrm{L(AD)} < 1$. An auditory distance of zero will give $\mathrm{L\,(AD)} = \left(1 + \mathrm{e}^{\beta}\right)^{-1} \approx 1$, while large values of auditory distance will give $\mathrm{L\,(AD)} \to 0$. The linear combination weights used to determine *AD* and the logistic function parameters α and β were selected to maximize correlation between L(AD) and a large set of subjective test results. Thus while the MNB structure can be viewed as an a priori model for human judgment, the details of how that structure operates (the weights) are found empirically. The relationships between L(AD) values and subjective speech quality test results can be found in [47].

While the MNB concept is broad, the specific MNB structures given in [29] operate at a single time scale, effectively measuring and then removing a single time-averaged frequency response of the SUT, and that measurement then figures into the calculation of AD. The MNB structures measure and remove spectral errors at 1–3 different frequency scales. *AD* is then calculated as a linear combination of 11 or 12 resulting linearly independent measurements.

3.5.4 Perceptual Evaluation of Speech Quality

As mentioned in Sect. 3.3, some SUTs can introduce time-varying delays into speech signals. The perceptual evaluation of speech quality (PESQ) algorithm was developed to respond to this situation. It is described in [32] and [50] and standardized in [33] which supercedes [43] and [49]. The algorithm combines the top-down variable delay estimation algorithm described in Sect. 3.2 above and in [32] with an enhanced version of the PSQM algorithm. The PESQ algorithm approximates the delay estimation and quality estimation processes as separable processes. The estimated delay history is removed from the speech signals before quality estimation begins. This delay history does not generally drive the quality estimation algorithm. (The exception is when a speech frame includes a delay increase that exceeds 16 ms. Results from such frames are not included in the final quality estimates produced by PESQ.) If the quality estimation process detects unusually high levels of distortion in specific segments of the speech signals, those segments are further analyzed to see if the delay estimate there can be improved.

The quality estimation stage of the PESQ algorithm is an enhanced version of the PSQM algorithm. One significant difference from the PSQM algorithm follows the creation of the Bark scale power spectral densities for the reference and test signals. The time-averaged frequency response of the SUT is measured and then partially normalized out of the reference signal. This reduces the algorithm's response to fixed filtering effects, consistent with the fact that listeners can often adapt to these effects. In fact, a listener may not even notice or object to mild fixed filtering, especially if no reference signal is available for comparison (as is the case in ACR subjective tests).

Another significant difference from the PSQM algorithm follows immediately. Here any time-varying gain is measured and partially normalized out of the test signal. This is accomplished by calculating the SUT gain for each pair of 32-ms reference and test speech frames, smoothing this gain history (time constant of 16 ms, low-pass corner frequency of 63 Hz), and then using the result to normalize the test signal. Roughly speaking, low-frequency (relative to 63 Hz) gain variations are removed (and presumably are not objectionable to listeners), while high-frequency (relative to 63 Hz) gain variations are retained and their effects are subsequently measured.

The frequency-dependent sensitivity of hearing and the perception of loudness are modeled together as given in (15) for the PSQM algorithm. In the case of the PESQ algorithm, however,

the parameter γ takes the value 0.23 for spectral samples at frequencies above 4 Bark (approximately 430 Hz) and a slightly larger value of γ below that frequency to model the slightly faster growth of perceived loudness at those lower frequencies and [50] attributes these γ values to [44]. This step completes the operations of the perceptual transformation. After processing multiple frames of the input speech signals, the perceptual transformations produce compressed loudness densities that are functions of time (i) and perceptual frequency (j) , $Lx(i,j)$ and $Ly(i,j)$.

The distance measure in the PESQ algorithm is more sophisticated than that used in PSQM. Both a difference time-frequency distribution

$$D(i,j) = Ly(i,j) - Lx(i,j) \tag{17}$$

and a corresponding masking time-frequency distribution

$$M(i,j) = 0.25 \cdot \min\left(Ly(i,j), Lx(i,j)\right) \tag{18}$$

are calculated. For each time-frequency location, the bipolar difference time-frequency distribution is then reduced in magnitude according to the positive masking time-frequency distribution:

$$\begin{aligned} \tilde{D}(i,j) &= D(i,j) - M(i,j) \quad \text{when } M(i,j) \le D(i,j), \\ &= D(i,j) + M(i,j) \quad \text{when } D(i,j) \le -M(i,j) \\ &= 0 \qquad\qquad\qquad\quad \text{otherwise.} \end{aligned} \tag{19}$$

Equation (19) provides a simple model for masking: a signal at a given time-frequency location can raise the threshold for detecting other signals at or near that time-frequency location.

The resulting masked difference time-frequency distribution $\tilde{D}(i,j)$ is called a disturbance density. It is further processed along two different paths so that the PESQ algorithm can respond differently to positive disturbances (due to added time-frequency components) and negative disturbances (caused by missing-frequency components) consistent with the human response to these two types of disturbances. The symmetric path retains both types of disturbances while the asymmetric path retains and emphasizes only the positive disturbances.

Both paths then use a time- and frequency-weighted L_p absolute value aggregation process across frequency:

$$\tilde{D}(i) = \beta_i \cdot \left[\sum_{j=1}^{N} \left(\left| \tilde{D}(i,j) \right| \cdot \alpha_j \right)^p \right]^{\frac{1}{p}}, \tag{20}$$

where α_j are frequency-specific weights and β_i are temporal weights computed from the power of the ith frame of the reference speech signal in a way that emphasizes disturbances that occur during quieter portions of the speech signal and deemphasizes disturbances that occur during louder portions of the speech signal. Once again, this is an operation that is consistent with masking. The norm p=3 is used in the symmetric branch and p=1 is used in the asymmetric branch.

The aggregation of $\tilde{D}(i)$ over time is done at two separate time scales in both the symmetric and asymmetric branches. First, groups of $N = 20$ results (spanning approximately 320 ms) are aggregated using

$$\tilde{D} = \left[\frac{1}{N} \sum_{i=1}^{N} \tilde{D}(i)^{p} \right]^{\frac{1}{p}} \tag{21}$$

with $p = 6$. This models an approximate "max" or "worst-case" type of response over this time window. These results are then aggregated over the entire duration of the speech signals (typically lasting 10 s, $N \approx 31$) in a similar manner using $p = 2$. This models a more forgiving assessment of the perceived disturbances over the longer term. The two final disturbance values created by the two paths (symmetric and asymmetric) are then combined in a linear combination with empirically determined weights to create a final estimate of perceived speech quality that typically ranges from 1 (worst quality) to 4.5 (best quality). The agreement between these values and subject speech quality test results is described in [50].

3.6 Output-Based Speech Quality Estimation

While Figure 3 describes a common approach to speech-based, signal processing-based speech quality estimation, it is not the only approach. One can also estimate speech quality using only the test signal that is output by the SUT. This is called the single-ended or output-based approach and it has both advantages and limitations.

The approach shown in Figure 3 requires that the input and output of the SUT be available at a common location. This is not a limitation for laboratory testing, but for field testing it requires that a separate, distortion-free, communications channel or storage medium be provided. Alternately the SUT can be removed from service, recorded speech signals can be injected into the SUT, and copies of those signals can be compared to the SUT output. Clearly the output-based approach transcends these limitations. The output-based approach can non-intrusively monitor and estimate speech quality using only

the SUT output. An additional advantage is that the issues of gain and delay estimation and compensation disappear. Finally, we note that the output-based approach may appear more pleasing in principle because it provides a closer parallel to the most common subjective testing environment (the ACR environment).

The limitations of the output-based approach stem from the fact that no reference signal is available. While listeners in ACR subjective tests routinely judge speech quality with no reference, the loss of the reference presents a bigger challenge to the signal processing approach. With less information available, it seems inevitable that resulting speech quality estimates must suffer. The perception-based approach described in Figure 5 is no longer possible. The problem moves from finding appropriate measurements of distance in an appropriate perceptual space to the problem of making a perceptually consistent classification of a single signal. Two solutions are summarized below. Examples of other work can be found in [51] and [52].

3.6.1 Auditory Non-intrusive Quality Estimation

The auditory non-intrusive quality estimation (ANIQUE) algorithm [53] incorporates elements of human speech production and human hearing and some of the key steps of are summarized here. The SUT output signal is processed though a bank of 23 critical band filters uniformly spaced by about 0.62 Bark between 125 Hz (1.2 Bark) and 3,500 Hz (14.8 Bark). These filters are implemented in the time domain. Each of the resulting 23 band-limited signals, $s_j(i)$, is then expressed as a carrier or instantaneous phase signal, $\phi_j(i)$, modulated by a temporal envelope, $\gamma_j(i)$, where i is the time sample index and j is the frequency band index. This decomposition follows directly from the introduction of $\hat{s}_j(i)$ which is the Hilbert transform of $s_j(i)$ and the analytic signal $z_j(i) = s_j(i) + \sqrt{-1} \cdot \hat{\boldsymbol{s}}_{\boldsymbol{j}}(\boldsymbol{i})$:

$$s_j(i) = \mathbb{R}\left(z_j(i)\right) = \gamma_{\boldsymbol{j}}(i) \cdot \cos\left(\phi_{\boldsymbol{j}}(i)\right), \tag{22}$$

where

$$\gamma_{\boldsymbol{j}}(i) = \sqrt{s_j^2(i) + \hat{s}_j^2(i)} \text{ and } \phi_{\boldsymbol{j}}(i) = \arctan\left(\frac{\hat{s}_j(i)}{s_j(i)}\right). \tag{23}$$

Based on the known importance of the temporal envelopes and the relative unimportance of the carrier signal or phase information for speech perception, the initial analysis proceeds based on the 23 speech envelopes $\gamma_{\boldsymbol{j}}(i)$ alone. More specifically, it is the spectral content of these envelopes (or modulation spectra) that drive the initial quality estimation. The modulation

spectrum of an envelope is estimated by breaking the envelope into 256-ms processing frames with 75% overlap. A Hamming window and an FFT are applied to each frame and the magnitude of the result is taken. This results in the modulation spectrum $\Gamma_j(k)$, where k is the envelope modulation frequency index.

The next step is to analyze each modulation spectrum and separate the lower frequency components (which can be attributed to natural human speech) from the higher frequency components (which can only be attributed to noise and distortion induced by the SUT) in order to form an articulation-to-nonarticulation ratio (ANR). (In this context, articulation refers to speech production, not speech perception.)

For each frame of the jth envelope, the lower frequency components are those between 2 and 30 Hz (a range consistent with the speed of movement in human articulation) and higher frequency components are those from 30 Hz to $\frac{1}{2}\mathrm{BW_j}$ where $\mathrm{BW_j}$ is the bandwidth of the jth critical band filter. This upper limit of $\frac{1}{2}\mathrm{BW_j}$ is consistent with the highest envelope modulation frequency that is relevant to the perception of speech quality. Thus the modulation power in 2–30 Hz and the modulation power in 30 to $\frac{1}{2}\mathrm{BW_j}$ Hz form the basis for the jth ANR, designated Λ_j. These 23 ANRs are then aggregated across the 23 critical band channels according to

$$\nu = \left[\sum_{j=1}^{23} \Lambda_j^{0.80}\right]^{0.38}, \tag{24}$$

where the exponents were determined empirically to maximize agreement between the ANIQUE output and a large set of subjective test scores.

Equation (24) provides a base quality estimate for each 256-ms frame of SUT output speech. Significant additional processing follows, including explicit algorithm branches to account for unnaturally abrupt starts and stops in the SUT output speech (possibly attributable to temporal discontinuities or transmission channel errors). Separate treatments for loud, faint, and inaudible speech frames are also provided. In addition, a normalization process serves to reduce the variance in ν attributable to input speech articulation and speech content, thus making ν a better indicator of the SUT output quality. Ultimately, all results are combined to generate the final output of the ANIQUE algorithm: a scalar variable that increases to indicate higher estimated speech quality. Relationships between this output and subjective test results are given in [53]. An enhanced version of the ANIQUE algorithm is standardized in [54].

3.6.2 ITU-T Recommendation P.563

A different approach is standardized in ITU-T Recommendation P.563 [55]. This algorithm combines the work of multiple research teams and can be viewed as a large and diverse, yet well-organized set of analysis tools, along with a set of signal-driven rules for automatically selecting, using, and combining the results of the various tools. After preprocessing of the SUT output speech, key parameters are extracted to move through a decision tree to classify the dominant impairment: constant noise, temporal impairments, temporally localized noise, speech with artificially high periodicity, unnatural male speech, or unnatural female speech. Once the dominant impairment has been identified, up to approximately 10 additional parameters may be extracted to enable a basic speech quality estimate customized to that type of impairment. This result is then combined linearly with up to approximately 10 additional parameters to form the final estimate of speech quality. The set of all parameters available covers a wide range from pitch-related parameters, to statistics of linear prediction coefficients, to SNR-related parameters, to tube dimensions in a vocal tract model.

One branch through the algorithm uses a linear prediction-based speech enhancer to form a pseudo-reference speech signal. This is then compared with the SUT output signal (the test signal) using an algorithm that has much in common with the PESQ algorithm. Ultimately, all relevant results are combined to generate the final output of the P.563 algorithm: a scalar variable that increases to indicate higher estimated speech quality.

4 • Conclusion

This chapter has presented key concepts associated with the signal processing approaches to the estimation of speech intelligibility and speech quality. Algorithms that estimate speech intelligibility have generally been developed for, and are most applicable to, acoustic transmission combined with noise, reverberation, and possibly some linear sound reinforcement equipment. Algorithms that estimate speech quality have generally been developed for, and are most applicable to, telecommunications equipment and systems. This segmentation of the field may be partially attributed to the inherent attributes of the two situations, but more generalized solutions are very likely possible and are certainly desirable. One rather demanding example is the need to estimate speech intelligibility in adaptive, non-linear digital communications systems

with non-stationary channel conditions, possibly compounded with acoustic background noise at the talker or listener ends.

This chapter has focused on a small number of example algorithms that estimate speech intelligibility or quality. It has provided some level of detail on each, especially regarding the various signal processing operations invoked and how they can model or at least emulate various relevant attributes of human hearing, speech production, or speech transmission. Necessarily, much diverse and significant work has gone unmentioned in this chapter. Examples of some of this work can be found in [56–66]. A closely related topic is signal processing algorithms that estimate the quality of transmitted music signals. Examples of work in this area can be found in [67–70].

It is natural to ask how well each of these algorithms work. In general, each algorithm described above has been found to be effective in some application area. Any quantitative answer to this question is encumbered with numerous associated issues which could constitute an entire separate field of study. One issue is truth data. It is generally held that carefully controlled subjective tests conducted by standard procedures provide the truth data that a signal processing algorithm should be judged against. However, multiple, equally valid subjective test methodologies are available. Any of these can only generate estimated mean human responses and some confidence interval about each estimated mean. Further, subjective tests generally tend to provide relative, not absolute, results and the rating of a fixed SUT can be very different between two tests, depending on the other SUTs included in each of those tests. This can present a serious mismatch to the signal processing algorithms which are by nature immune to these context effects.

Different techniques have been employed to bridge this mismatch. These can include the use of reference SUTs in subjective tests and per-test mathematical corrections, transformations, or calibrations of subjective test results and/or algorithm results. These techniques are generally required if useful comparisons are to be made, but they also complicate the situation considerably and lead to a multitude of domains where algorithm performance might be reported.

Another issue stems from the range of SUTs that an algorithm is tested on and the relationships between the classes of SUTs used in algorithm development and those used in algorithm testing. The amount of speech data used, the amount of averaging, and the level of analysis (e.g., per sentence, per talker, per SUT) can also have a significant influence on the results reported.

The references provided in this chapter for each algorithm often provide at least one view of how well the algorithm works. Due to the considerations listed here, these views are often not directly comparable to each other. One blind study of multiple algorithms is available in [71] and the results of other comparative studies are described in the internal working documents of various telecommunications standards bodies.

The field of signal processing-based estimation of speech intelligibility and quality has plenty of room for future advances. There is always room for more effective estimators, where "more effective" can mean better agreement with subjective test results over broader classes of SUTs, transmission channel conditions, and acoustic noise and reverberation environments. This would be a particular challenge for output-based estimators. The continued evolution and deployment of telecommunications services and applications (e.g., mobile, hands-free mobile, conferencing, Internet-based) motivates additional development work, including speech intelligibility estimators for telecommunications and speech quality estimators for telecommunications applications with acoustical or electrical interfaces, supporting narrowband (300–3,400 Hz) or wideband (50–7,000 Hz) transmission, and mixtures of multiple talkers and background noises, reverberations, or music. Finally, there is room for significant progress regarding how subjective tests are done, how their results are processed, and how estimation algorithm results are processed and compared with subjective test results so that the most meaningful conclusions can be drawn.

References

1. N. Johannesson, "The ETSI Computation Model: A Tool for Transmission Planning of Telephone Networks," *IEEE Commun. Mag.*, vol. 35, pp. 70–79, Jan. 1997.
2. ITU-T Recommendation G.107, "The E-Model, A Computational Model for Use in Transmission Planning," Geneva, 2003.
3. S. Möller, *Assessment and Prediction of Speech Quality in Telecommunications*. Boston: Kluwer Academic, 2000.
4. S. Quackenbush, T. Barnwell III, & M. Clements, *Objective Measures of Speech Quality*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
5. ITU-T Recommendation P.800, "Methods for Subjective Determination of Transmission Quality," Geneva, 1996.
6. ANSI S3.2-1989, "Method for Measuring the Intelligibility of Speech over Communications Systems," New York, 1990.
7. ISO TR-4870, "The Construction and Calibration of Speech Intelligibility Tests," Geneva, 1991.

8. H. Fletcher, *The ASA Edition of Speech and Hearing in Communication*. J. Allen, Ed. Woodbury, NY: Acoustical Society of America, 1995, Chap. A1.
9. N. French & J. Steinberg, "Factors Governing the Intelligibility of Speech Sounds," *J. Acoust. Soc. Am.*, vol. 19, pp. 90–119, Jan. 1947.
10. H. Fletcher, *Speech and Hearing in Communication*. New York: Van Nostrand, 1953, Chap. 17.
11. K. Kryter, "Methods for the Calculation and Use of the Articulation Index," *J. Acoust. Soc. Am.*, vol. 34, pp. 1689–1697, Nov. 1962.
12. K. Kryter, "Validation of the Articulation Index," *J. Acoust. Soc. Am.*, vol. 34, pp. 1698–1702, Nov. 1962.
13. ANSI S3.5-1969, "Methods for the Calculation of the Articulation Index," New York, 1969.
14. N. Jayant & P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984, Appendix E.
15. H. Steeneken & T. Houtgast, "A Physical Method for Measuring Speech-Transmission Quality," *J. Acoust. Soc. Am.*, vol. 67, pp. 318–326, Jan. 1980.
16. H. Steeneken & T. Houtgast, "Mutual Dependence of the Octave-Band Weights in Predicting Speech Intelligibility," *Speech Commun.*, vol. 28, pp. 109–123, 1999.
17. H. Steeneken & T. Houtgast, "Validation of the Revised STIr Method," *Speech Commun.*, vol. 38, pp. 413–425, 2002.
18. IEC 60268-16, "Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index," Geneva, 2003.
19. T. Houtgast & H. Steeneken, "A Multi-language Evaluation of the RASTI-Method for Estimating Speech Intelligibility in Auditoria," *Acustica*, vol. 54, pp. 185–199, 1984.
20. R. Goldsworthy & J. Greenberg, "Analysis of Speech-Based Speech Transmission Index Methods with Implications for Nonlinear Operations," *J. Acoust. Soc. Am.*, vol. 116, pp. 3679–3689, Dec. 2004.
21. ANSI S3.5-1997, "Methods for Calculation of the Speech Intelligibility Index," New York, 1998.
22. S. Voran, "Listener Ratings of Speech Passbands," in *Proc. 1997 IEEE Workshop on Speech Coding for Telecommunications*, pp. 81–82, Pocono Manor, PA, 1997.
23. J. Beerends, E. Larsen, N. Iyer, & J. van Vugt, "Measurement of Speech Intelligibility Based on the PESQ Approach," in *Proc. 3rd International Conference on Measurement of Speech and Audio Quality in Networks*, pp. 27–30, Prague, Czech Republic, 2004.
24. ITU-T Recommendation P.50, "Artificial Voices," Geneva, 1999.
25. N. Kitawaki, K. Nagai, & T. Yamada, "Objective Quality Assessment of Wideband Speech Coding Using W-PESQ Measure and Artificial Voice," in *Proc. 3rd International Conference on Measurement of Speech and Audio Quality in Networks*, pp. 31–36, Prague, Czech Republic, 2004.
26. M. Werner, T. Junge, & P. Vary, "Quality Control for AMR Speech Channels in GSM Networks," in *Proc. IEEE Interna-*

tional Conference on Acoustics, Speech, and Signal Processing, pp. 1076–1079, Montreal, 2004.
27. A. Karlsson, G. Heikkilä, T. Minde, M. Nordlund, & B. Timus, "Radio Link Parameter Based Speech Quality Index – SQI," in *Proc. 1999 IEEE Workshop on Speech Coding*, pp. 147–149, Porvoo, Finland, 1999.
28. S. Voran, "Compensating for Gain in Objective Quality Estimation Algorithms," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1068–1071, Montreal, 2004.
29. S. Voran, "Objective Estimation of Perceived Speech Quality, Part I: Development of the Measuring Normalizing Block Technique," *IEEE Trans. Speech Audio Process.*, vol. 7, pp. 371–382, Jul. 1999.
30. ITU-T Recommendation P.931, "Multimedia Communications Delay, Synchronization, and Frame Rate Measurement," Geneva, 1998.
31. S. Voran, "Perception of Temporal Discontinuity Impairments in Coded Speech – A Proposal for Objective Estimators and Some Subjective Test Results," in *Proc. 2nd International Conference on Measurement of Speech and Audio Quality in Networks*, pp. 37–46, Prague, Czech Republic, 2003.
32. A. Rix, M. Hollier, A. Hekstra, & J. Beerends, "Perceptual Evaluation of Speech Quality (PESQ) – The New ITU Standard for End-to-End Speech Quality Assessment, Part I – Time-Delay Compensation," *J. Audio Eng. Soc.*, vol. 50, pp. 755–764, Oct. 2002.
33. ITU-T Recommendation P.862, "Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for End-to-End Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs," Geneva, 2001.
34. S. Voran, "A Bottom-Up Algorithm for Estimating Time-Varying Delays in Coded Speech," in *Proc. 3rd International Conference on Measurement of Speech and Audio Quality in Networks*, pp. 43–56, Prague, Czech Republic, 2004.
35. ANSI T1-801-04-2005, "Multimedia Communications Delay, Synchronization, and Frame Rate," New York, 2005.
36. A. Gray Jr & J. Markel, "Distance Measures for Speech Processing," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 24, pp. 380–391, Oct. 1976.
37. S. Voran, "Advances in Objective Estimation of Perceived Speech Quality," in *Proc. 1999 IEEE Workshop on Speech Coding*, pp. 138–140, Porvoo, Finland, 1999.
38. S. Wang, A. Sekey, & A. Gersho, "An Objective Measure for Predicting Subjective Quality of Speech Coders," *IEEE J. Sel. Areas Commun.*, vol. 10, pp. 819–829, Jun. 1992.
39. B. Moore, *An Introduction to the Psychology of Hearing*. London: Academic Press, 1989, Chap. 3.
40. R. Bladon, "Modeling the Judgment of Vowel Quality Differences," *J. Acoust. Soc. Am.*, vol. 69, pp. 1414–1422, May 1981.
41. W. Yang & R. Yantorno, "Improvement of MBSD by Scaling Noise Masking Threshold and Correlation Analysis with MOS Difference Instead of MOS," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 673–676, Phoenix, 1999.

42. J. Beerends & J. Stemerdink, "A Perceptual Speech-Quality Measure Based on a Psychoacoustic Sound Representation," *J. Audio Eng. Soc.*, vol. 42, pp. 115–123, Mar. 1994.
43. ITU-T Recommendation P.861, "Objective Quality Measurement of Telephone-Band (300–3400 Hz) Speech Codecs," Geneva, 1996.
44. E. Zwicker & R. Feldtkeller, *Das Ohr als Nachrichtenempfänger*. Stuttgart: S. Hirzel Verlag, 1967.
45. S. Voran, "A Simplified Version of the ITU Algorithm for Objective Measurement of Speech Codec Quality," in *Proc. International Conference on Acoustics, Speech and Signal Processing,* pp. 537–540, Seattle, 1998.
46. S. Voran & C. Sholl, "Perception-Based Objective Estimators of Speech Quality," in *Proc. 1995 IEEE Workshop on Speech Coding for Telecommunications*, pp. 13–14, Annapolis, MD, 1995.
47. S. Voran, "Objective Estimation of Perceived Speech Quality, Part II: Evaluation of the Measuring Normalizing Block Technique," *IEEE Trans. Speech Audio Process.*, vol. 7, pp. 383–390, Jul. 1999.
48. ANSI T1-518-1998, "Objective Measurement of Telephone Band Speech Quality Using Measuring Normalizing Blocks (MNBs)," New York, 1998. Reaffirmed 2008.
49. ITU-T Recommendation P.861 Appendix II, "Objective Quality Measurement of Telephone-Band (300–3400 Hz) Speech Codecs Using Measuring Normalizing Blocks (MNBs)," Geneva, 1998.
50. J. Beerends, A. Hekstra, A. Rix, & M. Hollier, "Perceptual Evaluation of Speech Quality (PESQ) – The New ITU Standard for End-to-End Speech Quality Assessment, Part II – Psychoacoustic Model," *J. Audio Eng. Soc.*, vol. 50, pp. 765–778, Oct. 2002.
51. C. Jin & R. Kubichek, "Vector Quantization Techniques for Output-Based Objective Speech Quality," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 491–494, Atlanta, GA, 1996.
52. W. Li & R. Kubichek, "Output-Based Objective Speech Quality Measurement Using Continuous Hidden Markov Models," in *Proc. 7th International Symposium on Signal Processing and its Applications*, pp. 389–392, Paris, 2003.
53. D. Kim, "ANIQUE: An Auditory Model for Single-Ended Speech Quality Estimation," *IEEE Trans. Speech Audio Process.*, vol. 13, pp. 821–831, Sep. 2005.
54. ANSI ATIS-PP-0100005.2006, "Auditory Non-Intrusive Quality Estimation Plus (ANIQUE+) Perceptual Model for Non-Intrusive Estimation of Narrowband Speech Quality", New York, 2006.
55. ITU-T Recommendation P.563, "Single-Ended Method for Objective Speech Quality Assessment in Narrow-Band Telephony Applications," Geneva, 2004.
56. V. Peutz, "Speech Information and Speech Intelligibility," Preprint, Audio Engineering Society 85th Convention, Los Angeles, 1988.
57. R. Kubichek, "Mel-Cepstral Distance Measure for Objective Speech Quality Assessment," in *Proc. IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 125–128, Victoria, British Columbia, 1993.

58. A. De & P. Kabal, "Auditory Distortion Measure for Coded Speech – Discrimination Information Approach," *Speech Commun.*, vol. 14, pp. 205–229, Jun. 1994.
59. A. De & P. Kabal, "Auditory Distortion Measure for Coded Speech – Hidden Markovian Approach," *Speech Commun.*, vol. 17, pp. 39–57, Aug. 1995.
60. M. Hansen & B. Kollmeier, "Using a Quantitative Psychoacoustical Signal Representation for Objective Speech Quality Measurement," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1387–1390, Munich, 1997.
61. M. Hauenstein, "Application of Meddis' Inner Hair-Cell Model to the Prediction of Subjective Speech Quality," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 545–548, Seattle, 1998.
62. A. Rix & M. Hollier, "The Perceptual Analysis Measurement System for Robust End-to-End Speech Quality Assessment," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1515–1518, Istanbul, 2000.
63. J. van der Werff & D. de Leeuw, "What You Specify Is What You Get (Parts 1 & 2)," Preprint, Audio Engineering Society 114th Convention, Amsterdam, 2003.
64. J. Holub, M. Street, & R. Smid, "Intrusive Speech Transmission Quality Measurements for Low Bit-Rate Coded Audio Signals," Preprint, Audio Engineering Society 115th Convention, New York, 2003.
65. D. Sen, "Predicting Foreground SH, SL and BNH DAM Scores for Multidimensional Objective Measure of Speech Quality," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 493–496, Montreal, 2004.
66. A. Takahashi & H. Yoshino, "Perceptual QoS Assessment Technologies for VoIP," *IEEE Commun. Mag.*, vol. 24, pp. 28–34, Jul. 2004.
67. T. Thiede, W. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. Beerends, C. Colomes, M. Keyhl, G. Stoll, K. Brandenburg, & B. Feiten, "PEAQ – The ITU Standard for Objective Measurement of Perceived Audio Quality," *J. Audio Eng. Soc.*, vol. 48, pp. 3–29, Jan./Feb. 2000.
68. W. Treurniet & G. Soulodre, "Evaluation of the ITU-R Objective Audio Quality Measurement Method," *J. Audio Eng. Soc.*, vol. 48, pp. 164–173, Jan./Feb. 2000.
69. ITU-R Recommendation BS.1387-1, "Method for Objective Measurements of Perceived Audio Quality," Geneva, 2001.
70. B. Moore & C. Tan, "Measuring and Predicting the Perceived Quality of Music and Speech Subjective to Combined Linear and Nonlinear Distortion," *J. Audio Eng. Soc.*, vol. 52, pp. 1228–1244, Dec. 2004.
71. L. Thorpe & W. Yang, "Performance of Current Perceptual Objective Speech Quality Measures," in *Proc. 1999 IEEE Workshop on Speech Coding*, pp. 144–146, Porvoo, Finland, 1999.

29

Gaussian Models in Automatic Speech Recognition

Jeff Bilmes

Department of Electrical Engineering, University of Washington, Seattle, WA, USA

Abstract Most automatic speech recognition (ASR) systems express probability densities over sequences of acoustic feature vectors using Gaussian or Gaussian-mixture hidden Markov models. In this chapter, we explore how graphical models can help describe a variety of tied (i.e., parameter shared) and regularized Gaussian mixture systems. Unlike many previous such tied systems, however, here we allow sub-portions of the Gaussians to be tied in arbitrary ways. The space of such models includes regularized, tied, and adaptive versions of mixture conditional Gaussian models and also a regularized version of maximum-likelihood linear regression (MLLR). We derive expectation-maximization (EM) update equations and explore consequences to the training algorithm under relevant variants of the equations. In particular, we find that for certain combinations of regularization and/or tying, it is no longer the case that we may achieve a closed-form analytic solution to the EM update equations. We describe, however, a generalized EM (GEM) procedure that will still increase the likelihood and has the same fixed-points as the standard EM algorithm.

1 • Introduction

In most speech recognition systems, Gaussian and Gaussian-mixture hidden Markov models (HMMs) are used to represent density functions over sequences of acoustic observation vectors. The acoustic observations most often used are Mel-frequency cepstral coefficients (MFCCs) [36], perceptual linear prediction (PLP) [35], or variants thereof—these features are particularly well suited to Gaussian models and have been shown to work well in practice. Moreover, when training such models, the expectation-maximization (EM) update equations are well understood [15,52,6] and widely used.

As the accuracy of automatic speech recognition (ASR) systems has improved over the years, they have naturally become more complex as well. For example, they possess more underlying system parameters, most often in the form of more HMM states and/or more Gaussian components within each Gaussian mixture. The dimensionality of each Gaussian component has also typically increased. If the amount of training data does not increase correspondingly, then a parameter optimization procedure (such as EM) that operates without parameter constraints will produce a poorly trained system. The problem can be explained using a bias-variance argument [31]: the parameter variance increases as the number of parameters grows without an increase in training data. One must thus resort either to optimization under hard constraints or to regularization, or to both.

In speech recognition, optimization with hard constraints most often takes the form of parameter sharing, where a priori fixed decisions are made regarding which Gaussian components should share its parameters with some other Gaussian component. Often these decisions can be made based on phonetic or other high-level knowledge. For example, in a tri-phone system, the center state's Gaussian mixture might share many or all of the same Gaussian components regardless of context if the center phones are similar enough. Often the sharing is decided in a combined knowledge- and data-driven fashion [62]. Other forms of constrained optimization are possible, including partial Gaussian parameter sharing, where a sub-portion of different Gaussian components are shared [23], or sparse Gaussians [8], where Gaussian parameters are constrained to contain a certain set of zeros.

We note that parameter sharing, however, may come at a cost since learning algorithms often assume parameter independence [32,33,34,19], which intuitively means that adjusting one part of a model (either structurally or parameterwise) will not

influence another part. In this case, penalized log likelihoods will decompose into separate terms, which may be optimized separately. With parameter sharing, this need not be the case. Indeed, both dynamic Bayesian networks [14,45] and probabilistic relational models [20] can be thought of as network templates and rules, where the rules explain how a template may be expanded into a larger network, and how this larger network should share its parameters among the resulting expanded factors.

Another common procedure to avoid over-training a complex model given a fixed amount of data is regularization, which is generally applicable to regression, classification, and statistical model selection. Regularization is an example of Occam's razor, which roughly states that out of all possible models that can accurately explain an observable phenomenon (i.e., in our case, the data), we should prefer the simplest one possible. There are a variety of ways of measuring simplicity, including Kolmogorov complexity, the VC-dimension, entropy, minimum description length, number of total parameters, Bayesian information criterion, or a general Bayesian prior. From a pattern recognition perspective, we desire a model that achieves a low empirical error on training data while being as simple as possible. The trade-off between these two goals is determined using a "trade-off coefficient" which determines the regularization path [30] or the "accuracy-regularization frontier." This concept is the key behind support-vector and Kernel methods [57], where the function to minimize is a sum of a (hinge) loss function and a regularizer (which is typically the ℓ_2-norm due to its mathematical tractability). Regularization is also commonly used for training neural networks (weight decay [9]), regression (i.e., ridge-regression [31]), and splines [58].

Parameter sharing can also be seen as a form of regularization [31], as the goal of sharing is to optimize the parameters of a model that has been hard constrained to be "simple." In the parameter sharing case, however, simplicity corresponds roughly to the number of free parameters that are available to explain the data, but these free parameters may be pieced together in arbitrarily intricate ways. In other words, more regularization corresponds essentially to more sharing, or to reducing the number of free parameters, but collections of these parameters may combine in a large variety of ways to produce the portion of a model that governs a specific phenomenon needing to be represented (e.g., an individual phone in a speech utterance). There has not, however, been much discussion in the speech recognition literature regarding general forms of

regularization as it may be applied to Gaussian mixture modeling other than what has appeared in the Bayesian adaptation area (see, e.g. [26,11]). Recently, the notion of regularization has been applied to produce new forms of speaker adaptation [43,42] for both generative models (such as HMMs) and discriminative models (such as neural networks, support vector machines, and conditional random fields).

In this chapter, we explore how Gaussians, Gaussian mixtures, Gaussian mixture HMMs, and auto-regressive Gaussian mixture HMMs (what we end up calling a "generic Gaussian mixture") can benefit from the simultaneous use of fine-grained sub-Gaussian parameter tying and also parameter regularization. Moreover, we utilize aspects of graphical models to assist in our exposition. Graphical models are a visual mathematical abstraction that are used to describe certain properties of families of probability distributions [40,37,50]. Different types of graphical model (e.g., directed or undirected) can be used to represent factorization properties of such distributions; factorization is the key that allows the distributive law to be used so that algorithms can be developed that are both mathematically sound and algorithmically efficient. Alternatively, many researchers have also ascribed certain high-level meanings (such as causality) to various types of graphs [51]. There have in the past been descriptions of how graphical models can be used to describe certain aspects of speech recognition systems [5]. In this chapter, we use graphical models only to assist us in describing how Gaussian mixture HMMs can be decomposed into partially tied components.

Our main contribution is to derive general forms of EM learning equations in these models. We consider such models "acoustic" since in the field of automatic speech recognition, the Gaussian is where the rubber meets the road—specifically, we consider here only the portion of the statistical model that makes a connection between the acoustic speech signal (or better, the acoustic signal after having undergone some non-probabilistic processing, such as MFCC feature extraction [36]) and the inherent discrete processing that ultimately must correspond to word string (or other high- or meta-level) hypotheses. These models include, as described above, Gaussians, and their various flavors, including sparse Gaussians, auto-regressive Gaussians, and what are known as "trajectory models", where a Gaussian model regresses (or conditions) on other elements in the same or other feature streams. We treat in some detail the various forms of sub-Gaussian parameter sharing (using a tri-part decomposition of a Gaussian's parameters based on Cholesky factorization) that can arise under these scenarios. We also describe

the training algorithms using a regularized EM framework. In particular, we find that for certain combinations of regularization and tying, we no longer arrive at a closed-form solution to the EM update equations. We describe, however, a generalized EM (GEM) procedure that will still increase the likelihood and has the same fixed-points as the EM algorithm.

This chapter consists of the following sections. Section 2 gives a brief background on maximum likelihood training, Gaussians, and Gaussian mixtures. Section 3 derives EM update equations for training regularized Gaussians, using a three-way decomposition of the parameters of each Gaussian component. Next, Section 4 derives EM update equations for arbitrary patterns of tying, and these equations immediately lead to adaptation algorithms for mixture conditional Gaussians, and also a regularized maximum-likelihood linear regression [41] procedure. Section 5 describes various methods for training these parameters, and explains how one method is a generalized EM, and lastly Section 6 concludes. Of particular relevance to this chapter (both conceptually and in terms of useful notation) are references [6,5,4].

2 • Background

We begin with notation. Random variables will be depicted using capital letters (e.g., X, Y and Q), and lower case variables will be used to indicate their possible values (x, y, q). Given a set of M variables, the index set $1 \ldots M$ will be depicted using the matlab-like notation $1:M$. Given a subset of variable indices $A = \{a_1, a_2, \ldots, a_{|A|}\}$, where $a_i \in 1:M$, then $X_A = \{X_{a_1}, X_{a_2}, \ldots, X_{a_{|A|}}\}$.

Many of the models we describe correspond to temporal sequences of data. We have a size-N sample of training data $\mathcal{D}^{tr} = \{x^{(i)}, q^{(i)}\}_{i=1}^{N}$, where $x^{(i)} = x_{1:T_i}^{(i)}$ is a length T_i sequence of observation vectors. Each observation vector $x_t^{(i)} = \{x_{t,1}^{(i)}, x_{t,2}^{(i)}, \ldots, x_{t,M}^{(i)}\}$ will be of length M, and may correspond to a processed acoustic frame of speech (e.g., it could be a sequence of MFCCs and their deltas). The mth scalar element is denoted $x_{t,m}^{(i)}$, and in general we denote the mth element of vector x as $x_{.,m}$. Thus, we have that $x = x_{.,1:M}$. Also, $q_{1:T_i}^{(i)}$ is a length T_i sequence of integer labels or states, so that $q_t^{(i)}$ is the tth integer of the ith training sequence. When unambiguous, we will often use $\{x, q\}$ as a particular training pair. Each of the training pairs is assumed to be drawn from the same unknown underlying distribution, so that $(x^{(i)}, q^{(i)}) \sim p(x, q)\ \forall i$, where $p(x, q)$ factors according to the rules that define an HMM. We will also discuss static (non-temporal) data. In this case, we will treat $\{x^{(i)}, q\}$ as a training pair, where $x^{(i)}$ is a fixed length

vector. In this case, $p(x, q)$ will be a fixed length distribution. We will express conditional independence relations between random variables using the standard notation $X \perp\!\!\!\perp Y | Z$ [13], which is read "X is independent of Y given Z."

Most speech recognition systems utilize hidden Markov models (HMMs) with mixture density functions for their observation distributions. Specifically, an HMM is a probability distribution over $2T$ variables—there are T state variables $Q_{1:T}$ that are often hidden, and T other variables that are almost always observed $X_{1:T}$. The joint distribution over these $2T$ variables factorizes as follows:

$$p(x_{1:T}, q_{1:T}) = \prod_t p(x_t|q_t)p(q_t|q_{t-1}).$$

Here, $p(q_t|q_{t-1})$ gives the probability of moving from state q_{t-1} at time $t-1$ to state q_t at time t. In this work, we will be discussing the observation distribution $p(x_t|q_t)$ as that is where the acoustic information is typically produced as input into the HMM at time t. For example, with x_t observed, it is often the case that x_t consists of an M-dimensional MFCC feature vector for time t, although many other types feature vectors can be used as well.

In this chapter, x_t will be a continuous valued M-dimensional vector with density $p(x_t|q_t)$. This density will be a mixture model, where components of the mixture consist of Gaussian densities of some form. For example:

$$p(x_t|q_t) = \sum_\ell \alpha_{\ell q_t} \mathcal{N}(x_t; \mu_{\ell q_t}, \Sigma \ell q_t),$$

where $\mathcal{N}(x; \mu, \Sigma)$ indicates that x has an M-dimensional Gaussian density, and where $\alpha_{\ell q_t}$ is the mixture component weight. An M-dimensional Gaussian density has the form

$$p(x) = \mathcal{N}(x; \mu, \Sigma) = |2\pi\Sigma|^{-1/2} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)},$$

where μ is an M-dimensional mean vector, and Σ is an $M \times M$ covariance matrix. We will use $K = \Sigma^{-1}$ to refer to the inverse covariance (or the concentration) matrix of the density.

It is possible to view a Gaussian density either as a Bayesian network [50] or as a Markov random field—the properties of Gaussians being represented by various forms of graphical model were described in [5] and are very briefly summarized here. First, zeros in the covariance matrix correspond to marginal independence assumptions. For example, given the constraint $\Sigma_{ij} = 0$ for a particular i, j of the concentration matrix, then in the marginal we have $p(x_i, x_j) = p(x_i)p(x_j)$ or that $X_i \perp\!\!\!\perp X_j$. This property of

Gaussians is quite well known. Second, zeros in the concentration matrix correspond to conditional independence assumptions. In particular, if $K_{ij} = 0$ then $X_i \perp\!\!\!\perp X_j | X_{\{1:M\}\setminus\{i,j\}}$. In other words, zeros in the concentration matrix correspond to the pairwise Markov property of undirected graphical models [40]. Since the Gaussian is always positive, the Hammersley–Clifford theorem [40] applies and so also do all the standard Markov properties on undirected graphical models, and we can view a Gaussian as an undirected graph on a set of nodes where there are edges between two nodes i and j if and only if $K_{ij} \neq 0$.

A Gaussian may also be described by a Bayesian network. Assuming a particular ordering of the elements in the variable X, if we Cholesky factorize [27,29] the concentration matrix $K = U'DU$ where U is an upper triangular matrix with unity on the diagonal, and D is a diagonal matrix, then form $U = (I - B)$, we see that B is an upper triangular matrix with zeros on the diagonal, and the upper off-diagonal elements correspond to the edges in a Bayesian network that describes the Gaussian. For example, if $B_{ij} = 0$ with $j > i$, then in the Bayesian network form, there is no edge from node j to node i. With this form of Gaussian (one that we will use extensively in this chapter), the Gaussian distribution becomes [5]:

$$p(x) = (2\pi)^{-M/2} |D|^{1/2} e^{-\frac{1}{2}(x - Bx - \tilde{\mu})^T D (x - Bx - \tilde{\mu})}.$$

We may be interested, for example, in training a Gaussian model where K is constrained to have a certain pattern of zeros. In fact, *Sparse Gaussians* [8] are produced by learning a Gaussian distribution that factors with respect to a particular graph. The ubiquitously utilized and easy to train diagonal covariance matrix, for example, simply corresponds to using a completely disconnected graph. Other forms of constrained Gaussians can be formed simply by utilizing a different graph (i.e., various constraints on either the covariance, concentration, or Cholesky factorized concentration matrices). But as we will see in subsequent sections, regardless of the pattern of zeros in the Cholesky factorized form, the Gaussian will be easy to estimate in a maximum likelihood sense as long as all variables are observed.

A conditional (or autoregressive) Gaussian process is essentially a type of conditional probability density. Suppose that the vector x has been bi-partitioned into two sets U, V such that $U, V \subset 1 : M$ and $U \cap V = \emptyset$. Then $p(x_U | x_V)$ is a conditional Gaussian where the mean vector x_U is a linear function of the parent variable vector x_V, but otherwise X_U is a Gaussian. In other words

x_U has a different Gaussian distribution for each value x_V. This is expressed as

$$p(x_U|x_V) = |2\pi\Sigma_{U|V}|^{-1/2} e^{-\frac{1}{2}(x_U-(B_{U|V}x_V+\mu_{U|V}))^T\Sigma_{U|V}^{-1}(x_U-(B_{U|V}x_V+\mu_{U|V}))},$$

where $\Sigma_{U|V}$ is the conditional covariance matrix, which does not change with varying x_V. The conditional mean, however, does change with x_V as the affine transformation $B_{U|V}x_V + \mu_{U|V}$ is the conditional mean of x_U given x_V. This general form can arise in a number of ways. For example, after factoring a Gaussian using the chain rule of probability

$$p(x_{.,1:M}) = \prod_m p(x_{.,m}|x_{.,1:m-1}),$$

then $p(x_{.,m}|x_{.,1:m-1})$ corresponds to a conditional Gaussian (in this case $U = \{m\}$ and $V = \{1,\ 2, \ldots, m-1\}$), and in fact the parameters associated with each factor combine together to produce the Cholesky factorization of the concentration matrix K [3].

Using conditional Gaussians, either normal Gaussian distributions or mixtures thereof can be formed over a speech observation at time t, to form $p(x_t)$. Moreover, trajectory type models can be formed over time, where the conditional variables at time t can come from the past, future, present, or all of the above relative to t. The conditional variables might even come from an entirely separate feature stream. For example, linear mixture buried Markov models (BMMs) [7,3] can be formed by employing a mixture of conditional Gaussian distributions, where in this case we equate x_U with x_t, so x_U becomes the feature vector at the current time, and we also equate x_V with z_t, which is an L-dimensional vector of feature elements from previous times, future times, or as mentioned above even different feature streams (that may come from any time). We can denote this using sets, as $z_t \subset \{x_{t-\tau_{11}}, y_{t-\tau_{12}}, \ldots, x_{t-1}, y_{t-1}, y_t, x_{t+1}, y_{t+1}, \ldots, x_{t+\tau_{21}}, y_{t+\tau_{22}}\}$ where y_t might be a different feature stream altogether, and where τ_{ij} are arbitrary fixed integer constants that indicate how much of either the x or y stream are part of z_t. Much has been written on how conditional Gaussians and their mixtures can be used as trajectory models (examples include [53,5]). In the most general case, z_t may also contain a constant element of unity (1) so that both the regression coefficients and the fixed Gaussian mean can all be represented by the single $M \times (L+1)$ matrix B—in other words, if we have that $z_t = [z_t'\ 1]^T$ indicating that the very last element of z_t is unity, then $Bz_t = B'z_t' + \mu$ where μ is the last column of B, and B' is an $M \times L$ matrix consisting of all but

the last column of B. This produces a generalized conditional Gaussian of the form:

$$p(x|z) = (2\pi)^{-M/2}|D|^{1/2}e^{-\frac{1}{2}(x-Bz)^T D(x-Bz)},$$

where we assume that everything that might possibly be conditioned on is included in the large length L vector $\boldsymbol{z}$. We note that the diagonal matrix D corresponds to the inverse conditional variance parameters of the Gaussian (in this chapter, we sometimes refer to D as just the conditional rather than the inverse conditional variances when the meaning is clear).

In general, regardless of how we form a Gaussian (either as a conditional Gaussian $p(x|z)$ or as part of a standard Gaussian that has been factored using the chain rule), we will define a set of parameters of a Gaussian as the following triple: (1) the mean μ; (2) the (conditional) variance D; and (3) the coefficients of B. The matrix B could come from the Cholesky factorized concentration matrix $K = (I - B)^T D(I - B)$, or could also be the regression coefficients on some conditional vector $\boldsymbol{z}$, or it could come from both; we will not make a distinction between the two since they can be treated identically in the learning equations. Note, however, that in many cases B will be quite sparse—a particular sparsity pattern, then, can be viewed as a form of constrained optimization or itself a form of regularization, where fewer parameters are available in the model to adjust thus reducing parameter estimation variance. This decomposition of the Gaussian's parameters into these three portions (the mean, inverse conditional variance, and regression coefficients) will be the basis for our Gaussian parameter regularization and tying strategies described in Section 4. This treatment of the Gaussian parameters will moreover allow us to represent various training procedures for Gaussians and Gaussian mixtures in a unified way, regardless if the model consists of regular Gaussian components, various sparse Gaussians, or the conditional Gaussians used for temporal trajectory modeling (say in an HMM) or in a multi-stream model.

2.1 Maximum Likelihood Training

Given a set of training data, there are many ways to optimally adjust the parameters to be in best accordance with this data, including maximum likelihood (ML), maximum a-posteriori (MAP), and discriminative parameter training procedures [1,17, 16,39,38,31,60,55]. While discriminative parameter training is most appropriate for classification tasks, ML training is still commonly used and can sometimes perform as well or better than discriminative methods, often due to the mathematical or computational simplicity of such techniques, or to statistical estimation

issues [48]. Moreover, it is often extremely easy to train a generative model in the ML-framework due to the fact that the ML objective function decomposes exactly as does the generative model (e.g., an HMM). In fact, when all variables in question X and Q are observed, then unconstrained ML training becomes quite easy, namely $p(x, q) = p(x|q)p(q)$ can often be learned simply by counting and averaging.

In the models that we consider, namely $p(x, q)$ we will assume that X is observed and Q hidden when training. The most common form of training in this situation is to use the EM algorithm [15,6,46]. We very briefly outline this procedure in its basic form. Assuming that all of the parameters of the model are described by the variable λ, we start with an initial estimate of the parameters, say λ^p. The goal is to optimize the next set of parameters λ using an augmented function,

$$f(\lambda) = Q(\lambda, \lambda^p) = E_{p(Q|X,\lambda^p)}[\log p(X, Q|\lambda)] = \sum_q p(q|x, \lambda^p) \log p(X, Q|\lambda).$$

This is to be compared to the "incomplete" likelihood function, consisting only of the observed variables, namely $g(\lambda) = \log p(x|\lambda)$. It can be shown that $\partial f(\lambda)/\partial\lambda|_{\lambda=\lambda^p} = \partial g(\lambda)/\partial\lambda|_{\lambda=\lambda^p}$ so the gradients of both the incomplete and the augmented likelihood functions are the same at the location of the current parameters. Moreover, it can be shown that $f(\lambda)$ is a concave lower bound on the true likelihood function at the current point, and that optimizing $f(\lambda)$ is often much easier than optimizing $g(\lambda)$, and therefore is typically preferred. The process is often as simple as finding the unique peak of $f(\lambda)$ by taking a first derivative, setting it to zero, and then solving for λ (possibly with some Lagrangian constraints to ensure we stay within the parameters' probability simplex). This process is repeated. There are also guaranteed convergence properties—once converged, we are guaranteed to have found a local maximum of the likelihood function $g(\lambda)$, which follows from the above since the derivatives are identical.

Details of the derivation of the EM update equations for Gaussian mixtures, and HMMs using Gaussian mixtures are described in [15,46], and notation most similar to what we use in this chapter is described in [6]. In an HMM system, in order to train a Gaussian mixture HMM, one needs the quantities $p(Q_t = i, M_t = j|x_{1:T}, \lambda^p)$, the posterior probability that the Markov chain state at time t is i and the mixture component of state i at time t is j. For a plain Gaussian mixture system, the posterior needed is $p(M_t = j|x_t, \lambda^p)$, which is the posterior of the component for data instance x_t. There is a strong commonality

between Gaussian mixtures and Gaussian mixture HMMs, and in fact the only difference in the EM update equations is the form of this posterior. In fact, a Gaussian mixture HMM may be seen as a high dimensional multi-variate Gaussian, specifically $p(x_{1:T}|m_{1:T}, q_{1:T})$ is just an MT-dimensional Gaussian with a block-diagonal covariance matrix. Therefore, in accordance with our unified treatment of Gaussian models as described in the previous section, we will assume what we call a "generic Gaussian mixture", where we use a simple notation p_{mt} to designate the posterior probability under the previous parameters of the current component m for time or sample t, and we will do this using the same notation. As we will see, this will greatly simplify our equations. The general scenario is shown in Figure 1. Lastly, we will let $\mathcal{Q}$ be the general auxiliary function to be optimized in EM. For the purposes of this manuscript, it is also sufficient to provide a $\mathcal{Q}$ only for a generic Gaussian mixture.

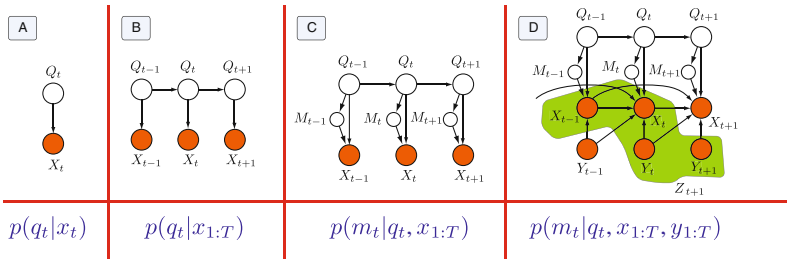


FIGURE 1 A variety of Gaussian-like models utilize the same form of posterior probability of the hidden variables ($q_{1:T}$ and $m_{1:T}$) given the evidence ($x_{1:T}$ or $y_{1:T}$). For each model, the posteriors are needed when training using regularized EM, and the equations for training the different models are identical up to the form of this posterior. In (**A**) (on the *left* of the figure), we have a simple mixture model, with posterior form $p(q_t|x_t)$. In (**B**), we have the basic Gaussian HMM, and its posterior. If we use a Gaussian mixture HMM (**C**), we need the posterior of the component m_t given the current hidden state q_t and the evidence. And last, in (**D**), we have a multi-stream Gaussian mixture auto-regressive (or trajectory) model. The main stream is $X_{1:T}$ with auxiliary stream $Y_{1:T}$, but the form of posterior needed is again still the same. In all cases, a forward–backward procedure [2] can be defined that can compute these posteriors. For our analysis, we will simply use p_{mt} to indicate the posterior of Gaussian component m at data sample (or time) t. Lastly, part D of the figure also shows that Z_t is a vector consisting of elements from X_{t-1}, X_t, Y_t, and Y_{t+1}. Our equations will use the generic notation z_t to indicate this vector, and to produce the other models it is sufficient to realize that $\boldsymbol{z}_t$ may be empty.

3 • Training Regularized Gaussian Decompositions

In this section, the decomposition of a Gaussian component's parameters, as mentioned in the previous section, into the three groups, μ, D, and B, is applied during EM training of a generic Gaussian mixture. Unlike Section 4, in this section we do not yet assume there is any sub-Gaussian parameter sharing.

A desirable property of any EM update equation is that each parameter group at each EM update can be estimated independently without needing any of the other new parameters. In other words, given a collection of previous parameters $\{\mu_m^p,\ D_m^p,\ B_m^p\}_m$, then ideally the update equations of each of the new parameters $\{\mu_m,\ D_m,\ B_m\}_m$ is expressed entirely in terms of only the data and the previous parameters. Otherwise, either multiple passes through the data might be required thereby making training much more computationally expensive, or even worse there might be cyclic dependencies in the parameter updates (which we ultimately will need to deal with in this chapter as well). A simple example may be demonstrated using the estimation of the mean $\mu \propto \sum_t x_t$ and variance $\sigma^2 \propto \sum_t (x_t - \mu)^2$, where we see that the variance depends on the mean. It is well known that the variance estimate can be reformulated using additional accumulators, namely $\sum_t x_t^2$ and $(\sum_t x_t)^2$, requiring only one pass through the data for both mean and variance, in this latter case accumulating both the sum and sum of squares of x_t. For certain patterns of sharing and regularization of a Gaussian's parameters, however, this type of trick will not always be possible. In any case, the general property as it will apply to our EM updates is depicted in Figure 2.

When there is no parameter sharing, the EM auxiliary function needing to be maximized becomes

$$\mathcal{Q} = \sum_m \sum_t \left(\frac{1}{2} \log(|D_m|) - \frac{1}{2}(x_t - B_m z_t)^T D_m (x_t - B_m z_t) \right) p_{mt} - \frac{\lambda}{2} \sum_m ||B_m||_F^2, \quad (1)$$

where $||A||_F = \mathrm{tr}(AA^T)$ is the Frobenius norm of matrix A, the last term acts as an ℓ_2 regularizer on the regression coefficients, and λ is an accuracy-regularization trade-off coefficient. This form of regularization is similar to ridge-regression [31] and to weight-decay [9] in the neural-networks literature. In our case, however, we assume that there is only one such global accuracy-regularization λ coefficient for all the Gaussians in the current model, and for all scalar elements within each component. An

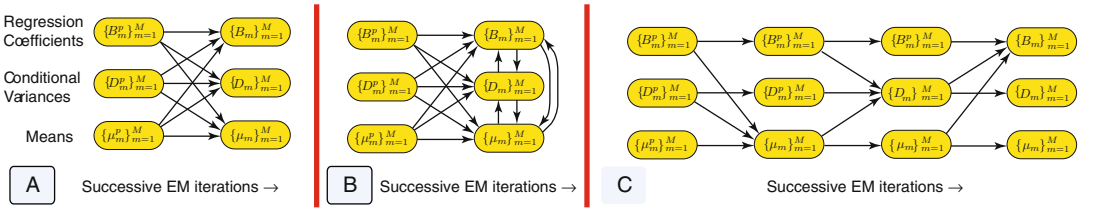


FIGURE 2 EM Gaussian decomposition parameter update dependencies: On the *left* (**A**), we see that each of the new parameters μ_m, B_m and D_m depends only on previous parameters μ_m^p, B_m^p and D_m^p and also on the training data set (not shown in the figure). In this case, it is possible to simultaneously update all new parameters, with only one pass through the training data. In the *middle* (**B**), the new parameters might depend on the other new parameters as well as old parameters, leading potentially to cyclic dependencies. On the *right* (**C**), we break the cycles by only updating one parameter group for each EM iteration, where each EM iteration requires a single pass through the data. This means that to update all the parameters in the model, three passes through the data are required. Still another strategy (*middle* figure, or **B**) updates all parameters simultaneously with one training data pass, but in place of each new parameter that is needed, we use the parameter from the previous EM iteration—this last step is not guaranteed to increase the likelihood of the data, although it does have the same fixed points as the left case, and may work well in practice.

alternative approach would employ one coefficient for each separate Gaussian component, or even one coefficient for each individual regression coefficient. In general, however, the value of such coefficients can only be found using cross-validation [31] on a development or held-out data set and doing this when there are multiple accuracy-regularization coefficients would turn the problem of finding these trade-off coefficients themselves into an intractable search problem.

Equation (1) does not include any constant terms which vanish once derivatives are taken. Also note, for now we have assumed, that the mean parameters are contained within the B matrix—we do this by assuming that one of the elements of z_t is always equal unity. This also means that we are regularizing in the above equation both the regression coefficients (e.g., cross-observation dependences as in a BMM) and the means. We might not wish to regularize the means, however, so in an expansion where the means and regression coefficients are separate (see below), we might not have a regularizer on the means (we utilize separate mean and regression regularization coefficients in Section 4). Also, since there is no sharing, there is a one-to-one correspondence between every value of m and each of the D and B matrices.

Maximizing the regularizer alone $-\frac{\lambda}{2}\sum_m ||B_m||_F^2$ will force the coefficients to converge toward zero. Therefore, it is important to assume that the training data is centered, that is, zero-mean normalized. Also, the coefficients will depend on the scaling of the data as well. Therefore, in practice, it will be important to perform both mean subtraction and variance normalization [44,10] on the observation vectors prior to any training.

Taking derivatives of Eq. (1) with respect to the B_m matrices (again, which includes the mean in the right most column of B_m), gives $\frac{\partial \mathcal{Q}}{\partial B_m} = 0$ or that

$$\sum_t D_m(x_t - B_m z_t) z_t^T p_{mt} - \lambda B_m = 0.$$

From this, we must attempt to solve for B_m, and there are several ways to proceed. One simple way next yields

$$\sum_t (x_t - B_m z_t) z_t^T p_{mt} - \lambda D_m^{-1} B_m = 0, \tag{2}$$

implying that

$$\sum_t x_t z_t^T p_{mt} - B_m \sum_t z_t z_t^T p_{mt} - \lambda D_m^{-1} B_m = 0.$$

We see that we have a matrix equation of the form $A - BC - DB = 0$ where $A = \sum_t x_t z_t^T\ p_{mt}$, $B = B_m$, $C = \sum_t z_t z_t^T\ p_{mt}$, $D = \lambda D_m^{-1}$, and the goal is to solve for B. This equation is known as the Sylvester matrix equation [25] (which is a generalization of the more well known Lyapunov equation [21] commonly used in control theory). The Sylvester equation has a unique solution if and only if there is no eigenvalue of C that is the negative of any eigenvalue of D. Therefore, since both C and D are positive definite in our case, this condition will certainly hold. A general (but not efficient) form of solution is given by $\bar{b} = G^{-1}\bar{a}$, where $\bar{b}$ (respectively $\bar{a}$) is a vector consisting of the columns of B (resp. $-A$) stacked on top of each other, and where $G = I \otimes D + C^T \otimes I$ where $\otimes$ is the Kronecker matrix product. Much more efficient solutions are given in [25]. Also, since D is diagonal, significant additional computational efficiencies can be obtained (see Section 3.2 below).

If $\lambda = 0$, however, we have a particularly simple form, which yields:

$$B_m = \sum_t x_t z_t^T p_{mt} \left(\sum_t z_t z_t^T p_{mt} \right)^{-1}. \tag{3}$$

So if $\lambda = 0$ (no regularization), the above shows that B_m can be estimated quite easily by summing posterior weighted outer products, and that B_m does not depend on any other current parameter. This assumes of course that T is large enough so that the matrix above is invertible, but that is typical for most data sets.

With regularization $\lambda > 0$, however, the new B_m has a much more intricate relationship with the new D_m matrix, as illustrated in the middle of Figure 2.

Next, taking derivatives of Eq. (1) with respect to the diagonal matrix D_m and setting $\frac{\partial Q}{\partial D_m} = 0$ gives

$$\sum_t p_{mt} \left(D_m^{-1} - (x_t - B_m z_t)(x_t - B_m z_t)^T \right) = 0,$$

or that

$$D_m^{-1} = \frac{\sum_t p_{mt} (x_t - B_m z_t)(x_t - B_m z_t)^T}{\sum_t p_{mt}}.$$

In this update, we again have a cross-dependency, namely D_m is dependent on the current estimation of B_m. In this case, therefore, it appears that even if $\lambda = 0$, we have a dependence between the parameters, thus again implying multiple passes through the data on each EM iteration, first to compute B_m and next to compute D_m. The following formula, however, shows that if $\lambda = 0$, we can compute the covariance matrix using the quantities that we have and without running through the data multiple times per iteration.

$$\begin{aligned} D_m^{-1} &\propto \sum_t \left\{ p_{mt} \left(x_t x_t^T - x_t z_t^T B_m^T - B_m z_t x_t^T + B_m z_t z_t^T B_m^T \right) \right\} \\ &= \sum_t p_{mt} x_t x_t^T - \left(\sum_t p_{mt} x_t z_t^T \right) B_m^T \\ &\quad - B_m \sum_t p_{mt} z_t x_t^T + B_m \left(\sum_t p_{mt} z_t z_t^T \right) B_m^T, \end{aligned}$$

but when $\lambda = 0$, we have

$$B_m \left(\sum_t p_{mt} z_t z_t^T \right) B_m^T = \left(\sum_t x_t z_t^T p_{mt} \right) B_m^T,$$

and therefore we get

$$D_m^{-1} \propto \sum_t p_{mt} x_t x_t^T - B_m \sum_t p_{mt} z_t x_t^T. \tag{4}$$

The equations above show that B_m and D_m can indeed be learnt simultaneously when $\lambda = 0$ without needing to make multiple passes over the data. Specifically, with only a single pass over the data, we can accumulate the three quantities $\sum_t x_t z_t^T p_{mt}$, $\sum_t z_t z_t^T p_{mt}^{-1}$, and $\sum_t p_{mt} z_t x_t^T$, which when done are sufficient for both the B_m and D_m updates in Eqs. (3) and (4).

Even though $\lambda = 0$ in this case, it still might appear that we have a problem in that the above equations require outer products of the form $z_t z_t^T$ and $x_t z_t^T$ for every m. Considering that there may be many hundreds or thousands of Gaussian components in a modern ASR system (corresponding to the different HMM states and mixture components therein), and that z_t could easily have many elements, these accumulators could be quite large. For example, in a BMM update equation [3], z_t consists of the collection of *all* context elements grouped together for all individual Gaussian elements, a vector that could be thousands of elements long. When $\lambda > 0$, solving the Sylvester equation would be prohibitive. A solution to this problem is given in Section 3.2, but we first discuss the relationship these models have with ridge regression.

3.1 Ridge Regression

Our procedure may in fact be seen as a multi-variate, multi-class, "soft", and learned-covariance generalization of ridge regression [31]. It is multivariate since we are essentially learning multiple sets of regression coefficients simultaneously. It is multi-class since the regression coefficients may correspond to entirely different Gaussians. It is "soft", since the coefficients are being learned as part of a larger EM procedure, so all data is being essentially weighted by posterior probabilities from previous model parameters. And lastly, unlike conventional ridge regression, we must learn covariance parameters as well. Since the relation with ridge regression is important to understand, we review ridge regression in some detail, and then expand on its generalization.

Ridge regression can be explained from either a frequentist perspective (a Gaussian model with a penalized log-likelihood) or a Bayesian perspective (the maximum a posteriori (MAP) of β under a Gaussian model and a Gaussian prior). In either case, the goal is to learn the parameters of a linear model for predicting the scalar variable x based on a vector variable $\boldsymbol{z}$ and a vector of parameters β of the same length. We are given training samples $\mathcal{D}^{tr} = \{x^{(i)}, z^{(i)}\}_{i=1}^N$ from which to learn β. In the Bayesian case, we find the β that maximizes the posterior $p(\beta|\mathcal{D}^{tr})$ or equivalently that maximizes $p(\mathcal{D}^{tr}|\beta)p(\beta)$. The likelihood $p(\mathcal{D}^{tr}|\beta)$ is composed of the model $p(x|z,\beta) = \mathcal{N}(x;\ \beta^T z, \sigma^2)$, a conditional Gaussian with mean $\beta^T z$ and variance σ^2. As before, we

assume that the last entry of z is always unity, so that we can absorb any mean parameter into the regression parameters β. We also assume a zero-mean and η^2-variance Gaussian Bayesian prior on β, meaning that $\beta \sim \mathcal{N}(0, \eta^2)$ where $p_\eta(\beta) = \mathcal{N}(0, \eta^2) = \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{1}{2}\beta^T\beta/\eta^2\right)$. The joint data and β distribution consisting of log-likelihood and prior can be written as

$$L(\beta) = \sum_i \log p(x^{(i)}|z^{(i)}, \beta) + \log p_\eta(\beta) \tag{5}$$

$$= -\frac{N}{2}\log(2\pi\sigma^2) - \frac{1}{2}\sum_i (x^{(i)} - z^{(i)}\beta)^2/\sigma^2$$

$$-\frac{1}{2}\log(2\pi\eta^2) - \frac{1}{2\eta^2}||\beta||^2, \tag{6}$$

where $||\beta||$ is the ℓ_2-norm. Finding the maximum a posteriori (MAP) value of β reduces to

$$\underset{\beta}{\text{argmin}} \sum_i \frac{1}{2\sigma^2}(x^{(i)} - z^{(i)}\beta)^2 + \frac{1}{2\eta^2}||\beta||^2$$

$$= \underset{\beta}{\text{argmin}} \sum_i \frac{1}{2}(x^{(i)} - z^{(i)}\beta)^2 + \frac{\sigma^2}{2\eta^2}||\beta||^2.$$

If we set $\lambda = \sigma^2/\eta^2$, we get the typical form of ridge regression:

$$\underset{\beta}{\text{argmin}} \sum_i \frac{1}{2}(x^{(i)} - z^{(i)}\beta)^2 + \frac{\lambda}{2}||\beta||^2. \tag{7}$$

This latest form is in fact the standard frequentist starting point for ridge regression, where we have an error term (or a likelihood) along with a complexity penalty that are related by an accuracy-regularization trade-off coefficient λ.

To increase the degree of regularization, we increase λ, and as $\lambda \to \infty$, $\eta^2 \to 0$, which also takes β toward zero for any finite-length data set. λ thus gives us a single normalized parameter to determine the amount of regularization. We note that the variance of the Bayesian prior changes as a function of the data variance since $\eta^2 = \sigma^2/\lambda$. If we wanted to use a fixed Bayesian prior regardless of the data variance, ridge regression would take the form

$$\underset{\beta}{\text{argmin}} \sum_i \frac{1}{2}(x^{(i)} - z^{(i)}\beta)^2 + \frac{\lambda\sigma^2}{2}||\beta||^2, \tag{8}$$

where in this case, we have an interpretation of our regularization coefficient as $\lambda = 1/\eta^2$, the inverse variance of the Bayesian prior.

Of course, the interpretation of λ in Eq. (7) or (8) does not matter here since the regularization coefficient is usually determined using cross-validation for any given data set [31]. Note, however, that the range of values attempted for cross-validation would matter, as in the former case λ is a ratio of variances, while in the latter case λ is an inverse variance.

Suppose now that we wish to solve multiple ridge-regressions simultaneously. We can think of this as multiple separate data sets or, alternatively, we can view this as a single multivariate generalization of ridge regression, where x is now a M-dimensional vector, B is an $M \times L$ matrix of regression coefficients, where each row of B is used to predict the corresponding element of x. In either case, when we wish to use a single λ regression coefficient to control such multiple regressions simultaneously, then the interpretation of λ might start to matter. If we want one global Bayesian prior variance to apply to all regressions simultaneously, then the form of regularization should be as we have defined it in Eq. (2) (for a specific value of m) where each of the variance terms are multiplied by lambda as in λD_m^{-1}. In this case, we interpret λ as the inverse of this globally shared Bayesian prior variance. If, on the other hand, we wish the Bayesian prior for each row of B_m to scale with the variance of the corresponding element in x, then Eq. (2) should take the form

$$\sum_t (x_t - B_m z_t) z_t^T p_{mt} - \lambda I B_m = 0, \tag{9}$$

where I is an appropriately sized identity matrix, in which case B_m has once again a particularly simple update equation, namely Eq. (3) becomes

$$B_m = \sum_t x_t z_t^T p_{mt} \left(\sum_t z_t z_t^T p_{mt} + \lambda I \right)^{-1}. \tag{10}$$

We see that as λ gets larger, the matrix that must be inverted becomes better conditioned, and it moves B toward the zero matrix.

For the remainder of this chapter, we will assume the globally shared Bayesian prior case, but we must keep in mind that some of the equations will change if we take the interpretation shown in Eq. (10).

3.2 Exploiting that D is diagonal

Significant computational simplifications can be achieved by noting that since D_m is a diagonal matrix, only the diagonal portion of the update equations needs to be computed. This requires no loss of generality because of the UDU' factorization mentioned above, whereby the B matrices may contain the

factored form of a full covariance matrix, or a generalized regression on z_t, or both.

We let $D_m(r)$ be the rth diagonal scalar element of the diagonal matrix D_m. Also, let $B_m(r)$ be a row-vector consisting of the non-zero (i.e., existing) elements from the rth row of the matrix B_m, and let $\boldsymbol{z}_t(r)$ be a column vector consisting of the elements of the vector $\boldsymbol{z}_t$ that correspond to the non-zero entries in the rth row of B_m. In other words, the scalar quantity $B_m(r)z_t(r)$ equals the rth element of the vector $B_m z_t$ but the computation of $B_m(r)z_t(r)$ does not include the terms that are just zero. This is a slight abuse of notation as the elements contained in $z_t(r)$ depend on the B_m matrix that it is currently being multiplied with, but the meaning should be clear nevertheless. The reason this works at all is the way we formed z_t as being a subset of all possible elements that x_t might condition on lumped together as one large vector—it has been the sparsity pattern of the B_m matrices which has not only included the appropriate regression coefficients, but also acted as a selector for appropriate elements within z_t. Note also that for $r \neq r'$, $z_t(r)$ and $z_t(r')$ might contain some of the same elements from z_t, that is, treating $z_t(r)$ as a set rather than a vector, it is possible to have $z_t(r) \cap z_t(r') \neq \emptyset$. Finally, we let $x_t(r)$ be the rth scalar element of the vector x_t. Using these definitions, we may re-write Eq. (1) as

$$\mathcal{Q} = \sum_m \sum_t \sum_r \left(\frac{1}{2} \log(D_m(r)) - \frac{1}{2} \left(x_t(r) - B_m(r) z_t(r)\right)^2 D_m(r) \right) p_{mt} - \frac{\lambda}{2} \sum_m \sum_r ||B_m(r)||^2. \tag{11}$$

It follows immediately that the update equations for the diagonal entries of the covariance matrices are

$$\frac{1}{D_m(r)} = \frac{\sum_{t=1}^T p_{mt} \left(x_t(r) - B_m(r) z_t(r)\right)^2}{\sum_{t=1}^T p_{mt}} = \frac{\sum_{t=1}^T p_{mt} [x_t(r)]^2 - B_m(r) \sum_{t=1}^T p_{mt} x_t(r) z_t(r)}{\sum_{t=1}^T p_{mt}}, \tag{12}$$

where the second equality follows only when $\lambda = 0$, similar to Eq. (4). Note that this is a scalar update equation for each of the diagonal elements of the matrix D_m. Depending on the sparse patterns of the matrices, the cost of the dot products $B_m(r)z_t(r)$ will be significantly smaller than before.

The update equations for the B_m matrices can also be significantly simplified. Since D_m is diagonal, the derivative of Eq. (1)

with respect to the rth row of B_m is uncoupled with the other rows. Also, only the non-zero portions of the row are computed (recall that B_m can be quite sparse so this can be a significant savings). Taking the derivative of Eq. 11 with respect to $B_m(r)$ thus yields:

$$\sum_{t=1}^{T} D_m(r)\big(x_t(r) - B_m(r)z_t(r)\big)z_t(r)^T p_{mt} - \lambda B_m(r) = 0.$$

This leads to the independent update equations for the non-zero portion of each row of the B matrix as follows:

$$B_m(r) = \left(\sum_{t=1}^{T} p_{mt} x_t(r) z_t(r)^T\right)\left(\sum_{t=1}^{T} p_{mt} z_t(r) z_t(r)^T + \frac{\lambda}{D_m(r)} I\right)^{-1},$$

where I is an appropriately sized identity matrix. In this case, only the outer products $z_t(r)z_t(r)^T$ and the SAXPY [27] operations $x_t(r)z_t(r)^T$ are required, yielding a significant reduction in both computational complexity and memory requirements. Also, unlike in the general case above, we see that we get a solution without needing to resort to algorithms that solve the Sylvester equation, regardless of the interpretation or value of λ. We see also again that adding the positive definite diagonal matrix $\lambda/D_m(r)I$ in the above ensures that the matrix is well conditioned (so inversion is possible). And as mentioned earlier, unlike typical ridge regression, we take here the interpretation of one global Bayesian prior variance in order to use a global λ parameter, which means we must appropriately scale by the variance $D_m(r)$.

All is not perfect, however, as when $\lambda > 0$, it appears that we once again obtain cycles in the parameter update dependencies, as in the center of Figure 2. Specifically, the equations show that both $B_m(r)$ depends on $D_m(r)$ and $D_m(r)$ depends on $B_m(r)$. Once again, we can break the cycle by using the previous parameters in place of current parameters and make multiple passes through the data, for each pass updating only one parameter group at a time, as shown on the right in Figure 2. Thus, in each step we are still guaranteed to improve the likelihood, although convergence may not be as fast. Moreover, we clearly have the same fixed points, as when at a particular fixed point, we have that $D_m^p(r) = D_m(r)$. We can also (Figure 2 middle) attempt to update all parameters simultaneously, but such a procedure does not guarantee in all cases an improvement in likelihood on each EM iteration.

4 • Training Shared Regularized Gaussian Decompositions

Normally, and as described in the previous section, each individual Gaussian component has its own unique set of parameters not shared with any other component in the model. In this section, we describe a model where some set of the Gaussian components is constrained to share a portion of their parameters with other components in the model. We wish moreover to analyze the EM update equations under such constraints. This procedure is often called parameter tying, but here the granularity of tying is somewhat finer than what is typically done in a speech recognition system. Specifically, an entire Gaussian component is commonly shared between two mixtures of Gaussians, or alternatively an entire mixture for an HMM state might be shared between two multi-state phone models in an HMM-based system [61]. Here, by contrast, each component Gaussian might share any of its μ, D, and B parameters with other components in the system. Moreover, a given component might share different groups with different other components. For example, a component's μ parameter might be shared differently than the component's D parameter. This leads to an enormous variety of parameter sharing patterns. There have been a number of other models that use partially shared Gaussian parameters, including the ones described in [28,23,49], but this is not done in the same way described herein. In particular, we will see that Cholesky factorized covariance matrices allow us to easily combine sharing and other forms of regularization.

It is crucial to realize that sharing is a form of constrained optimization—we still seek a maximum likelihood solution, but now we are constraining the parameters to live in a space where sub-portions of Gaussian parameters are forced to be identical. It is well known that constrained optimization can be more difficult (either computationally or mathematically) than unconstrained optimization [18], and we will see that our case is no exception. We will find, however, that the update rules are still fairly simple if we use the idea we saw in the previous section of breaking cycles over EM iterations, as it then is possible to find a solution without major modifications to the previously derived equations.

In the most general case of parameter sharing, we have three functions that map from the index m of a general class to the index k of a specific extant mean, diagonal covariance, and regression coefficient used by that index. Those mappings are, respectively, $T_\mu(m)$, $T_D(m)$, and $T_B(m)$. In other words, we have that $T_\mu(m) : \{1, \ldots, M\} \rightarrow \{1, \ldots, K_\mu\}$ where M is the total number of components

in the system, and K_μ is the total number of unique means in the system. An analogous definition holds for the D and B matrices, where we assume there are a total of K_D unique D matrices, and K_B total unique B matrices in the model. We also define the inverse mappings: $T_\mu^{-1}(k)$, $T_D^{-1}(k)$, and $T_B^{-1}(k)$, which refer to the set of all indices m that lead to a particular index k, that is

$$T_\mu^{-1}(k) \triangleq \{m : T_\mu(m) = k\},$$

and where the analogous definitions hold for $T_D^{-1}(k)$ and $T_B^{-1}(k)$.

Using this representation of tying via the index mapping, the auxiliary function becomes

$$\begin{aligned}\mathcal{Q} = \sum_m \sum_t \Bigg\{ & \frac{1}{2}\log(|D_{T_D(m)}|) \\ & - \frac{1}{2}(x_t - B_{T_B(m)} z_t - \mu_{T_\mu(m)})^T D_{T_D(m)} (x_t - B_{T_B(m)} z_t - \mu_{T_\mu(m)}) \Bigg\} p_{mt} \\ & - \frac{\lambda}{2} \sum_{k=1}^{K_B} ||B_k||_F^2 - \frac{\eta}{2} \sum_{k=1}^{K_\mu} ||\mu_k||^2. \end{aligned} \tag{13}$$

There are several differences between this equation and Eq. (1). First, we have separated out the mean vectors from the B matrices since, in this general case, there could be an entirely different tying pattern for the means and regression coefficients within B. Second, we now have two regularization coefficients, one for the regression coefficients λ and a separate one for the shared mean pool η—note again that this form of regression only makes sense when the data have been centered or zero-mean normalized, as mentioned earlier. Other than that, we still may have a unique posterior probability p_{mt} for each m since various combinations of mean, diagonal covariance, and regression coefficient could still produce a distinct value of p_{mt} for different m and fixed t—the constrained optimization procedure requires that the posterior be filtered down into the appropriate shared component of each Gaussian.

4.1 Covariance Updates

Taking derivatives and setting $\frac{\partial \mathcal{Q}}{\partial D_k} = 0$ implies that

$$\sum_t \sum_{m \in T_D^{-1}(k)} p_{mt} \Big(D_k^{-1} - (x_t - B_{T_B(m)} z_t - \mu_{T_\mu(m)}) (x_t - B_{T_B(m)} z_t - \mu_{T_\mu(m)})^T \Big) = 0,$$

or that

$$D_k^{-1} = \frac{\sum_t \sum_{m \in T_D^{-1}(k)} (x_t - B_{T_B(m)} z_t - \mu_{T_\mu(m)})(x_t - B_{T_B(m)} z_t - \mu_{T_\mu(m)})^T p_{mt}}{\sum_t \sum_{m \in T_D^{-1}(k)} p_{mt}}.$$

In this form, it is not possible to compute the numerator without using multiple passes through the data, so we next expand the numerator in order to see what must be computed. Noting that D_k^{-1} is diagonal, and that any additive contributions to D_k^{-1} in the form of $A+A^T$ can be added as $2A$, we get:

$$\begin{aligned} D_k^{-1} \propto \sum_{m \in T_D^{-1}(k)} \Bigg\{ & \left(\sum_t x_t x_t^T p_{mt} \right) - 2 \Bigg[\left(\sum_t x_t p_{mt} \right) \mu_{T_\mu(m)}^T \\ & + \left(\sum_t p_{mt} x_t z_t^T \right) B_{T_B(m)}^T - \mu_{T_\mu(m)} \left(\sum_t z_t^T p_{mt} \right) B_{T_B(m)}^T \Bigg] \\ & + B_{T_B(m)} \left(\sum_t z_t z_t^T p_{mt} \right) B_{T_B(m)}^T + \mu_{T_\mu(m)} \mu_{T_\mu(m)}^T \left(\sum_t p_{mt} \right) . \Bigg\} \end{aligned} \tag{14}$$

From the above, it is particularly easy to see that each major term in the sum does not require any of the other parameters until the very end of the accumulation. For example, computing quantities such as $\sum_t x_t x_t^T p_{mt}$, $\sum_t p_{mt} x_t z_t^T$, and $\sum_t z_t p_{mt}$ can be done in one pass through the data, and at the end combined with the resulting final B matrix or mean vector for the corresponding m in producing the final numerator for the kth covariance matrix.

If there are no B matrices (i.e., all the B entries are zero, which happens when we use purely diagonal covariance matrices), then we have a particularly simple form:

$$\begin{aligned} D_k^{-1} \propto \sum_{m \in T_D^{-1}(k)} \Bigg\{ & \left(\sum_t x_t x_t^T p_{mt} \right) \\ & - 2 \left(\sum_t x_t p_{mt} \right) \mu_{T_\mu(m)}^T + \mu_{T_\mu(m)} \mu_{T_\mu(m)}^T \left(\sum_t p_{mt} \right) \Bigg\} . \end{aligned} \tag{15}$$

4.2 B-Matrix Updates

Next, we find the update equations for the B matrices. Setting $\frac{\partial Q}{\partial B_k} = 0$ yields

$$\sum_{m \in T_B^{-1}(k)} D_{T_D(m)} \sum_t (x_t - B_k z_t - \mu_{T_\mu(m)}) z_t^T p_{mt} - \lambda B_k = 0,$$

or that

$$\begin{aligned} \sum_{m \in T_B^{-1}(k)} D_{T_D(m)} \Bigg\{ & \left(\sum_t x_t z_t^T p_{mt} \right) \\ & - B_k \left(\sum_t z_t z_t^T p_{mt} \right) - \mu_{T_\mu(m)} \left(\sum_t z_t^T p_{mt} \right) \Bigg\} - \lambda B_k = 0. \end{aligned}$$

Again, we have a similar problem in that we are trying to solve for a matrix B in an equation of the form $A - \sum_i D_i B C_i - E - \lambda B = 0$, which is a generalized form of the Sylvester equation. Note that even if we decide not to regularize so that $\lambda = 0$, then Sylvester-like equations still apply since B is still in relation with the other parameters via $A - E = \sum_i D_i B C_i$. In this case, other than the lack of a direct analytical solution for B_k, we see that it depends on other parameters from the current EM iteration, giving us the situation depicted in the center of Figure 2. In particular, B_k depends on all the other mean (μ) and variance (D) parameters that are involved in any of the components $m \in T_B^{-1}(k)$, as described in Figure 3. This means that the dependencies between current parameters as described in the center of Figure 2 will depend entirely on the tying patterns as given by the T functions, each individual parameter having a different set of dependences—in a sense, each parameter is independent of all other parameters given what might be called its variational or parameter Markov blanket.

When $\lambda > 0$ or in order not to solve the general Sylvester equation, we must once again use explicitly the fact that D_k for all k is diagonal—we may write Eq. (13) in a way that is analogous to the way Eq. (1) was written as Eq. (11). Using again $B_k(r)$ to denote the rth row of B_k, and the analogous definitions as used in Section 3.2, we immediately get:

$$B_k(r) = \left(\sum_{m \in T_B^{-1}(k)} D_{T_D(m)}(r) \left\{ \left(\sum_t x_t(r) z_t^T(r) p_{mt} \right) - \mu_{T_\mu(m)}(r) \left(\sum_t z_t^T(r) p_{mt} \right) \right\} \right) \left(\sum_{m \in T_B^{-1}(k)} D_{T_D(m)}(r) \left(\sum_t z_t(r) z_t^T(r) p_{mt} \right) + \lambda I \right)^{-1}, \quad (16)$$

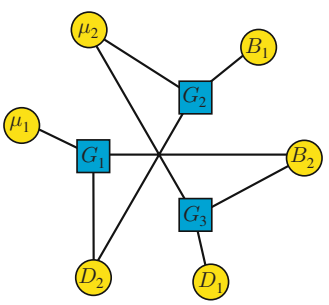


FIGURE 3 Sharing example. There are three Gaussian components, G_1, G_2, and G_3, with components G_2 and G_3 sharing the same mean μ_2, components G_1 and G_3 sharing the same regression matrix B_2, and G_1 and G_2 both sharing D_2. As shown in Eq. (16), the EM update equation for B_2 depends on the other new parameters μ_1, μ_2, D_1, D_2, but B_1's update depends only on μ_2 and D_2.

where I is an appropriately sized identity matrix. We once again see that an effect of the $\lambda > 0$ parameter ensures the invertibility of the matrix in which it is involved.

4.3 Mean Updates

Since the mean parameter can have a separate tying pattern, we must treat it separately from the B matrix. We set $\frac{\partial \mathcal{Q}}{\partial \mu_k} = 0$ to get

$$\sum_{m \in T_\mu^{-1}(k)} D_{T_D(m)}^{-1} \sum_t (x_t - B_{T_B(m)} z_t - \mu_k) p_{mt} - \eta \mu_k = 0,$$

leading to

$$\sum_{m \in T_\mu^{-1}(k)} D_{T_D(m)} \left\{ \left(\sum_t x_t p_{mt} \right) - B_{T_B(m)} \left(\sum_t z_t p_{mt} \right) \right\} = \sum_{m \in T_\mu^{-1}(k)} D_{T_D(m)} \sum_t \mu_k p_{mt} + \eta \mu_k,$$

or that

$$\mu_k = \left(\sum_{m \in T_\mu^{-1}(k)} D_{T_D(m)} \left(\sum_t p_{mt} \right) + \eta I \right)^{-1} \left\{ \sum_{m \in T_\mu^{-1}(k)} D_{T_D(m)} \left[\left(\sum_t x_t p_{mt} \right) - B_{T_B(m)} \left(\sum_t z_t p_{mt} \right) \right] \right\}.$$

When there is no mean sharing (i.e., the mean is used in only one component), we get that $T_\mu^{-1}(k)$ has only one entry, and if also $\eta = 0$, the inverse covariance matrix cancels out, and we end up with a greatly simplified form for the means:

$$\mu_k = \frac{\sum_t x_t p_{mt} - B_{T_B(m)} \sum_t z_t p_{mt}}{\sum_t p_{mt}}.$$

4.4 When Only the Variances are Tied

In certain cases, it may be that only a particular pattern of tying occurs. For example, when only the variances are tied, a number of simplifications are possible. The auxiliary $\mathcal{Q}$ function becomes:

$$\mathcal{Q} = \sum_m \sum_t \left\{ \frac{1}{2} \log(|D_{T_D(m)}|) - \frac{1}{2} (x_t - B_m z_t - \mu_m)^T D_{T_D(m)} (x_t - B_m z_t - \mu_m) \right\} p_{mt} - f(B, \mu),$$

where $f(B, \mu)$ indicates only the regularization terms corresponding to $\{B_k : 1 \le k \le K_B\}$ and $\{\mu_k : 1 \le k \le K_\mu\}$. Because B_k and μ in this case are not tied, we can, as in Section 3, join them

together in a single B matrix with the assumption that the final z has a constant of 1, and if we assume that they utilize the same regularization coefficient, we get:

$$\mathcal{Q} = \sum_m \sum_t \left\{ \frac{1}{2} \log(|D_{T_D(m)}|) - \frac{1}{2}(x_t - B_m z_t)^T \right.$$
$$\left. D_{T_D(m)}(x_t - B_m z_t) \right\} p_{mt} - \frac{\lambda}{2} \sum_m ||B_m||_F^2.$$

As in Section 3, we set $\frac{\partial \mathcal{Q}}{\partial B_m} = 0$ to get that:

$$\sum_t D_{T_D(m)}(x_t - B_m z_t) z_t^T p_{mt} - \lambda B_m = 0,$$

or once again that:

$$\sum_t (x_t - B_m z_t) z_t^T p_{mt} - \lambda D_{T_D(m)}^{-1} B_m = 0,$$

leading to the same update equations for B_m.

To find the update equations for the tied D_k matrix, we set $\frac{\partial \mathcal{Q}}{\partial D_k} = 0$ to get:

$$D_k = \frac{\sum_t \sum_{m \in T_D^{-1}(k)} (x_t - B_m z_t)(x_t - B_m z_t)^T p_{mt}}{\sum_t \sum_{m \in T_D^{-1}(k)} p_{mt}} \tag{17}$$

$$= \frac{\sum_{m \in T_D^{-1}(k)} \left(\sum_t p_{mt} x_t x_t^T - B_m \sum_t p_{mt} z_t x_t^T \right)}{\sum_{m \in T_D^{-1}(k)} \sum_t p_{mt}}, \tag{18}$$

where again the second equality holds only when $\lambda = 0$. This equation, therefore, as in Section 3 is relatively easy to compute. Also, note that when $\lambda = 0$, this is an EM rather than a GEM procedure (see Section 5), since in each case we are maximizing rather than just increasing the auxiliary function (since there is no parameter dependence). In general, we see that the form of parameter tying and regularization can interact in complicated ways to determine how we can optimize the parameters, and this is discussed further in Section 5.

4.5 Regularized MLLR

Speaker adaptation [41,26,24,47,59] is one of the most successful parameter adjustment methods in speech recognition. In the adaptation paradigm, the training and test set distributions are not identical [61]. Moreover, there is a large training set so that the training set distribution may be well estimated, but only a small amount of adaptation data from the test distribution is available—this small test distribution data set is used to slightly

adjust the parameters of a model to better represent the test set distribution. Of course, if too little restriction is placed on the amount of adjustment made to the model based on the adaptation data, then over-training can occur, and therefore adaptation must be based on a highly constrained model. In the widely used maximum-likelihood linear regression (MLLR) [41,59], a particular form of constraint is used, namely the mean vectors in a Gaussian mixture model are subject to a transformation that consists of no more than a scale, rotation, and shift (an affine transformation).

Given our derivation of shared conditional Gaussians above, we can extend MLLR so that not only the Gaussian means are adapted, but also the covariance matrix, all using non-zero mean Gaussian Bayesian priors—thanks to the Cholesky factorized covariance matrix.

In this framework, we moreover may easily adapt to C the regression coefficients, even if they correspond to a cross-stream conditional Gaussian. For example, the component parameters μ and B can be adapted while leaving the conditional variance parameters D fixed. This form of adaptation is seen essentially as another application of operations similar to the B matrix. In particular, suppose that for each m there is a mapping $T_A(m) : \{1, \ldots, M\} \to \{1, \ldots, K_A\}$ that maps from each component, say m, to a "regression class" [41], which gives an index that identifies the parameters that are to affect that component. The kth adaptation parameters then are to adapt to C all components $\{m : m \in T_A^{-1}(k)\}$ using a similar notation as in Section 4. For each k, there is a general upper triangular (so not symmetric) adaptation coefficient matrix A_k, a shift vector b_k, and a variance scale diagonal matrix C_k. These parameters influence the model as expressed in the following adjusted auxiliary function:

$$\begin{aligned}\mathcal{Q} = \sum_m \sum_t \sum_r \Bigg(& \frac{1}{2} \log(C_{T_A(m)}(r)^2 D_m(r)) \\ & - \frac{1}{2}\left(x_t(r) - A_{T_A(m)}(r) B_m(r) z_t(r) - b_{T_A(m)}(r)\right)^2 C_{T_A(m)}(r)^2 D_m(r) \Bigg) p_{mt} \\ & + \sum_k \left(-\frac{\lambda}{2}||A_k - I||^2 - \frac{\eta}{2}||b_k||^2 - \frac{\nu}{2}||C_k - I||^2 \right),\end{aligned}$$

where $A_k(r)$ (resp. $b_k(r)$, $C_k(r)$) is the rth row (resp. entry, diagonal element) of matrix A (resp. vector b_k, C_k). Again, any original mean parameters are contained within $B_m(r)z_t(r)$. A_k must be upper triangular so that the form A_kB is still upper diagonal with zeros along the diagonal, and this means that the Gaussian

normalization constant (a function of D_m) is still valid. Also, since C_k only occurs in squared form, C_k other than being a diagonal matrix is unconstrained. As can be seen, we adapt parameters of the mth component Gaussian by applying an affine transformation to the conditional parameters as in $A_{T_A(m)}(r)B_m(r)z_t(r) + b_{T_A(m)}(r)$ which becomes the new conditional mean. Also, the rth variance is scaled by $C_k(r)$. Of course, to avoid over-training on the data, we assume that K_A is relatively small, and if $K_A = 1$ we have one global A, C matrix pair and b vector that influences the model by performing one joint shift, scaling, and rotation of the parameters. This generalizes MLLR as well since we are not only allowing the adaptation to occur but we penalize the set of adaptation coefficients depending on its distance from the identity matrix using the Frobenius norm terms $||A_k - I||$ and $||C_k - I||$. This last regression term provides an additional way other than the number of regression classes that accuracy-regularization within MLLR may be controlled. This may be appropriate in cases with extremely small amounts of data, where even one (possibly sparse) A matrix would require too many parameters. Note that the approach outlined above may be seen as applying a Bayesian prior [31] on the adaptation coefficients in MLLR, but this is not the same form as MAP adaptation [26], where a Bayesian prior (in particular, a Dirichlet and Wishart prior) is applied to all of the Gaussian mixture parameters (including the mixture coefficients).

To perform adaptation, then, we can derive an EM update step very similar to the derivation in the shared parameter case in Sections 4.2 and 4.3. First, let $\bar{z}_t^m(r) = B_m(r)z_t(r)$ and $x_t^m(r) = x_t(r) - b_{T_A(m)}(r)$. Let $I(r)$ be a row-vector of zeros but with a one in the rth position (i.e., $I(r)$ is the rth row of the $M \times M$ identity matrix). We find $\frac{\partial \mathcal{Q}}{\partial A_k(r)}$ and $\frac{\partial \mathcal{Q}}{\partial b_k(r)}$ yielding

$$\frac{\partial \mathcal{Q}}{\partial A_k(r)} \left(\sum_{m,t,r} \left(x_t^m(r) - A_{T_A(m)}(r)\bar{z}_t^m\right)^2 D_m(r) p_{mt} - \sum_{k',r'} \frac{\lambda}{2} ||A'_k(r') - I(r')||^2 \right)$$
$$= \sum_t \sum_{m \in T_A^{-1}(k)} D_m(r)(x_t(r) - A_k(r)z_t^m)(z_t^m(r))^T - \lambda(A_k(r) - I(r)).$$

Solving for $A_k(r)$ gives

$$\sum_t \sum_{m \in T_A^{-1}(k)} D_m(r) x_t(r)(z_t^m(r))^T$$
$$- A_k(r) \sum_t \sum_{m \in T_A^{-1}(k)} (D_m(r)I) z_t^m (z_t^m)^T - \lambda A_k(r) + \lambda I(r),$$

yielding

$$A_k(r) = \left(\sum_t \sum_{m \in T_A^{-1}(k)} D_m(r) x_t(r) (z_t^m(r))^T + \lambda I(r) \right) \left(\sum_t \sum_{m \in T_A^{-1}(k)} (D_m(r) I) z_t^m (z_t^m)^T + \lambda I \right)^{-1}.$$

We can thus see that as $\lambda \to \infty$, we have that $A_k(r) \to I(r)$, so that $A_k \to I$, thus increasing λ decreases the amount of adaptation for each regression class. Updates for C_k are also relatively straightforward, where for each $C_k(r)$ we get a quadratic form. Each EM update step can be performed one or more times and, like MLLR [41], sets of EM updates can be iteratively interleaved with decodings to potentially further improve results in the unsupervised adaptation case.

It is easy to see that the above corresponds to placing nonzero-mean Gaussian priors on each of the three components of a Gaussian as we have used throughout this chapter, a property enabled by the use of Cholesky factorization. The Cholesky factorization has been used before [24,22]. Moreover, Bayesian priors have also been used for adaptation. For example, a prior may be placed on all parameters in the system [26]. Also, in maximum a posterior linear regression (MAPLR), a matrix Gaussian prior is placed on coefficients governing the affine transformation of the model means [11,56], or on an inner-factor diagonal of the covariance matrix [12]. A kernel ridge regression form of adaptation was also recently developed [54]. The technique presented above is different in that it does not use zero mean Gaussian priors.

5 • A Generalized EM Procedure

A generalized EM (GEM) is a procedure where rather than maximizing the auxiliary function by adjusting the parameters at each EM iteration, we instead only increase the auxiliary function. The above procedures that have been outlined show that, depending on the regularization coefficients and the parameter tying pattern, we may not be able to form a closed form update of each of the parameters as a function only of the previous EM iterations parameters, something as shown on the left of Figure 2. This circularity in the parameter updates cannot be broken even by using multiple passes through the data, as

none of the new parameters can be formed without access to the other new parameters.

There have been two strategies for parameter update, however, that break the cycles. The first strategy has suggested that we can still update all parameters simultaneously if we use the previous parameters. In other words, when one parameter type is needing another from the current iteration, we substitute its value with the corresponding parameter at the previous EM iteration. This procedure, again, is guaranteed to have the same fixed points as a standard EM iteration, but we have not provided any guarantees regarding whether it increases the likelihood function, nor if it ever reaches such a fixed point.

An alternate strategy is the one suggested on the right of Figure 2. In this case, we only update one parameter set at a time. For example, we might find the optimal values only for $\{D_k : 1 \leq k \leq K_D\}$, use these new parameters in Eq. (13). Next, we might optimize $\{B_k : 1 \leq k \leq K_B\}$, use these new parameters, and then last optimize $\{\mu_k : 1 \leq k \leq K_\mu\}$. This entire procedure can then be repeated. This alternating optimization can be done without needing to worry about cyclic parameter dependencies since none of the parameters of a given type depend on any of the other parameters of that same type in the same EM iteration (e.g., μ_k for any fixed k does not depend on any $\mu_{k'}$ for any k' in the EM current iteration).

We can easily see that this procedure is a generalized EM since each update is guaranteed not to decrease the original likelihood function—in particular, the auxiliary function, Eq. (13) is concave in each of μ_k, B_k, and D_k separately, and in each case we obtain the same standard EM bounds [15,46]. For example, computing the Hessian matrix of Eq. (13) with respect to $B_k(r)$, we get

$$\frac{\partial^2 \mathcal{Q}}{\partial B_k(r)^T \partial B_k(r)} = - \sum_{m \in T_B^{-1}(k)} D_{T_D(m)}(r) \left(\sum_t z_t(r) z_t(r)^T p_{mt} \right) - \lambda I,$$

a matrix that is clearly everywhere negative definite for all $\lambda \geq 0$, implying that $\mathcal{Q}$ is concave everywhere. Hessians with respect to the other parameters are also everywhere negative definite.

6 • Conclusions

We have described the EM/GEM update equations for generic Gaussian mixture systems, which include standard Gaussian mixtures, hidden Markov models, auto-regressive HMMs, multi-stream buried Markov models, and regularized MLLR. More specifically, we have described these equations where the

Gaussians are either regularized, partially parameter shared, or both.

We have seen that depending on the regularization parameters or the parameter tying pattern, the standard EM update equations might possess circular dependences, which requires modifications to the normal EM updates to break data dependency cycles. Several schemes were proposed for parameter update in this case, including when only one of each set of parameters is updated at a time. We have seen how this yields a GEM procedure that is still guaranteed to produce a local-maximum of the likelihood function. It would be interesting to relate the parameter tying mechanisms described here with Gaussian generalizations of Probabilistic Relational Models (PRMs) [20], as parameter sharing is a common feature in such models. In their case, however, parameters are typically discrete so the parameter dependency circularities do not arise. If one were to generalize PRMs to include partial Gaussian sharing as we have done above, the same procedures could potentially be adopted. Moreover, all of the above procedures would benefit from the application of ℓ_1-norm regularization [31], which would encourage sparse structure while retaining convexity in the EM update equations.

The above system of equations for tied and regularized Gaussians has been implemented as part of the graphical models toolkit (GMTK) and has been used in a wide variety of contexts. This material is based upon work supported by the National Science Foundation under Grant No. IIS-0093430 and by the Office of Naval Research under Grant No. N000140510388.

References

1. L.R. Bahl, P.F. Brown, P.V. de Souza, and R.L. Mercer. Maximum mutual information estimation of HMM parameters for speech recognition. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 49–52, Tokyo, Japan, December 1986.
2. L.E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Statist.*, 37(6):1554–1563, 1966.
3. J. Bilmes. Buried markov models: A graphical modeling approach to automatic speech recognition. *Comput. Speech Lang.*, 17:213–231, April–July 2003.
4. J. Bilmes and C. Bartels. Graphical model architectures for speech recognition. *IEEE Signal Process. Mag.*, 22(5):89–100, September 2005.
5. J. A. Bilmes. Graphical models and automatic speech recognition. In R. Rosenfeld, M. Ostendorf, S. Khudanpur, and M. Johnson, editors,

Mathematical Foundations of Speech and Language Processing. Springer-Verlag, New York, 2003.
6. J.A. Bilmes. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. Technical Report TR-97-021, ICSI, 1997.
7. J.A. Bilmes. Buried Markov models for speech recognition. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Process*, Phoenix, AZ, March 1999.
8. J.A. Bilmes. Factored sparse inverse covariance matrices. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, 2000.
9. C. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
10. C.-P. Chen, J. Bilmes, and D. Ellis. Speech feature smoothing for robust ASR. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, March 2005.
11. C. Chesta, O. Siohan, and C.-H. Lee. Maximum a posteriori linear regression for hidden markov model adaptation. In *European Conf. on Speech Communication and Technology (Eurospeech)*, 1999.
12. W. Chou and X. He. Maximum a posteriori linear regression (maplr) variance adaptation for continuous density hmms. In *European Conf. on Speech Communication and Technology (Eurospeech)*, 1513–1516, Geneva, Switzerland, 2003.
13. A. P. Dawid. Conditional independence in statistical theory. *J. R. Stat. Soc. B*, 41(1):1–31, 1989.
14. T. Dean and K. Kanazawa. Probabilistic temporal reasoning. *AAAI*, 524–528, 1988.
15. A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum-likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B.*, 39, 1977.
16. Y. Ephraim, A. Dembo, and L. Rabiner. A minimum discrimination information approach for HMM. *IEEE Trans. Info. Theory*, 35(5):1001–1013, September 1989.
17. Y. Ephraim and L. Rabiner. On the relations between modeling approaches for information sources. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 24–27, 1988.
18. R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, NY, 1980.
19. N. Friedman. The Bayesian structural EM algorithm. *14th Conf. on Uncertainty in Artificial Intelligence*, 1998.
20. N. Friedman, L. Getoor, D. Koller, and A. Pfeffer. Learning probabilistic relational models. In *IJCAI*, 1300–1309, 1999.
21. Z. Gajic. *Lyapunov Matrix Equation in System Stability and Control*. Academic Press, San Diego, 1995.
22. M.J.F. Gales. Maximum likelihood linear transformations for hmm-based speech recognition. *Comput. Speech Lang.*, 12:75–98, 1998.
23. M.J.F. Gales. Semi-tied covariance matrices for hidden Markov models. *IEEE Trans. Speech Audio Process.*, 7(3):272–281, May 1999.

24. M.J.F. Gales and P.C. Woodland. Mean and variance adaptation within the mllr framework. *Comput. Speech Lang.*, 10, 1996.
25. J.D. Gardiner, A.J. Laub, J.J. Amato, and C.B. Moler. Solution of the sylvester matrix equation $axbt + cxdt = e$. *ACM Trans. Math. Softw.*, 18(2):223–231, 1992.
26. J.L. Gauvain and C.H. Lee. Maximum a-posteriori estimation for multivariate Gaussian mixture observations of markov chains. *IEEE Trans. Speech Audio Process.*, 2:291–298, 1994.
27. G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins, Baltimore, 1996.
28. R. A. Gopinath. Maximum likelihood modeling with Gaussian distributions for classification. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 1998.
29. D.A. Harville. *Matrix Algebra from a Statistician's Perspective*. Springer, New York, 1997.
30. T. Hastie, S. Rosset, R. Tibshirani, and J. Zhu. The entire regularization path for the support vector machine. *J. Mach. Learn. Res.*, 5:1391–1415, October 2004.
31. T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. Springer, New York, 2001.
32. D. Heckerman, D. Geiger, and D.M. Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. Technical Report MSR-TR-94-09, Microsoft, 1994.
33. D. Heckerman and C. Meek. Embedded bayesian network classifiers. Technical Report MSR-TR-97-06, Microsoft Research, Redmond, WA, 1997.
34. D. Heckerman and C. Meek. Models and selection criteria for regression and classification. In *Proc. Thirteenth Conf. on Uncertainty in Artificial Intelligence*, Providence, RI. Morgan Kaufmann, August 1997.
35. H. Hermansky. Perceptual linear predictive (PLP) analysis of speech. *J. Acoust. Soc. Am.*, 87(4):1738–1752, April 1990.
36. M.J. Hunt, M. Lennig, and P. Mermelstein. Experiments in syllable-based recognition of continuous speech. *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 1980.
37. F.V. Jensen. *An Introduction to Bayesian Networks*. Springer, New York, 1996.
38. B.-H. Juang, W. Chou, and C.-H. Lee. Minimum classification error rate methods for speech recognition. *IEEE Trans. on Speech and Audio Signal Processing*, 5(3):257–265, May 1997.
39. B.-H. Juang and S. Katagiri. Discriminative learning for minimum error classification. *IEEE Trans. on Signal Process.*, 40(12): 3043–3054, December 1992.
40. S.L. Lauritzen. *Graphical Models*. Oxford Science Publications, Oxford, 1996.
41. C.J. Leggetter and P.C. Woodland. Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models. *Comput. Speech Lang.*, 9:171–185, 1995.

42. X. Li and J. Bilmes. Regularized adaptation of discriminative classifiers. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 2006.
43. X. Li, J. Bilmes, and J. Malkin. Maximum margin learning and adaptation of mlp classifers. In *European Conf. on Speech Communication and Technology (Eurospeech)*, 2005.
44. R.J. Mammone, X. Zhang, and R.P. Ramachandran. Robust speaker recognition. *IEEE Signal Process. Mag.*, 13(5):58–71, September 1996.
45. K. Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD thesis, U.C. Berkeley, Dept. of EECS, CS Division, 2002.
46. R.M. Neal and G.E. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. In M.I. Jordan, editor, *Learning in Graphical Models*, 355–368. Kluwer Academic Publishers, Dordrecht, 1998.
47. L. Neumeyer, A. Sankar, and V. Digalakis. A comparative study of speaker adaptation techniques. In *European Conf. on Speech Communication and Technology (Eurospeech)*, 1127–1130, Madrid, Spain, 1995.
48. A. Ng and M. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Neural Information Processing Systems (NIPS)*, 14, Vancouver, Canada, December 2002.
49. P. Olsen and R. Gopinath. Modeling inverse covariance matrices by basis expansion. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 945–948, 2002.
50. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, Los Altos, 2nd printing edition, 1988.
51. J. Pearl. *Causality*. Cambridge University Press, Cambridge, 2000.
52. L.R. Rabiner and B.H. Juang. An introduction to hidden Markov models. *IEEE ASSP Mag.*, 1986.
53. S. Roweis and Z. Ghahramani. A unifying review of linear Gaussian models. *Neural Comput.*, 11:305–345, 1999.
54. G. Saon. A non-linear speaker adaptation technique using kernel ridge regression. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 2006.
55. F. Sha and L. K. Saul. Large margin Gaussian mixture modeling for phonetic classification and recognition. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, Tolouse, France, 2006.
56. O. Siohan, T. Myrvol, and C. Lee. Structural maximum a posteriori linear regression for fast hmm adaptation. *Comput. Speech Lang.*, 16:5–24, 2002.
57. V. Vapnik. *Statistical Learning Theory*. Wiley, New York, 1998.
58. G. Wahba. *Spline Models for Observational Data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, 1990.
59. P.C. Woodland. Speaker adaptation: Techniques and challenges. In *Proc. IEEE ASRU*, 1999.

60. P.C. Woodland and D. Povey. Large scale discriminative training for speech recognition. In *ICSA ITRW ASR2000*, 2000.
61. S. Young. A review of large-vocabulary continuous-speech recognition. *IEEE Signal Process. Mag.*, 13(5):45–56, September 1996.
62. S. Young, J. Jansen, J. Odell, D. Ollason, and P. Woodland. *The HTK Book*. Entropic Labs and Cambridge University, 2.1 edition, 1990's.

30

Speech Synthesis

Thierry Dutoit[1] *and Baris Bozkurt*[2]

[1] *Faculte Polytechnique de Mons, Belgium*
[2] *Izmir Institute of Technology (IYTE), Izmir, Turkey*

Text-to-speech (TTS) synthesis is the art of designing talking machines. It is often seen by engineers as an easy task, compared to speech recognition.[1] It is true, indeed, that it is easier to create a bad, first trial text-to-speech (TTS) system than to design a rudimentary speech recognizer. After all, recording numbers up to 60 and a few words ("it is now", "am", "pm") and being able to play them back in a given order provides the basis of a working talking clock, while trying to recognize such simple words as "yes" or "no" immediately implies some more elaborate signal processing. If speech synthesis were really that simple, however, one could only blame the TTS R&D community for not having been able to massively produce a series of talking consumer products as early as the eighties.

The major point is that users are generally much more tolerant to ASR errors than they are willing to listen to unnatural speech. There is magic in a speech recognizer that transcribes continuous radio speech into text with a word accuracy as low as 50%; in contrast, even a perfectly intelligible speech synthesizer is only

[1] In an international conference on speech processing, a famous scientist once brought up a tube of toothpaste (whose brand was precisely Signal™) and, pressing it in front of the audience, he coined the words: "This is speech synthesis; speech recognition is the art of pushing the toothpaste back into the tube … ".

moderately tolerated by users if it delivers nothing else than "robotic voices".[2]

This importance of *naturalness* versus *meaning* is actually very typical of the synthesis of natural signals (as opposed to their recognition). One could thus advantageously compare speech synthesis to the synthesis of human faces: while it is quite easy to sketch a cartoon-like drawing which will be unanimously recognized as a human face, it is much harder to paint a face which will be mistaken for a photograph of a real human being. To a large extent, you can modify the size, position, and orientation of most of the elements of a portrait drawing without breaking the intelligibility barrier (just think of the cubists ...), but even a slight change to a photorealistic painting will immediately make the complete work look like what it actually is: a painting of a face, not a real photograph of it.

Delivering intelligibility *and* naturalness has thus been the holy grail of speech synthesis research for the past 30 years. This chapter outlines the specific problems, encountered when trying to reach these goals, and shows how existing solutions benefit from fine-tuned signal processing techniques.

Section 1 discusses concatenative synthesis and introduces the underlying signal processing requirements, which are then examined with more details: prosody modification (Section 2), smoothing of discontinuities at concatenation points (Section 3), estimation of concatenation costs for corpus-based speech synthesis (Section 4), and the more advanced voice quality control and speaker modification issues (Section 5).

1 • Concatenative Text-to-Speech Synthesis

Figure 1 introduces the functional diagram of a fairly general TTS synthesizer. It consists of a natural language-processing module (NLP, also termed as *front-end*), capable of producing a symbolic linguistic representation of the text to read: a phonetic transcription, together with the desired intonation and rhythm (often termed as prosody). The NLP module is followed by a digital signal processing module (DSP, also termed as *back-end*), which transforms the symbolic information it receives into speech.

---

[2]Strictly speaking, though, there is no such thing as a "robotic voice" : robots have the voice we give them. In practice, the concept of a "robotic voice" is so strongly established that the TTS researchers are now trying to build "robots that do not sound like robots".

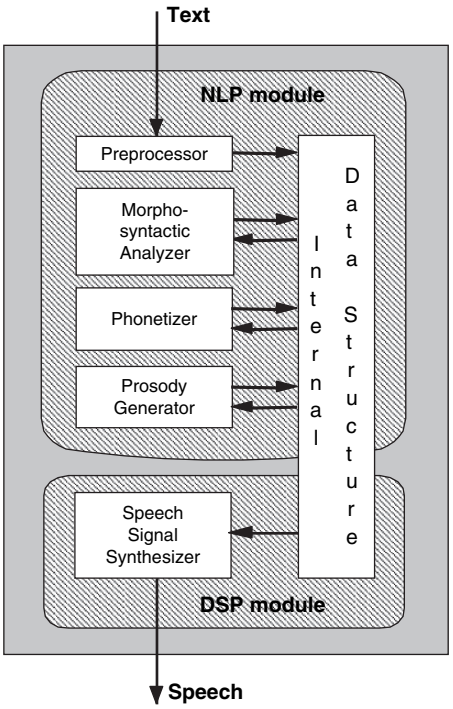


FIGURE 1 The functional diagram of a fairly general text-to-speech conversion system.

A preprocessing (or *text normalization*) module is necessary as a front-end, since TTS systems should in principle be able to read any text, including numbers, abbreviations, acronyms, and idiomatics, in any format. The preprocessor also performs the apparently trivial (but actually intricate) task of finding the end of sentences in the input text. It organizes the input sentences into manageable lists of word-like units and stores them in the internal data structure.[3] The NLP module also includes a morpho-syntactic analyzer, which performs part-of-speech tagging and groups the input sentence into syntactically related sets of words. A phonetizer provides the sequence of phonemes to be pronounced, and a prosody generator computes their duration and intonation (both are *suprasegmental features* and are jointly termed as *prosody*). Once phonemes and prosody have been predicted, the DSP module actually synthesizes them, by

[3]In modern TTS systems, all modules exchange information via some *internal data structure* (most often, a multi-level data structure, in which several parallel descriptions of a sentence are stored with cross-level links; sometimes feature structures as used in unification grammars). More on this can be found in Dutoit (1997, chapter 3).

producing speech samples which, when played via a digital-to-analog converter, will hopefully be understood and, if possible, be mistaken for real, human speech.[4]

Producing speech samples automatically, however, does not merely reduce to the playback of a sequence of pre-recorded words or phonemes. Even though we write in terms of isolated words and think speech in terms of discrete phonemes, we produce *continuous* speech, as a result of the coordinated and continuous action of a number of muscles. These articulatory movements are not produced independently of each other; they are frequently altered in a given context to minimize the effort needed to produce a sequence of articulatory movements. This effect is known as *coarticulation*. Coarticulatory phenomena are due to the fact that each articulator moves continuously from the realization of one phoneme to the next. They appear even in the most carefully uttered speech. In fact, they *are* speech. Thus, producing intelligible speech requires the ability to produce continuous, coarticulated speech.

More generally, transients in speech are more important for intelligibility than stable segments (Harris, 1953; Shannon et al., 1995), while modifying stable segments (e.g., the center of vowels) can very easily affect naturalness. Speech, indeed, is never really periodic, or stable. Even sustained vowels exhibit small frequency and amplitude variations (respectively termed as *jitter* and *shimmer*) and have substantial inharmonic components due to non-complete closure of the vocal folds after the so-called glottal closure instant; the presence of these noise-like components is correlated with the actual pitch period. As a result, the intuitive concept of "adding a bit of noise" to intonation curves, to amplitude curves, or to voiced speech waveforms in order to "make it sound more natural" merely leads to … more noisy voiced speech: inharmonicity in voiced speech is not pure randomness.

Speech signal synthesizers can be roughly classified into two categories: the *expert-based* and the *instance-based* ones.

The former computes speech samples from phonetics either by using acoustic invariants obtained by expert phoneticians (this is often termed as the *rule-based*, or *formant* synthesis approach;

---

[4]In this chapter, we only consider the problems inherent to the design of the speech signal synthesizer in the DSP module; issues related to the NLP module are covered in several general textbooks given in the references, as well as in the Oxford Handbook of Computational Linguistics (Mitkov, 2003). The JASA article by Klatt [87] is still a good reference too, at least for pointing out what problems have to be solved.

see Klatt, 1987; O'Shaughnessy, 1984; Allen et al., 1987) or by describing speech in terms of the movements of the articulators, and perform synthesis using equations derived from mathematical simulations of the vocal tract (this is often termed as the *articulatory* synthesis approach; see Flanagan et al., 1975; Sondhi and Schroeter, 1997; Badin et al., 1998). Expert-based speech synthesis is mostly used for fundamental research on speech, since an improvement of the resulting speech quality necessarily coincides with additional insight into speech and vocal tract dynamics. We do not focus on this approach here.

On the opposite, instance-based (often termed as *concatenative*) speech synthesizers produce synthetic speech by gluing chunks of natural speech (preliminarily recorded and chopped out of human speech) together, in a specific order, so as to deliver a given message. Speech chunks readily embody natural jitter, shimmer, and inharmonicity. They can also be tailored so as to account for coarticulation, by constraining units to contain most coarticulatory effects. As a result, although concatenative synthesis only implicitly refers to the phonetics of speech (it could just as well be used for the synthesis of musical sounds), it is by far the winning paradigm in the industry. As Lindblom (1989) soundly notices: "*After all, planes do not flap their wings!*"

1.1 Diphone-Based Synthesis

Constraining speech chunks to embody coarticulation effects can be achieved to some extent by using *diphones* as basic units. A diphone is a speech segment which starts in the middle of the stable part (if any) of a phoneme and ends in the middle of the stable part of the next phoneme (Dixon and Maxey, 1968). Diphones therefore have the same average duration as phonemes (about 100 ms), but if a language has Nphonemes, it typically has about N^2 diphones,[5] which gives a typical diphone database size of 1,500 diphones (about 3 minutes of speech, i.e., about 5Mb for speech sampled at 16 KHz with two bytes per sample). Some diphone-based synthesizers also include a reduced set of multi-phone units of varying length to better represent highly coarticulated speech (such as in /r/ or /l/ contexts).

In order to create such a synthesizer, one needs to set up a list of required diphones; a corresponding list of words is carefully completed, in such a way that each segment appears at least once (twice is better, for security). Unfavorable positions, like inside-stressed syllables or in strongly reduced (over-coarticulated) contexts, are usually excluded. A corpus is then read by a

[5]A bit less in practice: not all diphones are encountered in natural languages.

professional speaker (avoiding speaking style variations as much as possible, and even possibly with flat pitch, so as to facilitate speech analysis) and digitally recorded and stored. The elected segments are spotted, either manually with the help of signal visualization tools, or automatically thanks to segmentation algorithms, the decisions of which are checked and corrected interactively. Segment waveforms are then collected into a diphone database. This operation is performed only once.

At run-time, once the synthesizer receives some phonetic input (phonemes, phonemes durations, intonation curves) from the NLP module, it sets up a list of required diphones, together with their required duration and fundamental frequency contour (Figure 2). The available diphones would only match these prosodic requests by chance, since they have generally been extracted from words and sentences which may be completely different from the target synthetic sentence. Some *prosody modification* is therefore necessary. How this is achieved will be examined with more details in Section 2.

Additionally, since segments to be concatenated have generally been extracted from different words—that is, in different phonetic contexts—they often do not match in terms of amplitude and/or timbre. Even in the case of stationary vocalic sounds, for instance, a rough sequencing of waveforms typically leads

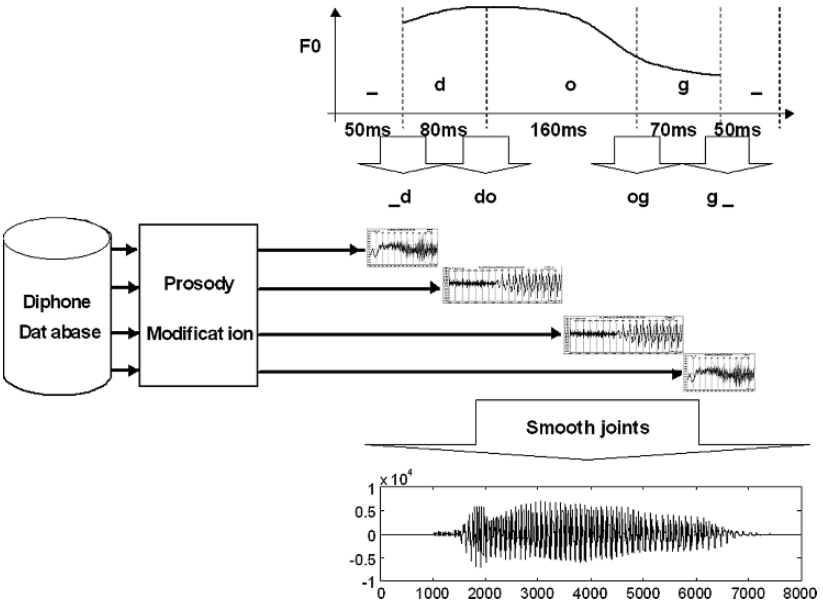


FIGURE 2 A schematic view of a diphone-based speech synthesizer.

to audible discontinuities. Part of this problem can be solved during the construction of the synthesis segments database by *equalization,* in which related endings of diphones are given similar energy levels at the beginning and at the end of segments (by setting the energy of all the phones of a given phoneme to their average value, for instance), in such a way as to eliminate amplitude mismatches. In contrast, timbre (i.e., spectral envelope) conflicts are better tackled at run-time, by *smoothing* (Figure 2) individual couples of segments when necessary rather than equalizing them once for all, so that some of the phonetic variability naturally introduced by co-articulation is still maintained. Smoothing will be discussed in Section 3.

1.2 Unit Selection-Based Synthesis

An extension of the diphone-based approach called *automatic unit selection* (Hunt and Black, 1996) or *corpus-based synthesis* has more recently been introduced, and opened new horizons to speech synthesis. As for diphone-based synthesis, the design of such a system requires the recording of a speech corpus, although much larger: 1–10 h typically, or 150–1,500 Mb. Segmentation in this case can only be achieved semi-automatically, if not completely automatically. The result is a diphone database which comprises much more than one instance of each diphone; all instances differ by their phonetic context and prosodic realization.

At run-time, given a phoneme stream and target prosody for an utterance, this algorithm selects, from this very large speech database, an optimum set of acoustic units (typically isolated diphones[6] or sequences of contiguous diphones, hence the NUU acronym which is sometimes used for this technique: nonuniform units) that best matches the target specifications (Figure 3).

For every *target unit* required (typically, every diphone to be synthesized), the speech database proposes lots of *candidate units*, each in a different context (and in general not exactly in the same context as the target unit). When candidate units cannot be found with the correct pitch and/or duration, prosody modification can be applied. Candidate units usually do not concatenate smoothly (unless a sequence of such candidate units can be found in the speech database, matching the target requirement), some smoothing can be applied. Recent synthesizers, however, tend to avoid prosodic modifications and smoothing, which sometimes create audible artifacts, and keep the speech data as is (Balestri

[6]Although sub-phonetic units such as half-diphones have also been proposed as alternatives when complete diphones are not available with the required pitch and duration (Conkie, 1999).

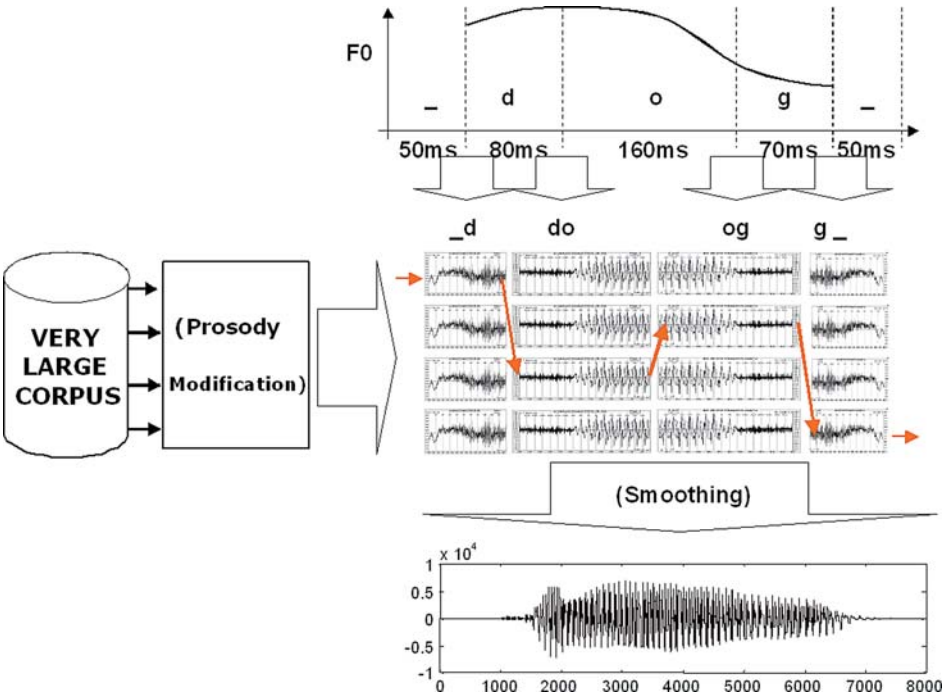


FIGURE 3 A schematic view of a unit selection-based speech synthesizer. The prosody modification and smoothing modules have been mentioned between parentheses, since they are not always implemented.

et al., 1999) (thereby accepting some distortion between the target prosody and the actual prosody produced by the system, and some spectral envelope discontinuities).

The biggest challenge of unit selection synthesis lies in the search for the "best" path in the candidate units lattice. This path must minimize two costs simultaneously: the overall *target cost,* defined as the sum of elementary target costs between each candidate unit chosen and the initial target unit, and the overall *concatenation cost*, defined as the sum of elementary concatenation costs between successive candidate units. The elementary target cost is typically computed as a weighted sum of differences between the linguistic features of the target and candidate units. The elementary concatenation cost is usually computed as a weighted sum of acoustic differences between the end of the left-candidate unit and the beginning of the right-candidate unit. This will be examined in Section 4.

The remarkable naturalness of speech produced by unit selection-based synthesizers comes from the fact that they tend to avoid as many concatenation points as possible, by selecting the largest sequences of consecutive units in the database. This greatly decreases the overall concatenation cost: the elementary concatenation cost between successive candidate units in the database is simply set to zero, so as to force the system to choose consecutive units. Another great advantage of such systems is that they avoid a precise (and often unnatural) prediction of phoneme durations and intonation curves. These are obtained as a by-product of unit selection itself provided the correct linguistic features are used for the estimation of the target cost. Most TTS systems sold today are based on this technique, while more rudimentary diphone-based systems are still used for low cost consumer products (as well as for research: free diphone-based TTS systems are available[7]; this is hardly the case for unit-selection-based systems).

It should be noted, however, that the development of larger and larger segmented speech corpora cost a lot. Hence, it is becoming increasingly difficult for companies to deliver tailored voices on demand. Similarly, the voice quality offered by today's TTS systems is somehow uniform. Producing expressive speech can only be achieved with the parallel recording of several speech corpora, each with an identifiable voice quality; this adds to the cost of a system. An alternative to the development of various speech corpora is precisely the use of speech processing techniques for voice quality modification, or for speaker modification. This problem is the focus of Section 5.

2 • Prosody Modification

The term *prosody* refers to certain properties of the speech signal which are related to audible changes in pitch, loudness, and syllable length. As we have seen, the concatenation of acoustic units taken from natural speech avoids the difficult problem of modeling the way humans generate speech. It also introduces another problem: the possible need for a modification of their prosody, mainly of their intonation and duration.[8]

---

[7]See for instance the synthesizers available from the MBROLA (http://tcts.fpms.ac.be/synthesis/mbrola.html), FESTIVAL (http://www.cstr.ed.ac.uk/projects/festival/), FestVox (http://festvox.org/), or FreeTTS (http://freetts.sourceforge.net/docs/) projects.

[8]In addition to the possible need for smoothing at concatenation points (see Section 3) and for compressing of the unit inventory using a speech

Prosody modification is a requirement for diphone-based synthesizers, in which only one instance of each diphone is available and must be used in all cases (i.e., with various target durations and intonation curves). It tends to be omitted in unit selection-based synthesizers, assuming the speech unit inventory is rich enough for the requested diphone to be available with approximately correct pitch and duration. Prosody modification is also required for *prosody transplantation*, which aims at modifying speech so as to keep the voice of the original speaker (the original speaker is still identifiable) while imposing prosody from another speaker. Prosody transplantation is necessary in industrial applications where synthetic speech must be inserted into fixed, pre-recorded sentences. In order to avoid abrupt voice changes, prerecorded speech is re-synthesized using the synthesizer's voice, but keeping its original, highly natural intonation.

Of the two parameters (pitch and duration), pitch is clearly the more sensitive. Even slightly modifying the shape of a pitch curve indeed readily breaks the naturalness of the speech segment it applies to. This effect is still more audible when the pitch modification ratios on both sides of concatenated units differ too much. Applying a constant pitch modification ratio on a complete segment (as in Bozkurt et al., 2004) does less harm, provided the ratio is not too high (typ. $<= 2$) or too low (typ. $>= 0.5$). Similarly, applying duration modification is less sensitive, provided the ratio is reasonable ($<= 2$ and $>= 0.5$).

Neither pitch nor duration modification is a straightforward operation. Simply resampling speech data (creating additional speech samples by interpolation of original ones, while keeping the same sampling frequency for playing them) does modify duration, and also the spectral envelope. All formants get compressed/dilated on the frequency axis, which has exactly the same audible effect as that of changing the "pitch" wheel of a tape recorder: human voices turn into doggy or mousy voices.

Two main classes of algorithms are in use for prosody modification, namely *time-domain* and *frequency-domain* algorithms.

2.1 Time-Domain Prosody Modification

The pitch-synchronous overlap-add (PSOLA) algorithms (Charpentier and Stella, 1986; Moulines and Charpentier, 1990) have drawn considerable attention, given their high segmental and suprasegmental efficiency, associated with a virtually unequaled simplicity in the case of their time-domain version:

---

coding technique, concatenative synthesis techniques tend to require large amounts of memory. Speech coding is covered in Chapter 31.

TD-PSOLA (Hamon et al., 1989). They extend the early time-domain harmonic scaling (TDHS) of Malah (1979) and the synchronized overlap-add (SOLA) technique of Roucos and Wilgus (1985), which highlighted the efficiency of simple block operations for time scaling of speech. The distinctive feature of the TD-PSOLA method is the idea that it is possible to perform both the required pitch and duration modifications directly on continuous waveforms.

2.1.1 Pitch Modification

If we restrict ourselves to a purely periodic signal $s(n)$, it is possible to get a perfect pitch-modified version of $s(n)$ by summing *OLA frames* $s_i(n)$, extracting pitch synchronously from $s(n)$ with a weighting window $w(n)$, and changing the time-shift between frames from the original pitch period T_0 to the desired period T:

$$s_i(n) = s(n)w(n - iT_0) \tag{1}$$

$$\tilde{s}(n) = \sum_{i=-\infty}^{\infty} s_i(n - i(T - T_0)). \tag{2}$$

If $T \neq T_0$, the operation results, according to the Poisson formula (Papoulis, 1962), in a re-harmonization of the spectrum of $s_i(n)$ (which, if we assume perfect periodicity, does not depend on i when the original voiced sound is purely periodic) with the fundamental frequency $1/T$:

$$\text{if } s_i(n) \overset{F}{\longleftrightarrow} S_i(\omega) \text{ then } \tilde{s}(n) \overset{F}{\longleftrightarrow} \frac{2\pi}{T} \sum_{i=-\infty}^{\infty} S_i\left(i\frac{2\pi}{T}\right) \delta(w - i\frac{2\pi}{T}). \tag{3}$$

Consequently, provided that $w(n)$ can be chosen so that the spectrum of $s_i(n)$ closely matches the spectral envelope of $s(n)$, equation (2) provides a simple and efficient way to change the pitch of a periodic signal (Figure 4).

Notice $w(n)$ also needs to be smoothly zero-ended, so as to avoid clicks at each OLA operation. Typically, Hanning windows are used. Another practical constraint is that F_0 modification produces better audio quality when the size of OLA frames is about two times the local pitch period, and when they are centered on the glottal closure instant, or at least on an instant of important speech excitation (as it is the case on Figure 4). As a result, TD-PSOLA requires the prior computation (or the semi-automatic setting) of *pitch marks* on all speech units. However, computing glottis closure instants (the only reference instant available in voiced speech) has two major

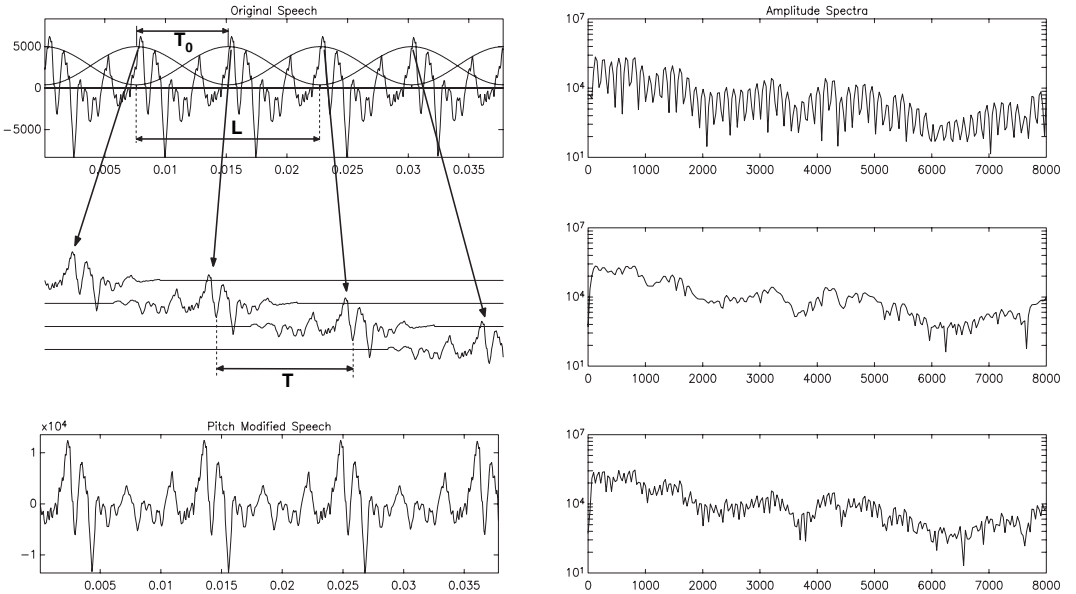


FIGURE 4 The TD-PSOLA re-harmonization process. The pitch-modified waveform (*bottom* plot) has the same spectral envelope as the original waveform (*top* plot). The center plot shows the OLA frames and their Fourier Transform, which is approximately the spectral envelope of the initial signal.

drawbacks: finding glottal closure instants requires expert assistance and it is not very precise if performed automatically (see Cheng and O'Shaughnessy, 1986, for a review), and time consuming if performed manually. A possible alternative to this problem is to measure glottal activity directly, with the help of an electro-laryngograph (Hess and Indefrey, 1987), instead of trying to deduce it from speech. Another solution is to use the local center of gravity of speech as the center of OLA frames (Stylianou, 1998b).

2.1.2 Duration Modification

Duration modification can be achieved in the time-domain by replicating or eliminating some OLA frames, as shown in Figure 5. For duration-only modification, OLA frames need not be pitch-synchronous (see the WSOLA method of Verhelst and Roelands, 1993, and the PICOLA algorithm of Morita and Itakura, 1986, implemented in the new ISO MPEG-4 audio standard). When two noncontiguous frames are overlap-added, a cross-correlation analysis is performed to find the optimal lag or lead between the frame centerings of the source waveforms.

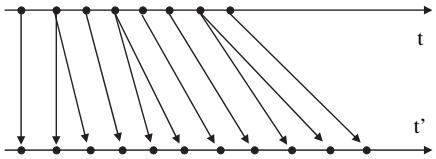


FIGURE 5 OLA-based duration modification process (lengthening in this case, as some frames are duplicated for synthesis). Dots on the top and bottom time lines correspond to centers of frames on the original and modified signals, respectively.

For simultaneous duration and pitch modification, pitch-synchronous frames are used. In this case, in order to prevent the algorithm from introducing artificial short-term autocorrelation in the synthesis signal when frames are duplicated, Charpentier and Moulines (1989) have proposed reversing the time axis of every other duplication of an unvoiced OLA frame. With such an artifice, speech can reasonably be slowed down by a factor of four, even though some tonal effect is encountered in voiced fricatives, which typically combine voiced and unvoiced frequency regions and therefore cannot be inaudibly time reversed.

TD-PSOLA is widely used in concatenative synthesis because of its high synthesis quality and extraordinarily low computational load (typically seven operations per sample; real-time synthesis can be achieved with an Intel 386 processor). Synthesized speech, however, is not perfect: spectral mismatches cannot easily be eliminated at segmental boundaries, and tonal quality appears when large prosodic modifications are applied on the concatenated units.

2.2 Frequency-Domain Prosody Modification

Sinusoidal approaches (e.g., Macon, 1996) and hybrid harmonic/stochastic representations (as Stylianou's HNM, 1998a) have also been proposed for speech synthesis. HNM (Harmonic+ Noise Model), in particular, is increasingly used because of its additional smoothing (see Section 3.1) and speech compression capabilities. In this model, the speech spectrum is divided into two bands. A time-varying maximum-voiced frequency determines the limit between the two bands. In the lower band, the signal is represented as a sum of harmonically related sinewaves with slowly varying amplitudes and frequencies. The frequencies, amplitudes, and phases of these sinewaves are estimated from real speech signal by minimizing a weighted time-domain least-squares criterion. The upper band, which contains the noise part, is modeled by a linear prediction (LP) model and it is modulated by a time-domain amplitude envelope. This makes it possible to account for the nonperiodic components of speech,

which include frication noise and period-to-period variations of the glottal excitation.

3 • Smoothing

Concatenating speech segments possibly extracted from different words or phonetic contexts is not straightforward. If done without care, it produces audible clicks, which cannot always be removed by simple OLA between the end of the left segment and the beginning of the right segment. This is due to three types of mismatches between successive segments, respectively related to phase, pitch, and spectral envelope mismatches (Dutoit and Leich, 1993), and illustrated in Figure 6 for the concatenation of vowels [a] sampled at 16 kHz.

Phase mismatch originates in overlap-adding frames that are not centered at the same place within the period (Figure 6a): even if the wave shapes to be concatenated are identical, the

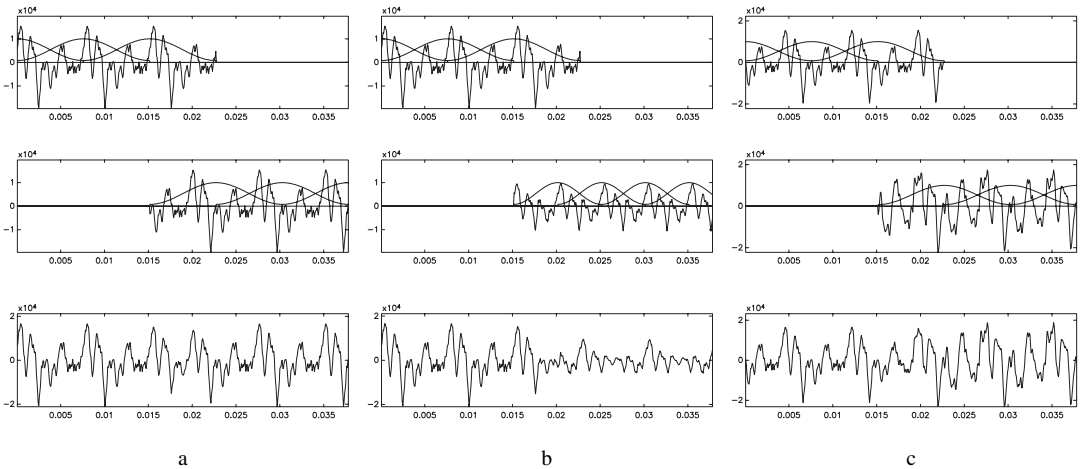


FIGURE 6 *Top* = last frames of the first segment to be concatenated; *center* = first frames of the second one; *bottom* = after OLA; (**a**) *Phase mismatch*: Waveforms are identical but the OLA windows are not centered on the same relative positions within the period. (**b**) *Pitch mismatch*: Both segments have exactly the same spectral envelope but were pronounced with different pitches. OLA windows are positioned coherently. The synthetic signal is obtained after changing the interval between the frames of the right segment, in order to impose a constant pitch. (**c**) *Spectral envelope mismatch*: The left segment is the diphone "ma", the right one is "am". The pitch is constant, and the windows are again positioned coherently. The spectral discontinuity is concentrated in one period.

OLA operation results in an unexpected discontinuity in this case. This can be avoided by adequate pitch-marking of boundary frames (as done in TD-PSOLA for all speech segments). Another solution is to adjust the position of OLA windows on the left and right segments so as to maximize the cross-correlation between windowed frames (as in WSOLA), thereby maximizing the chance that overlap-added frames are synchronous (Meyer et al., 1993). This solution, which corresponds to finding a common relative synchronization point, is easier than the absolute synchronization on glottal closure instants, but it is more time-consuming (it is usually performed online, since synchronization points cannot be pre-computed for all possible pairs of speech segments). It is also possible to use Stylianou's (1998b) center of gravity as an absolute reference point for centering frames.

Pitch mismatch arises when overlapped signals have very different fundamental periods, even if they have identical spectral envelopes and are windowed on similar relative positions. This situation cannot easily be avoided in a database of several thousand segments (Figure 6b). It is nevertheless possible to minimize this factor by recruiting professional speakers to record the segments database and training them to read the corpus with the most constant pitch possible. Another solution is to apply frequency-domain pitch modification before smoothing.

Last but not least, spectral envelope mismatches are due to coarticulation and speaker variability and appear whenever speech segments to be concatenated have been extracted from rather different phonetic contexts (Figure 6c). They can be minimized, by using context-dependent segment boundaries (Conkie and Isard, 1994). It is also possible to smooth them out, by using frequency/parametric domain (3.1) or time-domain (3.2) methods. A recent trend in unit selection-based synthesis is simply to ignore them.

3.1 Frequency/ Parametric Domain Smoothing

Speech models can advantageously be used for smoothing discontinuities around concatenation points, provided a parametric form for acoustic units which makes it possible to modify their local spectral envelope without introducing audible artifacts. There has been a considerable amount of research effort directed at the design of adequate speech models for TTS. Linear prediction (LP) has been used first (Markel and Gray, 1976) for that purpose, given its relative simplicity. However, the buzziness inherent in LP degrades perceived voice quality.

The *linear prediction PSOLA* (LP-PSOLA) approach (Moulines and Charpentier, 1988) was introduced as a means of combining the high compression and smoothing capabilities of the LP

model with the natural prosodic modifications provided by TD-PSOLA. LP-PSOLA techniques thus associate a parametric LP synthesizer with a TD-PSOLA algorithm. Prosodic modifications are performed with the TD-PSOLA algorithm on the excitation signal of the all-pole synthesis filter rather than on direct speech. The output of the overlap-adder is simply connected to an LPC synthesis filter, on the coefficients of which linear smoothing rules are applied. Several LP-PSOLA techniques can be derived from this principle, depending on the type of excitation signal used, as in the *multipulse linear prediction coding PSOLA* algorithm (MPLPC/PSOLA) of Moulines and Charpentier (1988), or the *residue-excited-PSOLA* (RELP-PSOLA) technique (Macchi et al., 1993).

The *HNM* model (see Section 2.2) also allows to smooth out discontinuities by linearly interpolating harmonic amplitudes and LP coefficients. Special care has to be taken for phase interpolation, due to phase wrapping (Stylianou et al., 1997).

3.2 Time-Domain Smoothing

The standard implementation of TD-PSOLA cannot smooth spectral envelope mismatches. As a result, synthesizing highly fluent speech with TD-PSOLA often requires a painstaking trial-and-error process in which the most probable diphone sequences are systematically tested, and diphones are rejected when an important spectral envelope mismatch is encountered. Other versions of the rejected diphones are then extracted from words with a different phonetic context than the initial one; tests are run again, and so on.

A solution consists of resynthesizing the speech segment database so as to provide segments with interesting properties which can be turned into account at synthesis time. Resynthesis is performed once on the database, off-line. This is the case of the so-called *multiband resynthesis OLA* (*MBROLA*) technique (Dutoit and Leich, 1993), in which a sinusoidal model is used to resynthesize all periods with the same pitch and with the same initial harmonic phases. This maximizes the similarity between periods having the same spectral envelope, and therefore happens to provide an efficient solution to pitch and phase mismatches as well. Additionally, when periods don't have similar spectral envelope (i.e., in case of spectral envelope mismatch), it is possible to perform period-based smoothing in the time domain (Figure 7). This feature compensates the loss of segmental quality due to the resynthesis operation. It is also one of the features of MBROLA which contributed to the success of the MBROLA project (http://tcts.fpms.ac.be/synthesis/mbrola/), by making it possible for nonspeech experts to contribute to the development of new synthetic voices.

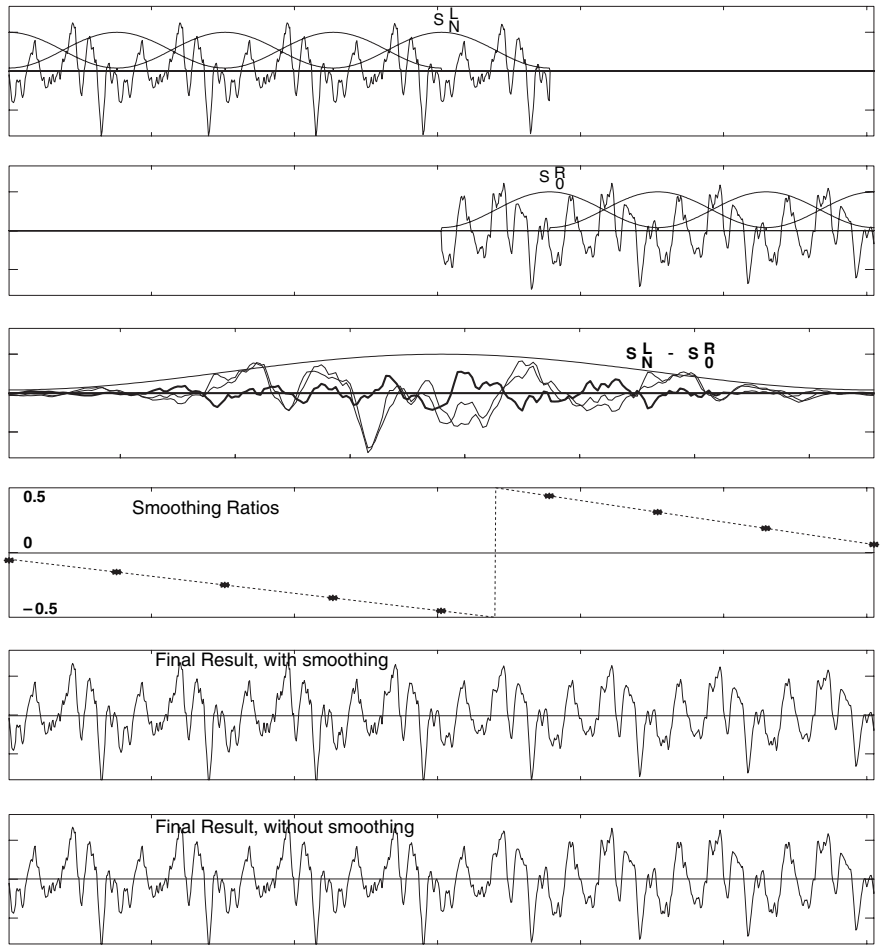


FIGURE 7 The time-domain linear smoothing process. S_N^{L} is the last frame of the left speech segment; S_0^{R} is the first frame of the right segment. The difference between waveforms, $S_N^{\mathrm{L}} - S_0^{\mathrm{R}}$, is distributed on several frames around the concatenation point.

For a formal comparison between different smoothing algorithms, see Dutoit (1994) and Syrdal et al. (1998).

4 • Estimation of Unit Selection Costs

The quality of speech produced by corpus-based concatenative synthesis is very sensitive to the definition of *target* and *concatenation* costs (Figure 8). These costs, as we shall see, should not be minimized separately, but rather through a weighted

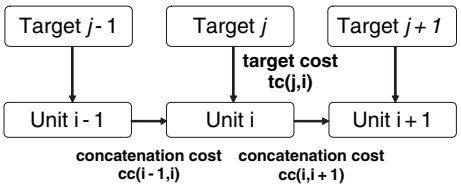


FIGURE 8 Target and concatenation costs.

sum. Target costs should favor the selection of units whose acoustics will match the context in which they will be inserted. Ideally, target costs should thus reflect the acoustic (or even perceptual) difference between target units and candidate units. Since the acoustics of the target are not available (the synthesizer is precisely in charge of producing them), only linguistic features can be used for estimating the segmental difference. We will not focus here on the definition of target costs, as they rely on linguistic descriptions of speech which fall beyond the scope of this chapter. The reader can refer to Möbius (2000).

In contrast, ideal concatenation costs should reflect the perceptual discontinuity (as opposed to the acoustic one) between successive candidate units, so as to favor the selection of units whose "cut-and-paste" concatenation will not be perceived. In particular, if the complete target sentence, or a significant part of it, is available in the corpus, all its consecutive units should obviously be selected. This is easily achieved by setting concatenation cost to zero for originally consecutive units. As a result, the unit selection algorithm tries to use as many consecutive units as possible from the corpus, provided the related target cost is not excessive (which would correspond to a "fluid" pronunciation of the required phonemes, but with the wrong intonation or phrasing, and therefore lead to low intelligibility).

Since Hunt and Black (1996), many proposals have been made for best defining concatenation costs. Costs are systematically based on some parameterization (using some speech model) of the initial and final speech frames of concatenated units, combined to some distance between the values of these parameters. Among the parameters used for this purpose, the most prominent are formant values, LPC spectra, Line Spectrum Frequencies (LSFs), Mel-Frequency Cepstrum Coefficients (MFCCs), various other frequency-modified spectral representations. Distances between the values of these multidimensional parameters vary from the Euclidian distance to the Kullback–Leibler divergence, through weighted Euclidian distances or the Mahalanobis distance. Although it is not clear yet which one leads to optimal results in all cases (i.e., for all

recording conditions and all languages, if only possible), several studies have been made for comparing combinations of these options on specific corpora, by computing correlations between the values of these parameters/distances combinations and the impression of discontinuity reported by human listeners (see for instance the comparisons performed by Donovan (2001)). In the work of Klabbers and Veldhuis (2001), it was found that using LPC power-normalized spectra and a symmetrical Kullback–Leibler distance function gave the best results. This is close to Stylianou and Syrdal's (2001) choice for a symmetric Kullback–Leibler distance between FFT-based power spectra. In Wouters and Macon (1998), an Euclidian distance on MFCCs was chosen. Vepa and King (2004) report on using the Mahalanobis distance on LSFs. Clearly, this question is still open; its solution will obviously require (and lead to) a better modeling of perceived discontinuity in speech.

5 • Voice Quality/Speaker Modification and Other Open Issues

Voice quality in speech synthesis research usually refers to the perceived degree of characteristics like breathiness, creakiness, and loudness. Voice quality variations are considered to be mainly due to the variations in the phonation (production of the glottal excitation signal) process at the glottis. Voice quality issues have often been studied within the context of formant synthesis in the speech synthesis area since in this type of parametric approach, glottal excitation signals are synthesized parametrically and therefore can rather easily be controlled. Few studies address voice quality issues in concatenative synthesis (Kawai and Tsuzaki, 2004; Stylianou, 1999; Campbell and Marumoto, 2000). But voice quality analysis/synthesis is drawing more and more attention in the domain of concatenative speech synthesis since it is one of the most important features of naturalness in speech. In addition, it is especially important for emotional/expressive speech synthesis (an area which is gaining popularity) since voice quality codes as much information about the state of the speaker as does the prosodic information (Campbell and Marumoto, 2000). There is a strong correlation between voice quality variations and prosodic variations in speech since both are features of the phonation process. Therefore, advances in one of the fields would potentially result in advancement in the other. However, current state of art in voice quality analysis/modification/synthesis of recorded speech is not

yet advanced enough to be widely used since tools for automatically estimating and modifying features of the glottal excitation of real speech are necessary; this is a challenging problem due to nonlinear processes in speech production.

From a signal processing point of view, voice quality variations mainly correspond to variations in spectral tilt, in the relative amount of aperiodic components in speech, and in some spectral variations in the low frequency part of the spectrum (like variations in the glottal formant frequency F_g in the bandwidth of first formant, in amplitudes of the first few harmonics, etc.) (Doval and d'Alessandro, 2003). The simplest approach consists of studying how these few features vary in natural speech. In Figure 9, we demonstrate (on synthetic speech signals) how variations in two spectral features (glottal formant frequency F_g and high frequency band energy) correspond to a variation of phonation in the tense-soft dimension. The time-domain signals presented at the *bottom* of Figure 9 include the glottal excitation (glottal flow derivative) signal and the speech signal obtained by filtering this glottal excitation signal with an all-pole vocal tract filter with resonances at 600, 1,200Hz, 2,200, and 3,200Hz. A variation from tense phonation to soft phonation corresponds to a decrease both in glottal formant frequency and high frequency band energy (and vice versa). For a detailed study of acoustic feature variations due to voice quality variations, see Klatt and Klatt (1990).

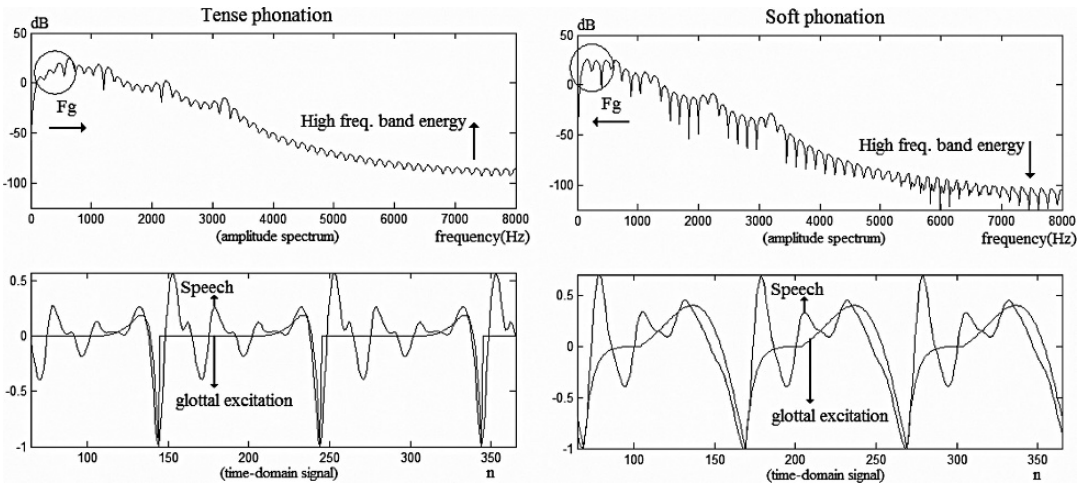


FIGURE 9 Spectral variations due to variations in phonation. *Top* figures show the amplitude spectrum and *bottom* figures show the time domain signals for glottal excitation and speech (Bozkurt, 2005).

5.1 Voice Quality and Concatenative Synthesis

Voice quality studies in concatenative speech synthesis research are concentrated mainly on voice quality labeling of speech corpora. In corpus construction for concatenative synthesis, one of the main difficulties is the need for long recording sessions and the physical difficulties this brings; voice quality changes during sessions (for example, due to the voice fatigue effect at the end of sessions). When long sessions are split into smaller sessions to avoid fatigue, this leads to the problem of matching voice qualities among sessions. Voice quality variations may be induced by many other reasons which are related to physical properties of the vocal folds or even psychological factors (when the speaker gets bored, he/she may start speaking fast-tense to be able to finish the recording quickly). Such variations in speech corpora potentially result in an increased degree and frequency of acoustic discontinuities in synthetic speech. This problem has been addressed in some recent studies (Kawai and Tsuzaki, 2004; Stylianou, 1999; Campbell and Marumoto, 2000) and is an open research area.

Most often, voice quality variations are treated as very slow variations (for example, variations between two recording sessions of several hours) but actually, voice quality differences do not exist only between segments from two different phrases. Voice source characteristics may show quite fast variations and voice quality discontinuities may exist even when two segments of the same sentence are concatenated. One of the reasons why voice quality discontinuities are not seen as a major concern for the current TTS systems is that unit selection is based on phonological structure matching and therefore implicitly selects segments with substantial voice quality continuity. However, voice quality and prosody are not completely interdependent and the realization of various voice quality effects for the same prosodic pattern and phonological segment is possible, and even desirable for expressive speech synthesis.

There are mainly two ways to avoid the discontinuities introduced by voice quality mismatches: using offline signal processing to equalize some voice quality variations, or including voice quality features in concatenation costs for unit selection to guarantee voice quality continuity.

The advantage of offline processing for voice quality equalization is that it does not introduce additional complexity in unit selection or to run-time signal processing. Stylianou (1999) proposes a method to detect some voice quality problems like roughness and breathiness throughout long recording sessions and correct them by using auto-regressive filters. Given a large speech corpus from the same speaker, this technique automat-

ically selects a recording session as reference, sets statistical decision thresholds by using Gaussian mixture models, and corrects speech segments where needed.

The second solution stays rather as an unstudied subject apart from studies investigating which acoustic features are correlated with voice quality variations. One such study is Kawai and Tsuzaki (2004) which investigates the correlation of various measures (MFCC, Spectral tilt, band limited power) to perception scores of voice quality discontinuities constructed by concatenating phrases recorded over a long period of time (up to 202 days) and concludes that band-limited power (8-16Khz) is a good measure for detecting voice quality discontinuities. This area is likely to be studied in the concatenative expressive speech synthesis area in the upcoming years.

One other possible solution to the voice quality discontinuity problem is voice quality smoothing on the fly, which is another unstudied subject. For such an operation, some means of achieving voice quality modification without audible degradation of in the segmental quality of speech are clearly needed. Most of the concatenative synthesis algorithms, which include some spectral smoothing, smooth some of the voice quality discontinuities automatically (for example, spectral envelope smoothing results in smoothing some spectral tilt discontinuity automatically). However, such smoothing is too local for removing all voice quality variations. Higher quality smoothing can be possible once the separate components of discontinuity are understood (for example, the discontinuity in the aperiodic component of speech).

5.2 Voice Quality Modification

Voice quality modification is not only useful for smoothing voice quality variations but also for synthesizing variations on purpose. A potentially very useful and interesting approach in expressive speech synthesis is to use a hybrid synthesis paradigm: concatenative synthesis with more control parameters (like energy, spectral tilt, relative energy of the aperiodic component, glottal formant frequency) than the conventional concatenative synthesizers (which mostly include only duration and intonation information). Such an approach would stretch the limits of what is possible with the available data while keeping naturalness at a higher level compared to formant synthesis.

Natural voice quality modification is one of the unreached goals of speech synthesis research. Defining high-level rules that will drive low-level parameter modifications is too complex with our current understanding of the phenomenon. Signal processing for low-level parameter modification is being addressed in some few studies.

As in many parameter modification schemes, there are mainly two classes of approaches: spectral modification techniques and parametric methods performing some form of decomposition–modification–recomposition. The latter is usually known to introduce some audible artifacts and spectral modifications are preferred when possible. Spectral techniques are specially advantageous for voice quality modification since the perceptual relevance of spectral parameters is high, the representations are rather simplified (various time-domain models can be unified into a single spectral representation, as presented in Doval and d'Alessandro, 1997). They are also less sensitive to phase distortions, while time-domain voice quality modification applied on phase distorted speech signals often results in very low quality speech. D'Alessandro and Doval (2003) draw guidelines for the spectral modification of voice quality in a recent study. Their method includes modification on three dimensions: glottal formant, spectral tilt and PAPR (periodic to aperiodic ratio) the last two being perceptually more relevant for modification. However, the implementation of their ideas in a speech synthesizer is still not tested and is potentially a promising research area.

5.3 Speaker Modification

For commercial speech synthesis systems, customers often ask for the possibility of modifying speaker characteristics on the fly. Voice quality modification tools are potentially very useful for modifying some characteristics of the available voices (for example, making a synthetic voice sound softer to increase pleasantness, or just the converse: making it tenser to draw more attention).

Creating a "new" speaker from an existing one is rather more complex, although some signal processing tricks like changing average pitch frequency and applying spectral envelope modification would provide some speaker variations which are acceptable in some special cases (see Laroche, 2003, for instance).

For concatenation-based systems, creating a completely new voice to the system requires the recording and segmentation of a large speech corpus for the new speaker. An alternative solution is to use a voice conversion algorithm to be able to create a large speech database for a new speaker, given a large speech database from another speaker and a small speech database from the new speaker. Most of the algorithms proposed for this voice conversion task performs frame-by-frame spectral mapping. Long-time suprasegmental variations like intonation, voice quality, and timing are not realistically accounted for. Despite those difficulties, using voice conversion technology for

adding new voices to concatenative speech synthesis systems deserves to be listed among future technical goals for speech synthesis research.

5.4 Other Open Issues

The interest of signal processing in speech synthesis by far does not reduce to the topics mentioned in the previous sections. Among other interesting implications, we would like to point out some which deserve special attention:

- *Monitoring tools for speech recordings for TTS database construction*. Efficiency in long recording sessions is essential because database construction is an expensive process. Some of the factors reducing efficiency are variations in general characteristics of the speaker's voice (pitch, loudness, and speech rate), low-quality segments produced due to fatigue or other reasons like dryness of vocal folds. Tools that can monitor pitch, loudness, speech rate, jitter–shimmer, breathiness, etc. and provide warning messages when some limits are exceeded can be potentially useful. We have found no reference to such tools in the literature.
- *Perceptual models for speech features*. Except the considerable effort made on finding the optimum perceptual distance measure for spectral discontinuities for improving quality of unit selection systems, perceptual speech models are still not much used. For example, pitch discontinuities, maybe the most important type of discontinuities in concatenative synthesis but no known measure of perceived pitch discontinuity seems to be available. In the best of all worlds, speech synthesis technology should be equipped with perceptual models and perform smoothing operations in a perceptual domain. Such an approach would perform better smoothing when necessary and reduce the amount of unnecessary smoothing operations which sometimes introduce degradations.
- *High-quality fundamental frequency modification*. F_0 modification still stays as one of important problems in concatenative speech synthesis. This is due to the fact that perceptual relevancy of fundamental frequency is very high, and existing modification algorithms have limited quality. Improvements in the quality of fundamental frequency modification have a great potential for reducing the size and complexity of speech corpora.
- *HMM-based speech synthesis* (Tokuda et al., 2004). Speech recognition and speech synthesis research can be considered to be somehow two independent research areas when we consider the amount of shared tools. HMM-based speech synthesis is a step forward for profiting from existing advances in speech

recognition technology in speech synthesis. One example is the investigation of using speaker adaptation techniques of speech recognition for voice conversion of speech synthesis in such a hybrid system.

- *Advanced articulatory modeling/synthesis*. Although the recent advances in concatenative TTS have resulted in a reduction of popularity and dedicated effort on articulatory speech models, there is still a need for a better insight in speech production phenomena in order to expand our reach in TTS synthesis and speech processing in general. Articulatory speech synthesis should be stated among the future technologies providing plenty of interesting problems to solve and demanding considerable research effort.

For a complementary list of forthcoming speech synthesis research areas, the reader could advantageously refer to Sproat et al.'s (1999) report on the *NSF Workshop for Discussing Research Priorities and Evaluation Strategies in Speech Synthesis*.

6 • Further Reading and Relevant Resources

Most of the information presented in this chapter can be found with a lot more details in Dutoit (1997), Van Santen et al. (1997) or Sproat (1998).

References

Allen, J., Hunnicut, S. and Klatt, D. (1987). *From Text to Speech: The MITTALK System*, Cambridge University Press.

Badin, P., Bailly, G., Raybaudi, M. and Segebarth, C. (1998). "A three-dimensional linear articulatory model based on MRI data". *Proceedings of the International Conference on Speech and Language Processing*, vol. 2, pp. 417–420, Sydney, Australia, November 1998.

Balestri, M., Paechiotti, A., Quazza, S., Salza, P. and L., Sandri, S. (1999). "Choose the best to modify the least: a new generation concatenative synthesis system", Proceedings of Eurospeech, Budapest, Hungary.

Bozkurt, B. (2005). *Zeros of the Z Transform (ZZT) Representation and Chirp Group Delay Processing for the Analysis of Source and Filter Characteristics of Speech Signals*, PhD Dissertation, Faculté Polytechnique de Mons, Belgium.

Bozkurt, B., Dutoit, T., Prudon, R., d'Alessandro, C. and Pagel, V. (2004). "Improving quality of MBROLA synthesis for non-uniform units synthesis". In Narayanan, S. and Alwan, A. (eds.), *Text to Speech Synthesis: New Paradigms and Advances*. Prentice Hall.

Campbell, N. and Marumoto, T. (2000). "Automatic labelling of voice-quality in speech databases for synthesis.", *Proceedings of the International Conference on Spoken Language Processing (ICSLP'00)*, pp. 468–471, Beijing, China.

Charpentier, F. and Stella, M.G. (1986). "Diphone synthesis using an overlap-add technique for speech waveforms concatenation", *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 86*, pp. 2015–2018.

Charpentier, F. and Moulines, E. (1989), "Pitch-synchronous waveform processing fechniques for text-to-speech synthesis using diphones", *Proceedings of Eurospeech 89*, Paris, vol. 2, pp. 13–19.

Cheng, Y.M. and O'Shaughnessy, D. (1986). "Automatic and reliable estimation of glottal closure instant and period", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no 12, pp. 1805–1815.

Conkie, A. (1999). "A robust unit selection system for speech synthesis", *Proceedings of the 137th meet. ASA/Forum Acusticum*, Berlin.

Conkie, A. and Isard, S. (1994). "optimal coupling of diphones", *Proceedings of the 2nd ESCA/IEEE Workshop on Speech Synthesis,* Mohonk, Sept. 1994.

D'Alessandro, C. and Doval, B. (2003). "Voice quality modification for emotional speech synthesis." *Proceedings of the European Speech Communication and Technology (Eurospeech'03)*, pp. 1653–1656. Geneva, Switzerland.

Dixon, N.R. and Maxey, H.D. (1968). "Terminal analog synthesis of continuous speech using the diphone method of segment assembly", *IEEE Transactions on ASSP*, AU-16, no. 1, pp. 40–50.

Donovan, R.E. (2001). "A new distance measure for costing spectral discontinuities in concatenative speech synthesisers". *Proceedings of 4th ISCA Tutorial and Research Workshop on Speech Synthesis*, Perthshire, Scotland.

Doval, B. and d'Alessandro, C. (1997). "Spectral correlates of glottal waveform models: an analytic study". *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, pp. 1295–1298.

Dutoit, T. (1994). "High quality text-to-speech synthesis : a comparison of four candidate algorithms". *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'94)*, pp. 565–568. Adelaide, Australia.

Dutoit, T. (1997). *An Introduction to Text-To-Speech Synthesis*, Kluwer Academic Publishers.

Dutoit, T. and Leich, H. (1993). "MBR-PSOLA: text-to-speech synthesis based on an MBE resynthesis of the segments database", *Speech Communication*, 13, pp. 435–440.

Flanagan, J.L., Ishizaka K. and Shipley K.L. (1975). "Synthesis of speech from a dynamic model of the vocal cords and vocal tract". *Bell System Technical Journal,* 54, pp. 485–506.

Hamon, C., Moulines, E. and Charpentier F., (1989). "A diphone system based on time-domain prosodic modifications of speech", *Proceedings*

of the International Conference on Acoustics, Speech, and Signal Processing 89, S5.7, pp. 238–241.

Harris, C.M. (1953). "A study of the building blocks in speech", *Journal of the Acoustical Society of America,* 25, pp. 962–969.

Hess, W. and Indefrey, H. (1987). "Accurate time-domain pitch determination of speech signals by means of a laryngograph", *Speech Communication*, 6, pp. 55–68.

Hunt, A.J. and Black A.W. (1996). "Unit selection in a concatenative speech synthesis system using a large speech database". *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'96)*, vol. 1, pp. 373–376. Atlanta, Georgia.

Kawai H. and Tsuzaki M. (2004). "Voice quality variation in a long-term recording of a single speaker speech corpus". In Narayanan, S. and Alwan, A. (eds.), *Text to Speech Synthesis: New Paradigms and Advances.* Prentice Hall.

Klabbers, E. and Veldhuis, R., (2001). "Reducing audible spectral discontinuities", *IEEE Transactions on Speech and Audio Processing*, 9(1):39–51.

Klatt, D.H. (1987). "Text-to-speech conversion", *Journal of the Acoustical Society of America*, 82(3), 737–793.

Klatt D.H. and Klatt L.C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers. " *Journal of the Acoustical Society of America*, 87(2):820–57.

Laroche, J. (2003). "Frequency-domain techniques for high-quality voice modification", *Proceedings of the International Conference on Digital Audio Effects (DAFx-03),* London.

Lindblom, B.E.F. (1989), "Phonetic Invariance and the Adaptive Nature of Speech", in B.A.G. Elsendoorn and H. Bouma eds., *Working Models of Human Perception*, Academi Press, New York, pp. 139–173.

Macchi, M., Altom, M.J., Kahn, D., Singhal, S. and Spiegel, M., (1993). "Intelligibility as a 6th Function of speech coding method for template-based speech synthesis", *Proceedings of Eurospeech 93*, Berlin, pp. 893–896.

Macon, M.W. (1996). *"Speech Synthesis Based on Sinusoidal Modeling"*, Ph. D Dissertation, Georgia Institute of Technology.

Malah, D. (1979). "Time-domain algorithms for harmonic bandwidth reduction and time-scaling of pitch signals", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no 2, pp. 121–133.

Markel, J.D. and Gray A.H. (1976). *Linear Prediction of Speech*, Springer.

Meyer, P., Rühl, H.W., Krüger, R., Kluger, M. Vogten, L.L.M., Dirksen, A. and Belhoula, K., (1993). "PHRITTS – A text-to-speech synthesizer for the German language", *Proceedings of Eurospeech 93*, Berlin, pp. 877–890.

Mitkov, R. (2003). Handbook of Computational Linguistics, R. Mitkov, ed., Oxford University Press.

Möbius, B. (2000). "Corpus-based speech synthesis: methods and challenges". *Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (Univ. Stuttgart),* AIMS 6(4), 87–116.

Morita, N. and Itakura, F. (1986). "Time-scale modification algorithm for speech by use of pointer interval control overlap and add (PICOLA)

and its evaluation," *Proceedings Of Annual Meeting of Acoustical Society of Japan*.

Moulines, E. and Charpentier, F. (1988). "Diphone synthesis using a multipulse LPC technique", *Proceedings of the FASE International Conference*, Edinburgh, pp. 47–51.

Moulines, E. and Charpentier F. (1990). "Pitch synchronous waveform processing techniques for text-to-speech synthesis using diphones". *Speech Communication*, 9, 5–6.

O'Shaughnessy, D. (1984). "Design of a real-time French text-to-speech system", *Speech Communication*, 3, 233–243.

Papoulis, A. (1962). *The Fourier Integral and Its Applications*, McGraw Hill, p. 47.

Roucos, S. and Wilgus, A. (1985). "High-quality time scale modification of speech", *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 85,* pp. 236–239.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J. and Ekelid, M. (1995). "Speech recognition with primarily temporal cues", *Science*, 13;270(5234):303–4.

Sondhi, M.M. and Schroeter, J. (1997). "Speech production models and their digital implementations", *The Digital Signal Processing Handbook*. CRC and IEEE Press.

Stylianou, Y. (1998a). "Concatenative speech synthesis using a Harmonic plus Noise Model". *Proceedings of the 3rd ESCA Speech Synthesis Workshop*, 261–266. Jenolan Caves, Australia.

Stylianou, Y. (1998b), "Removing phase mismatches in concatenative speech synthesis", *Proceedings of the 3rd ESCA Speech Synthesis Workshop*, pp. 267–272.

Stylianou, Y. (1999). "Assessment and correction of voice quality variabilities in large speech databases for concatenative speech synthesis." *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'99)*, 377–380. Phoenix, A2.

Stylianou, Y. and Syrdal, A.K. (2001). "Perceptual and objective detection of discontinuities in concatenative speech synthesis", Proceedings of ICASSP, Salt Lake City, UT.

Stylianou, Y., Dutoit, T. and Schroeter, J. (1997). "Diphone concatenation using a Harmonic plus Noise Model of speech", *Proceeding of Eurospeech '97,* pp. 613–616.

Sproat, R., ed. (1998). *Multilingual Text-to-Speech Synthesis*. Kluwer Academic Publishers.

Sproat, R., Ostendorf, M. and A. Hunt, eds. (1999). *The Need for Increased Speech Synthesis Research: Report of the 1998 NSF Workshop for Discussing Research Priorities and Evaluation Strategies in Speech Synthesis*.

Syrdal, A., Stylianou, Y., Garisson, L., Conkie A. and Schroeter J. (1998). "TD-PSOLA versus Harmonic plus Noise Model in diphone based speech synthesis". *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'98)*, 273–276. Seattle, WA.

Tokuda, K., Zen, H. and Black, A. (2004). "An HMM-based approach to multilingual speech synthesis." In Narayanan, S. and Alwan, A. (eds.), *Text to Speech Synthesis: New Paradigms and Advances*. Prentice Hall.

Van Santen, J.P.H., Sproat, R., Olive, J., Hirshberg, J. eds. (1997), *Progress in Speech Synthesis*, Springer.

Vepa, J. and King, S. (2004). "Join cost for unit selection speech synthesis". In Alwan A. and Narayanan S. (eds.), *Speech Synthesis*. Prentice Hall.

Verhelst, W. and Roelands, M. (1993). "An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech", *Proceedings of ICASSP-93*, Vol. II, pp. 554–557.

Wouters, J. and Macon, M. (1998). "Perceptual evaluation of distance measures for concatenative speech synthesis", *Proceedings of ICSLP*, Vol 6, pp. 2747–2750, Denver Co.

31

Speech Coders

Roch Lefebvre and Philippe Gournay

Université de Sherbrooke, Sherbrooke, QC, Canada

1 • Introduction

Telecommunication systems make intensive use of speech coders. In wireless systems, where bandwidth is limited, speech coders provide one of the enabling technologies to reach more users and furnish better services. In wireline systems, where bandwidth can be less of an issue, speech is also digitized and compressed to a certain extent depending on the system.

In general, a speech encoder takes sampled speech at its input and produces a compressed bitstream for transmission or storage. At the receiver (or at playback from a storage device), a speech decoder takes the bitstream and produces a synthesized speech signal that is desired to sound as close to the original speech as possible. This is illustrated in Figure 1. The compression ratio of the speech encoder and the quality of the synthesized speech depend on the system requirements and application needs. Very often, the application requirement is that synthesized speech should have quality as high as possible with a bit rate as low as possible. Any voice application with limited channel capacity and/or memory space for storage can greatly benefit from existing low-rate speech coders.

The main objective of this chapter is to give the reader an appreciation of the ways signal processing is used inside and around modern speech coders to meet specific application needs and to solve acoustical problems in the communication link. Examples will focus mostly on standards since these are well documented and thoroughly tested. Only the main concepts of speech coders

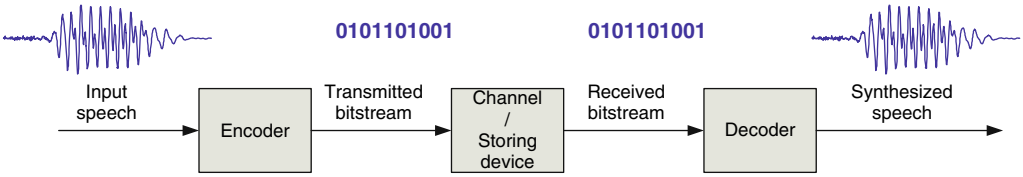


FIGURE 1 Simplified block diagram of a voice transmission or storage system using a speech coder.

will be presented here. For in-depth reviews of speech coders, the reader is referred to the several excellent tutorials written on the subject. For example, see [1,2].

The outline of this chapter is as follows. In Section 2, principles of modern speech coders are briefly reviewed. Section 3 presents methods to address issues related to input conditions. These include the presence of acoustical noise, the effect of audio bandwidth and the variability of the input signal characteristics. In Section 4, problems related to channel conditions, such as packet loss and variable transmission delay, are addressed. Finally, Section 5 discusses current developments and future challenges.

2 • Speech Coding Concepts

Advances in speech coding span several decades. It would be impractical to present here all the advances that have been made in this field. However, looking at most voice communication systems using speech coders today, we observe that the speech coders used in most systems are based on variants of linear predictive (LP) coding [3]. The basic principle in LP coding is to model the speech signal using a time-varying linear filter with a proper excitation signal at the input of the filter. The time-varying linear filter is often called the *synthesis* filter in the sense that it produces the synthetic speech, provided it has an appropriate excitation at its input. Moreover, the excitation signal is very often modeled as the weighted sum of two components: one component selected from an *adaptive codebook* and the other component from the *innovative codebook*. The adaptive codebook is nothing more than the past excitation signal, which is also available at the decoder – or at least approximately if there are channel errors. The adaptive codebook has thus a content which varies in time, as the past excitation is shifted after encoding new speech samples. The innovative codebook is also called the fixed codebook. In the simplest form, its content does not change

with time. The innovative codebook essentially models the unpredictable component of the speech samples, that is, the information that is not modeled by the synthesis filter and the adaptive codebook. What distinguishes many modern speech coders is the structure and the search method of the innovative codebook. Figure 2 shows a simplified version of a modern speech decoder using the above-mentioned principles. We see from this figure that the encoded bitstream carries three types of information: (1) the LP, or synthesis filter parameters; (2) the adaptive codebook parameters (typically, a set of delays and gains); and (3) the innovative codebook parameters (typically codevector indexes and the associated gains). From this information, the decoder can reconstruct or synthesize a frame of speech. At the encoder, the input speech will be split into frames, and the parameters in Figure 2 will thus be calculated and encoded for every frame. A typical speech frame will have between 10 and 30 ms in duration.

In Figure 2, only the speech decoder is illustrated. It shows the relative simplicity of a speech decoder, both conceptually and algorithmically. At every frame, the decoder constructs the total excitation signal and then filters this excitation signal through the synthesis filter to produce a frame of synthesized speech. The complexity is more at the encoder, where the parameters actually have to be calculated from the input speech frame, and then encoded for transmission. This is usually carried out in a sequential manner. First, the synthesis filter parameters are calculated to match the input speech. This involves windowing the speech signal, then calculating a correlation matrix for that

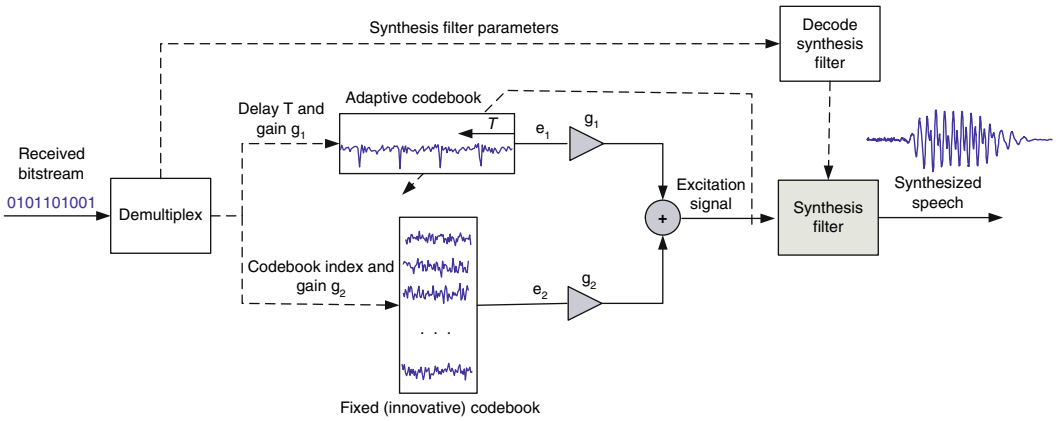


FIGURE 2 Typical speech decoder based on an LP filter model, using two components to encode the excitation: an adaptive codebook and a fixed (innovative) codebook.

speech frame, and inverting the matrix to find the optimal filter parameters. Then, the excitation parameters are calculated in such a way that when the excitation is passed through the LP filter, this produces synthetic speech as close to the input speech as possible according to a given error criterion. This search process is called *analysis-by-synthesis* in the speech coding literature. The adaptive codebook parameters (optimal delays T and gains) are calculated first. This codebook typically models periodic components in speech which are strong when the vocal cords are active. The adaptive codebook thus achieves a significant coding gain with relatively few bits (a delay and a gain). After the adaptive codebook contribution is determined, the innovative codebook is searched for its optimal contribution.

The innovative codebook is shown in Figure 2 as comprising seemingly random components. To make it useful and implementable in real time, this codebook has to be structured. Several approaches have been proposed to build this structure [4–6]. The most commonly used one in speech coders deployed today is referred to as the ACELP model [7]. There are more than 15 international standards today based on ACELP for wireless and wireline communication systems.

The main idea of the ACELP model is to construct the innovative codebook by using an interleaved pulse positions (IPP) structure. Each speech frame is further divided into subframes, typically of 5 ms each. Then, for each subframe, the candidates in the innovative codebook are formed by dividing the samples in the subframe in interleaved *tracks* and by placing few non-zero pulses on each track, each with equal amplitude but individual sign. The number of bits allocated to the ACELP codebook will then depend on the number of tracks and the number of pulses per track. The advantage of this structure is that it requires very little storage (the codebook entries are described by algebraic formulas and not stored in a table). Further more, this pulse structure allows for very fast search algorithms which makes the innovative excitation search implementable in real time even on mobile devices with limited CPU power and battery life. More details on the ACELP structure and search algorithms can be found in [5–7].

In itself, a speech coder introduces perturbations in the speech signal since the encoding process is *lossy*. This means that the decoded, synthesized speech waveform differs from the original speech waveform. These perturbations can however be made perceptually very small. Methods to reduce the perceived distortions in a speech coder include pre- and post-filtering as well as proper weighting of the error signal when selecting the optimal

excitation parameters. Modern, state-of-the-art speech coders can produce high-quality speech at compression ratios of 10:1 and often much more. For example, the EFR-GSM standard narrowband speech coder [8], widely used in cellular networks, operates at 12.2 kbits/s. The compression ratio of this coder is 10.5 (128 kbit/s for the uncompressed narrowband signal, with 16 bits per sample and 8 kHz sampling frequency, divided by 12.2 kbit/s). This compression ratio can easily be doubled with the use of voice activity detectors and discontinuous transmission in the case of silence (more than 50% of conversational speech is silence).

Outside the perturbations introduced by the speech coder itself, several factors can influence the speech quality. In the following sections, we group these factors into two general categories:

- Input conditions
- Channel conditions

We will see how signal processing can be used to address these different issues to maintain or improve the speech quality in different system scenarios.

3 • Issues Related to Input Conditions

In Figure 1, we implicitly assume that the input signal is "clean" speech, which is a suitable requirement for the good performance of speech coders based on the model of Figure 1. This is especially important for low-rate speech coders, which take advantage of the speech signal characteristics. We also assume that the input signal is indeed speech and not, for example, loud acoustical noise or even music, as can be the case in telephony applications when a caller is put on hold. The input conditions have significant influence on the synthesized speech quality. In this section, we consider different variants of input conditions and discuss how signal processing can be used to solve the related issues.

3.1 Signal Dynamics

Digitized speech exhibits specific dynamics in the time domain and in the frequency domain. Figure 3 shows about 3/4 of a second of a typical speech signal. Large variations in the time envelope can be observed. In segments with slowly evolving time envelope, the signal typically exhibits local stationarity, for example, as a pseudo-periodic (voiced) signal or as a pseudo-random (unvoiced) signal. This is shown in Figure 4, where the voiced and unvoiced segments of Figure 3 are shown in more detail in both time and frequency domains.

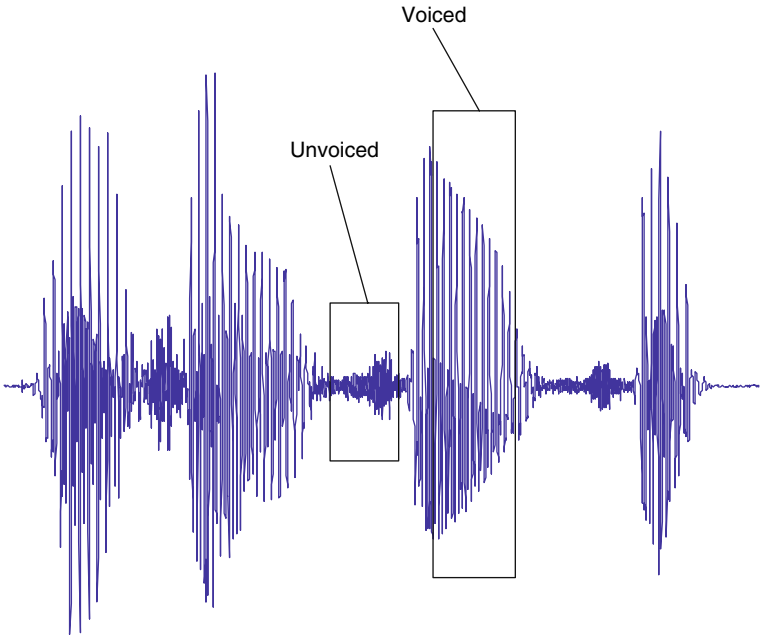


FIGURE 3 Example of a typical speech signal (about 3/4 of a second is shown), with succession of voiced and unvoiced segments.

The speech example in Figures 3 and 4 is sampled at 8 kHz, typical in telephony. The audio bandwidth is thus about 4 kHz (full band). This full-band, non-equalized condition is typically referred to as "flat" speech. Several speech coding algorithms have been optimized for "IRS-filtered" speech. IRS filtering refers to pre-conditioning, or equalizing, the speech signal with a frequency response typical of a telephone handset [9]. Linear-phase FIR filters are typically used to implement the IRS characteristics. Figure 5 shows the gain of the frequency response for the ITU-T IRS8 filter used for narrowband speech signals. Frequencies outside the 300–3,400 Hz telephone band are strongly attenuated, while the telephone band is tilted toward the high frequencies. Low-rate speech coders which have been optimized for IRS-filtered speech can suffer slight quality degradation when encoding flat speech. This is because the quantizers of the speech encoder's parameters do not have enough bits to cover all possible signal characteristics and are thus specialized. Of particular concern is the encoding of the synthesis filter coefficients. It is thus recommended to use a specific speech coding algorithm in the condition for which it was optimized, by properly pre-conditioning the input speech, for example, with

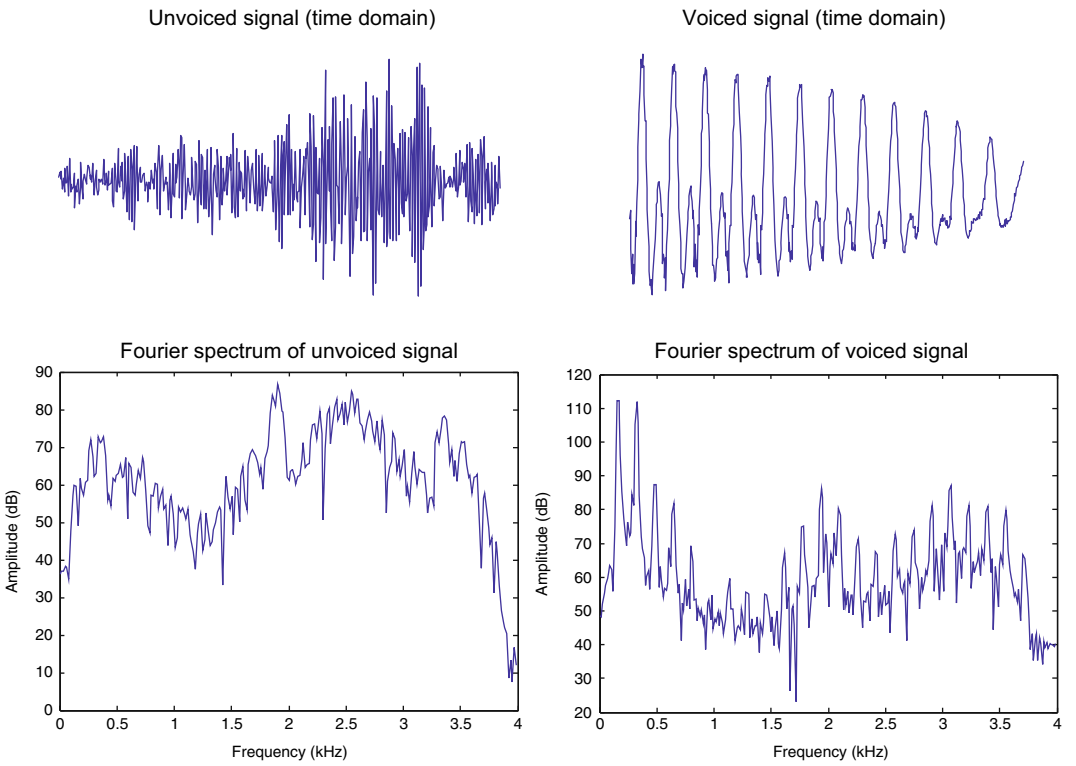


FIGURE 4 Unvoiced and voiced segments of Figure 3, shown in time and frequency domains (time-domain signals not to scale).

an IRS filter characteristic. For example, ITU-T recommendation G.729 was optimized for IRS conditioned signals.

3.2 Acoustical Noise

The input speech of a communication system is rarely exempt from noise. With mobile phones especially, acoustical noise from the speaker's environment is mixed with speech. In extreme cases – talking in a crowded hall for example – the noise level can even surpass the speech level. In most cases, the noise level and characteristics can still perturb the speech signal sufficiently to reduce the performance of the speech coder in the system. Hence, in many systems, different techniques are used to reduce the amount of acoustical noise entering the system. One approach consists of reducing the noise at the transducers, using specific microphones and circuitry. Directional and differential microphones, in particular, can be used to focus on the speech source, which can be seen as unidirectional, while blocking a large part of the acoustical noise, which is generally omnidirectional. With

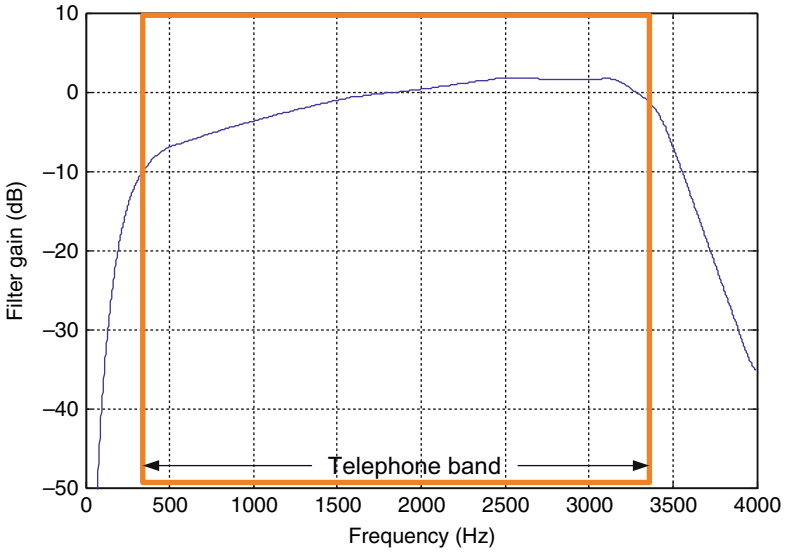


FIGURE 5 Gain of frequency response for IRS filter. Frequencies outside the telephone band are attenuated.

residual noise still present after using directional microphones, noise reduction algorithms can be used to further reduce the level of noise in the speech signal. We consider here the typical case where there is only one microphone, which means that the speech and noise are mixed in a single observation signal.

Several single-microphone noise reduction algorithms have been proposed. A large majority of the ones in use are variants of spectral subtraction methods [10]. The main idea is to subtract an estimate of the noise signal from the "speech-plus-noise" mixture. This subtraction is performed not in the time domain, but in the frequency domain – thus the name spectral subtraction. This is because in the frequency domain, amplitude and phase components can be processed separately. Since in single-microphone systems there is only one observation of the noise and speech mixture, phase information from the noise cannot be extracted in active voice regions – where noise reduction is to be useful. However, if the noise signal is relatively stationary, then an estimate of its amplitude spectrum can be obtained from several successive noise segments where speech is inactive. Hence, a typical spectral subtraction algorithm operates as in Figure 6.

In Figure 6, the noise estimation often relies on a voice activity detector (VAD) to determine whether the current frame has active speech or is only noise. If the current frame is noise only, the noise spectrum is updated. The noise spectrum is hence a smoothed

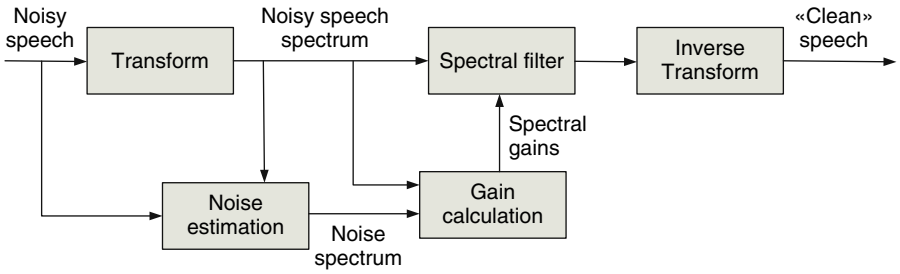


FIGURE 6 Basic blocks of spectral subtraction algorithm for single-microphone noise reduction.

average of the previous frames containing noise only. The update memory length depends on the algorithm. If the current frame is active speech, then it is not included in the noise estimation. Then, in basic spectral subtraction, the noise spectrum (estimate) is subtracted from the noisy speech spectrum (both are amplitude spectra only – phases are left unchanged). The clean speech is recovered by applying the inverse transform. An equivalent, and widely used, variant consists of calculating a gain, between 0 and 1, for each frequency bin of the noisy speech spectrum. When the speech-to-noise ratio is high for a given frequency bin, the gain is set close to 1 (spectrum unchanged). When the speech-to-noise ratio is low, the gain is set closer to 0 (noise attenuation). Using this gain or "spectral filter" approach avoids in particular the problem of negative valued amplitudes as in direct spectral subtraction.

As a specific example of a noise reduction algorithm based on spectral subtraction, consider the VMR-WB wideband speech coder, standardized by the 3GPP2 [11] in 2005. The noise subtraction algorithm of VMR-WB is described in [12].

3.3 Audio Bandwidth

For decades, the audio bandwidth of speech has been limited to roughly 300–3,400 Hz, in particular due to link limitations and handset frequency response. As shown in Figure 7, speech occupies a larger bandwidth compared to the telephone band. Transmitting this larger bandwidth can bring the users much closer to experiencing "face-to-face" quality communications, which strongly benefits applications such as teleconferencing and videoconferencing. The lower band – below 300 Hz – increases presence, while the higher band – above 3,400 Hz – increases intelligibility as several fricatives have significant content in the high frequencies. However, as the spectrum in Figure 7 shows, spectral dynamics and variability increase with bandwidth. In the example given, which is a segment of voiced speech, the telephone

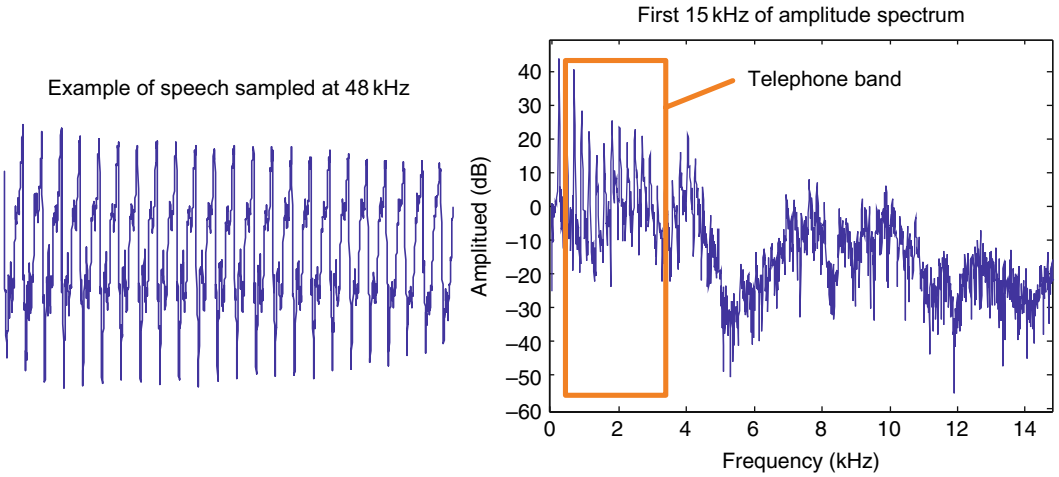


FIGURE 7 Speech example (voiced) sampled at 48 kHz.

band exhibits the harmonic structure discussed above, while the higher band has less structure. This observation suggests that the LP model of Figure 2 is suitable for the low band, but other models may be more applicable to the higher band of speech. Parametric models will be discussed below in the section covering the encoding of non-speech signals.

Depending on the sampling frequency and resulting audio bandwidth, speech can be characterized as

narrowband (or telephone band) for 8 kHz sampling (3 kHz band);
wideband for 16 kHz sampling (7 kHz band);
FM grade for 32 kHz sampling (15 kHz band);
or *full band* for 44.1 or 48 kHz sampling (20 kHz band).

The most significant jump in quality results from using wideband speech compared to narrowband speech. More improvement can be achieved by using a higher bandwidth, but it depends on the quality of the microphones and speakers used in the system. A large fraction of the speech energy lies below 8 kHz. Providing the energy above 8 kHz to the listener requires a proper acoustical environment, including efficient speakers and relatively low environmental noise. Of course, increasing the audio bandwidth usually requires additional bit rate.

Wideband speech can be encoded efficiently using essentially the same coding model as for narrowband speech, namely using the model of Figure 2 at the decoder. There are however some aspects specific to a wider bandwidth. In particular, as shown in Figure 7, the spectrum has larger dynamics (peak to valley

ratios) and cannot be considered fully harmonic even in voiced speech. There are also more formants, i.e., spectral envelope peaks, in wideband speech than in narrowband speech. As a consequence, the synthesis filter in Figure 2 has to model a richer spectral envelope in wideband speech. This is accomplished by increasing the filter order, from a typical value of 10 coefficients in narrowband speech coders to a value between 16 and 20 coefficients in wideband speech coders. Of course, increasing the filter order implies more coefficients to encode and transmit, hence more bits to use for transmitting the synthesis filter. Indeed, a 10th-order synthesis filter typically requires 20–25 bits to transmit, while a 16th-order filter, used in a wideband speech system, will require at least 35 bits.

Increasing the synthesis filter order is not sufficient to increase performance of a wideband speech coder. In the analysis-by-synthesis process performed at the encoder to search for the optimum parameters in Figure 2 in a given speech frame, a weighting filter is used to amplify low-energy bands in the speech spectrum. These low-energy bands are typically the frequencies which lie between the speech formants as well as the high frequencies in the spectrum. Since wideband speech has more spectral dynamics than narrowband speech, the weighting filter has to be optimized for this. Specifically, the weighting filter in a narrowband speech coder typically has the form

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)}, \tag{1}$$

where $1/A(z)$ is the transfer function of the synthesis filter, and γ_1 and γ_2 are adjustable parameters with the constraints $\gamma_1 \leq 1.0$, $\gamma_2 \leq 1.0$, and $\gamma_2 < \gamma_1$. An example is given in Figure 8, using wideband speech (16 kHz sampling) as input. The frequency response of the synthesis filter $1/A(z)$, shown on the left spectrum, follows the envelope of the input speech. Without a weighting filter (i.e., with $W(z) = 1$), the encoder would try to find the best excitation parameters that produce a synthesized frame matching the original speech in the mean squared error sense. In this case, the overall speech encoder would on average introduce white noise in the original speech. Introducing weighting filter $W(z)$ in the excitation parameters search changes things quite a bit. Instead of trying to match the original speech, the encoder now tries to find the best excitation parameters which match a *weighted speech* signal which has the spectral envelope of $W(z)/A(z)$ as shown in dotted line in the left part of Figure 8. It can be easily shown that applying this weighting filter in the excitation parameters search results in shaping the speech coder's

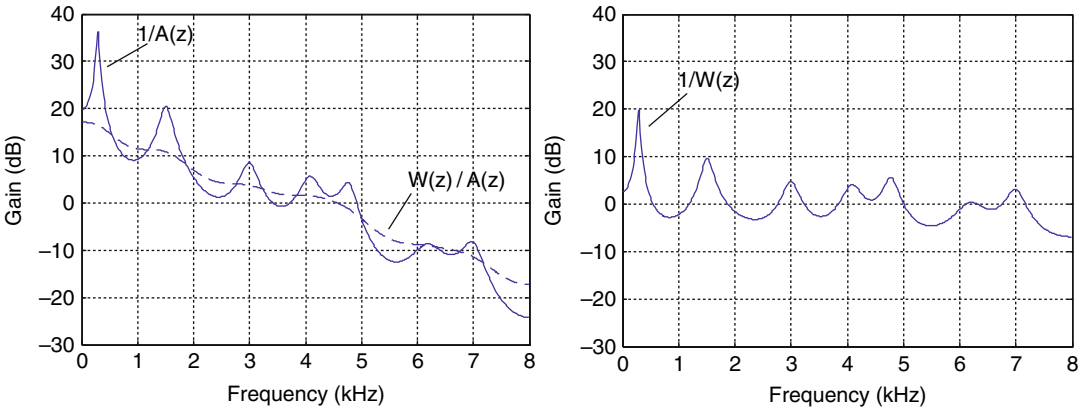


FIGURE 8 Typical example of weighting filter $W(z)$ optimized for a narrowband speech coder. Here, $\gamma_1 = 1.0$ and $\gamma_2 = 0.85$. Left: Frequency response of synthesis filter $1/A(z)$ superimposed with frequency response of *weighted* synthesis filter $W(z)/A(z)$. Right: Frequency response of the inverse weighting filter, $1/W(z)$.

noise with the inverse filter $1/W(z)$, which is shown in the right part of Figure 8. Now, the speech coder's noise, which would otherwise be white, follows more closely the input speech energy distribution – its formants. This reduces the perceived distortion due to the masking properties of the hearing system [13].

Figure 8 shows that this noise weighting strategy works well with narrowband signals, i.e., below 4 kHz. However, for wideband signals having more spectral dynamics, the noise introduced by the speech coder may still be too high in high frequencies due to the large spectral tilt (more than 40 dBs between 500 Hz and 7 kHz in the envelope $1/A(z)$ of Figure 8). Hence, different improvements to the weighting filter have been proposed in the case of wideband speech. A simple modification consists of adding an explicit correction for the spectral tilt in the weighting filter [14]:

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)} T(z), \tag{2}$$

where $T(z)$ is defined as $T(z) = 1 - \alpha \frac{R(1}{R(0)} z^{-1}$. Here, $R(k)$ is the autocorrelation of the input speech frame at lag k and $0 < \alpha < 1$ is a parameter controlling the amount of tilt reduction in $W(z)$. Figure 9 shows the effect of adding the tilt factor $T(z)$ in the weighting filter $W(z)$. Here a value of $\alpha = 0.6$ is chosen, with the same values of γ_1 and γ_2 as in Figure 8. Using this tilt factor, the encoder now has to match a signal that is more "flat" (dotted curved in left part of Figure 9) compared to not using the tilt factor as in Figure 8. The effect is that the noise introduced in the synthesized speech by the encoder will not only follow the formant structure (peaks and valleys) but also exhibit a more proper spectral tilt. The noise in the low-energy spectral bands

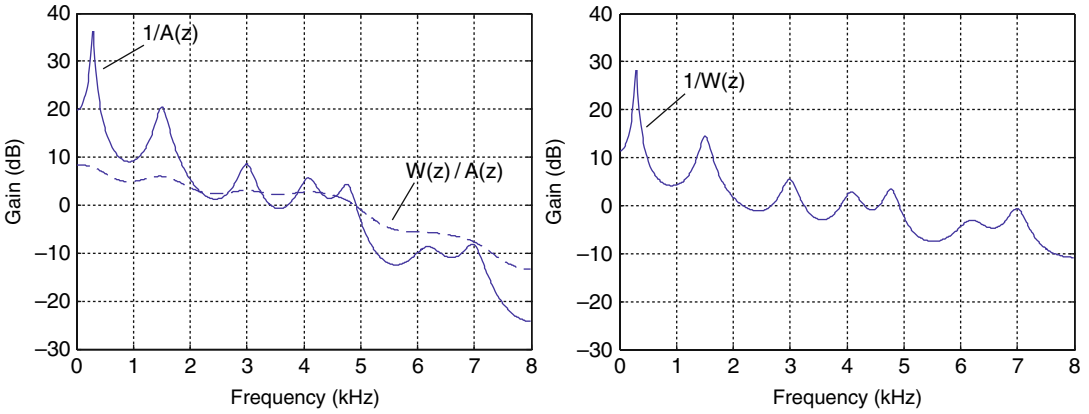


FIGURE 9 Weighting filter with tilt factor as in Equation (2. Here, $\gamma_1 = 1.0$, $\gamma_2 = 0.85$, and $\alpha = 0.6$. *Left*: Frequency response of synthesis filter $1/A(z)$ superimposed with frequency response of *weighted* synthesis filter $W(z)/A(z)$. *Right*: Frequency response of the inverse weighting filter, $1/W(z)$.

(here, in the high frequencies) will therefore be lower and better masked.

A simpler approach to decouple spectral tilt from formant noise shaping is exemplified in the AMR-WB wideband speech coding standard [5]. Here, the input speech is first pre-emphasized using a fixed first-order filter $P(z)$. Then, the synthesis filter $1/A(z)$ is actually computed from the pre-emphasized speech samples. Then, the weighting filter is computed as

$$W(z) = \frac{A(z/\gamma_1)}{P(z)}. \tag{3}$$

This reduces the filtering complexity while achieving essentially the same results as with Equation (2) and noise shaping similar to that in Figure 9.

3.4 Encoding Non-speech Signals

In several applications in communication systems, the audio signal can be different than speech. Examples include call waiting and on-hold music, acoustical noise (covered in a section above), and more recently musical ringtones which give cell phones an audible personality. The growing use of multimedia cell phones also increases the need for low-rate speech and audio coding, where ideally any kind of audio – speech, music, mixed content – can be compressed at low bit rates, transmitted and resynthesized with high quality using a single technology. However, at low rates, speech coders produce annoying artifacts when encoding music. On the other hand, perceptual audio coders [13] cannot be used for speech transmission systems requiring low delay and especially low rates. Perceptual audio coders use more general models than speech coders, but require higher bit rates and delay. In wireless communication systems, in particular, bit rates are critical and must be kept very low.

Seeing the importance of multimedia, and in particular audio, on emerging wireless and cellular systems, the Third Generation Partnership Project (3GPP) initiated in the early 2000s a new standardization activity to seek a new low-rate general audio coding algorithm. This resulted in 2004 in two new recommended standards for low-rate speech and audio transmission on cellular networks. EAAC+ [15], on the one hand, and AMR-WB+ [16] on the other hand were both recommended for complementary applications; EAAC+ for higher rates (32 kbits/s and above) and music signals, and AMR-WB+ for lower rates (24 kbits/s and below) and all types of content from speech to music and including mixed content such as movies soundtracks. The underlying model in these two standards is quite different.

EAAC+ is based on perceptual audio coding, where the input audio is decomposed into frames and each frame is transformed in the frequency domain using a form of Fourier transform [13]. In EAAC+, the Fourier transform used is the modified discrete cosine transform (MDCT). The frequency components are quantized following a bit allocation which depends on a perceptual model estimated from the frequency spectrum of the input signal. The perceptual model essentially calculates a masking curve which determines the maximum allowable noise level per frequency band to keep the noise inaudible. Then, if the bit budget is sufficient for the frame, each frequency component is quantized such that the quantization noise falls below the masking curve. To improve quality at low rates, EAAC+ uses a very low rate parametric approach called spectral band replication (SBR) to encode the high frequencies. Specifically, the signal spectrum is split into two equal and overlapping bands; the lower band is encoded using perceptual audio coding as described above, while the higher band is encoded using SBR. In SBR, the higher band is regenerated at the decoder applying frequency transposition to the lower band and using side information to scale the transposed frequency components. Only a few kbit/s are used as side information, and no bits are used to describe the frequency transposition. Hence, most of the bit rate is used to encode the lower band, where a large fraction of the signal energy usually lies especially in the case of speech signals. A detailed description of the SBR algorithm used in the EAAC+ 3GPP standard can be found in [17]. EAAC+ is a general audio coding algorithm, but it is covered in this chapter since it addresses the problem of encoding speech, music, and mixed signals at low rates.

AMR-WB+, also known as Extended AMR-WB, is another 3GPP standard recommended in 2004 for low-rate general audio transmission on cellular networks [16]. AMR-WB+ uses a hybrid encoding model which alternates between the ACELP model (see Section 2) and a transform-coding model called transform coded eXcitation (TCX) [18]. The TCX model is such that it allows seamless switching with the ACELP model. The ACELP model in AMR-WB+ is actually the AMR-WB speech coding standard from 3GPP, which is a multirate coding standard for speech sampled at 16 kHz (see Section 4.1). The AMR-WB+ encoder selects between the ACELP coding mode (AMR-WB) and the transform-coding mode (TCX) based on the input signal characteristics. It is worthwhile to note that mode selection is not based on speech/music classification; when encoding speech, the TCX model is chosen in a significant proportion, and when encoding music, the ACELP model is also chosen in some frames. Mode selection at the encoder can be made in either closed loop or open loop. In closed-loop mode selection, different coding modes are applied to the input signal, and the best mode is selected based on segmental signal-to-noise measures. In open-loop mode selection, the input signal is classified, and only the selected coding mode is applied based on the classification. Thus, open-loop mode selection results in a reduced-complexity encoder compared to closed-loop mode selection. Different frame lengths can be used and are also determined by the mode selection. As in EAAC+, AMR-WB+ uses a very low rate parametric approach to encode the high frequencies. Instead of frequency transposition as in SBR, AMR-WB+ uses spectral folding from the lower band to synthesize the higher band, along with low-rate side information to properly scale the folded frequency components. Only 800 bits per second per channel are used to encode the high frequencies in AMR-WB+. It can operate from 6 to 36 kbit/s in mono and from 7 to 48 kbit/s in stereo. The sampling frequency can range from 8 to 48 kHz. The algorithmic details of AMR-WB+ can be found in [16].

The EAAC+ and AMR-WB+ standards are both designed for mono and stereo operations, and target general audio at low-rates. For low rate wideband speech and audio communications, another available codec, standardized in 1999, is ITU-T recommendation G.722.1 [19]. The encoder operates on monophonic wideband signals sampled at 16 kHz and supports bit rates of 24 and 32 kbit/s. Lower rates (16 kbit/s) are also supported by the encoder, but significant degradation of speech quality prevented the ITU-T to recommend its use at these rates. Furthermore,

this standard is recommended for use in hands-free systems with channels having very low bit error rates. This is suitable for teleconferencing applications, one of the main targets for this standard. G.722.1 uses a transform-based encoder as in perceptual audio coding. However, the frequency components are encoded using very simple scalar quantizers and entropy coding, without the use of sophisticated perceptual models. Hence, one of the main characteristics of G.722.1 is its low complexity, reported at 15 MOPS (million operations per second) for both the encoder and the decoder.

4 • Issues Related to Channel Conditions

The previous sections concentrated on the speech encoder and decoder, without consideration for the transmission channel. In most telecommunications applications, the transmission channel will not provide an error-free service. Bit errors will occur and packets will be delayed or lost. In data communications, this problem is alleviated by adding transport layers above the network layers in order to ensure that all bits are received correctly. This implies retransmission of lost data which adds considerable amounts of delay. This is exemplified by the TCP/IP model for the Internet. For real-time speech and audio communications, delay is critical, so retransmission is unacceptable. Hence, other solutions have emerged to alleviate the channel imperfections and maintain a suitable service quality for speech communications on lossy channels. These different solutions are presented in the following sections.

4.1 Adaptation to Channel Conditions

Adaptive multirate (AMR) speech coders support several coding rates and are capable of switching seamlessly and rapidly (on a frame-by-frame basis) from one rate to another. They were originally developed for wireless communications [20,21], where the quality of the communication channel (noise, interferences) varies greatly over time.

4.1.1 Adaptive Multirate Coders

The motivation for developing AMR solutions was the following. Over communication channels, the available bit rate is shared between source and channel coding. Because of transmission errors, increasing the bit rate dedicated to source coding to the detriment of the bit rate dedicated to channel coding does not necessarily increase the resulting audio quality. Hence, the underlying principle of AMR is to dedicate to source encoding the maximum bit rate allowed by the communication channel. When the communication channel is clean, most of the bit rate is dedicated to source coding. That amount decreases as the

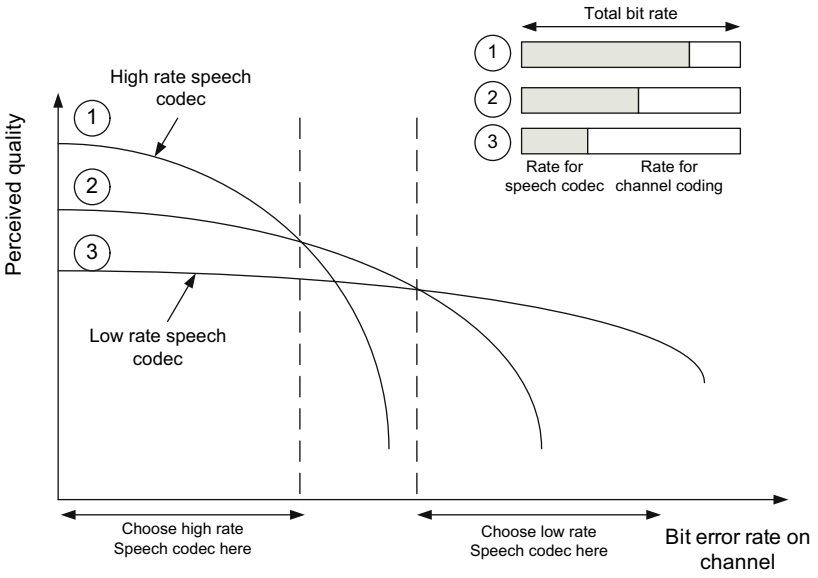


FIGURE 10 Principle of adaptive multirate (AMR) speech coding for channel with varying quality (bit error rate). Constant total bit rate (speech codec + channel codec) is assumed.

quality of the communication channel degrades, leaving more room for error correction. A similar strategy can be used when the communication channel is congested (varying available bit rate). This is illustrated in Figure 10. Case-1 corresponds to good channel conditions (low bit errors), while case-3 corresponds to poor channel conditions (high bit errors).

At the system level, AMR speech coders generally require a link adaptation unit to adjust the channel coding and speech coding schemes according to the quality of the communication channel [22].

There are two examples of standard AMR speech codecs. The "AMR narrowband" codec [23] is an ETSI and 3GPP standard. It uses the ACELP (algebraic code-excited linear prediction) coding algorithm [7]. It operates on narrowband (300–3,400 Hz) speech signals at eight rates in the range of 4.75–12.2 kbits/s. The coding modes at 12.2, 7.9, and 6.7 kbits/s correspond to the European ETSI EFR-GSM, the North American TIA IS-641, and the Japanese ARIB EFR-PDC codecs, respectively. The standard includes algorithms at the decoder for concealing lost frames. In addition to the AMR capability, the standard also includes some source-controlled variable bit rate features: voice activity detection (VAD), discontinuous transmission (DTX), and comfort noise generation (CNG). Finally, the AMR standard describes an associated channel coding, capable of error correction and bad frame detection, for full-rate (22.8 kbits/s) and half-rate

(11.4 kbits/s) channels of GSM. Introduction of AMR as an evolution of GSM systems has improved significantly the quality of service in these systems.

The "AMR wideband" (AMR-WB) codec [5] was the first codec adopted as a standard for both wireless and wireline services. It was standardized by ETSI/3GPP in 2001 and approved by the ITU-T in 2002 under reference G.722.2. It also interoperates with the latest 3GPP2 wideband standard, VMR-WB at mode 3 [11]. Compared to AMR narrowband, the audio bandwidth in AMR-WB is increased from 300–3,400 Hz to 50–7,000 Hz. This increases intelligibility and naturalness of speech and creates a feeling of transparent communication. The AMR-WB codec operates at eight rates in the range of 6.6–23.85 kbits/s and supports VAD/CNG/DTX.

RTP payload and file storage formats for AMR and AMR-WB supporting forward error correction (FEC), interleaving, and VAD/DTX/CNG were defined by the IETF [24].

4.1.2 Embedded Codecs

In communication systems using the AMR approach, whenever a decision is taken to reduce the source bit rate somewhere in the network (for example, to avoid congestion), that decision must be relayed both forward to the decoder and backward to the encoder. Rate switching of the encoder is therefore necessarily delayed, and some data can be lost in the meanwhile. There are also some contexts in which the AMR approach is not appropriate. In point-to-multipoint (or multicast) applications, for example, not all destinations get the same bit rate due to channel limitations or end device capability. In order to accommodate all destinations in a multicast application, a sender using an AMR codec would have to use the bit rate taken by the destination with the smaller bandwidth. This would penalize all the other destinations.

In an embedded codec, the bitstream is split into a core layer and one or several enhancement layers. Encoding is typically performed in stages, where a core encoder is first applied to the input signal. Then, enhancement layers are applied in successive stages to progressively reduce the distortion and improve the quality. Decoding all layers provides the highest quality. The encoder generates all layers at all time. However, at some stage during transmission, some of the upper layers can be stripped off the bitstream. The decoder is still able to decode the layers it has received, with the quality gracefully degrading as the number of received layers decreases. The encoder does not need to know whether some layers have been discarded or not somewhere in

the communication chain. The encoder just provides a constant bitstream comprising all layers, and the network nodes or the receiver decides how many layers can be transmitted to the next node or how many layers can be used for decoding. The decision can be based on network congestion or on the ability of the decoder to interpret all layers. For example, an embedded coder could comprise a standard core layer along with improvement layers. Older receivers could decode only the core, standard layer, while newer receivers could decode all layers for improved service quality.

Embedded coding was first studied for sample-based waveform coders. The basic PCM (pulse code modulation) coder, known as ITU-T recommendation G.711, can very easily be embedded, as some of the least significant bits can be stripped off a PCM codeword with gradually increasing degradation of quality. With minor modifications to the encoder and decoder, an ADPCM (adaptive differential pulse code modulation) codec can also be turned into an embedded codec, at the expense of only slight degradation of its performances [25]. This principle is used by the ITU-T G.727 standard and ANSI T1.310 ADPCM codec which operate at 2–5 bits per sample (16–40 kbits/s). As for the ITU standard G.722, it describes a wideband (7 kHz) subband ADPCM codec at 64 kbits/s with three layers at 48, 56, and 64 kbits/s.

The embedded coding principle has also been applied to other coding algorithms. Reference [26], for example, describes a CELP-based embedded coder using one adaptive codebook (ACB) and three cascaded innovative codebooks (ICB), the ACB and first ICB forming the core layer and the two other ICBs forming the upper, refinement layers. Reference [27] describes a two-layer embedded coder, with a core using linear-prediction (ACELP) speech coding and one enhancement layer using transform-based (MDCT) audio coding.

The ITU-T G.729.1 codec [28] is, to date, the most accomplished example of an embedded codec. Its bitstream is organized into 12 layers ranging from 8 to 32 kbits/s. The first two layers (8 and 12 kbits/s) are narrowband and speech specific while the others (14–32 kbits/s with 2 kbits/s steps) are wideband and can deal with more general audio signals. The core layer (8 kbits/s) is compatible with the ITU-T G.729 standard (this was decided to ensure interoperability with already deployed VoIP systems, which widely support G.729). The first two layers (8 and 12 kbits/s) are populated by a cascade CELP codec. The third layer (14 kbits/s) performs a parametric time-domain bandwidth

extension to achieve wideband speech quality. The upper layers are populated by a transform codec. The input and output signals can be sampled at either 8 or 16 kHz. Another interesting feature of the ITU G.729.1 embedded codec is that its bitstream includes some frame erasure concealment (FEC) information (namely a voicing classification, a phase, and an energy) for improved concealment of (and recovery after) lost frames similar to that of the 3GPP2 VMR-WB codec.

4.1.3 Variable Bit Rate Codecs

In most communication systems, it is desirable to accommodate as many users as possible while providing each of them with an acceptable quality of service, whatever the condition of the communication channel (traffic load, noise, and interference levels). Channel-controlled AMR codecs offer a means for adapting the source bit rate to the channel conditions. However, AMR codecs do not take into account the fact that some parts of the speech signals can be compressed more efficiently than others. Source-controlled variable bit rate speech coders were designed to minimize the average data rate (ADR) required by each user while continuously maintaining voice quality above a certain level [29].

Significant reduction in ADR is very easily achieved by simply taking into account the fact that, during a conversation, a speaker is silent almost 60% of the time. Most, if not all, modern speech coders use the VAD/DTX/CNG approach to take advantage of that small activity factor typical of conversational speech. Voice activity detection (VAD) partitions the speech signal into active (speech) and inactive (silence and background noise) periods. Discontinuous transmission (DTX) is used during inactive periods, as silence and background noise vary much more slowly than active speech and require very little bit rate. Comfort noise generation (CNG) is used at the decoder to recreate the acoustic background noise. Information about the acoustic background noise is sent within SID (silence descriptor) frames that are encoded at a much lower bit rate than speech (typically, 1 kbits/s) and also sent at a much lower update rate (typically, once every eight frames).

Further reduction in average bit rate can be obtained when the varying nature of speech is also taken into account [30]. During unvoiced segments, for example, the long-term predictor (adaptive codebook in Figure 2) used by most CELP codecs is unnecessary and the corresponding bits can be saved. Also, those noise-like segments are much less sensitive to a coarse quantization of their waveform. Onsets on the other hand are critical

events that, though they do not occur very frequently, need to be quantized accurately.

A number of narrowband speech coding standards using the source-controlled variable bit rate approach were designed for wireless telephony over CDMA systems. The first one was the North American TIA/EIA/IS-96 standard (also known as QCELP) created in 1993 [31]. In this standard, each 20-ms input frame is encoded at either 0.8, 2, 4, or 8.5 kbits/s (rate set I) depending on the nature of the input signal. The lowest rate is for background noise. The channel rates corresponding to those source rates are 1.2, 2.4, 4.8, and 9.6 kbits/s. The TIA/EIA/IS-733 standard (also known as Q-13) was established in 1995 and uses another rate set: 1.0, 2.7, 6.2, and 13.3 kbits/s (rate set II). The TIA/EIA/IS-127 (EVRC) standard was created in 1997 to replace IS-96 (same rate set except that the 2-kbits/s quarter rate is not used). EVRC uses the RCELP (relaxed CELP) technique [32] which consists of modifying the input speech signal (specifically, simplifying its pitch contour) so that the pitch information can be coded more efficiently (only one pitch value per frame need to be transmitted instead of four as in conventional CELP). The SMV (selectable mode vocoder) [33] standard was created in 2001 to replace EVRC. It uses the same rate set as the IS-96 standard. It also supports several network-controlled operating modes (premium, standard, economy, and capacity saving) that, by making different usages of the rates, provide different average data rates. Due mainly to its high computational complexity, the SMV standard was never really used however. EVRC-B is an enhancement of EVRC that, mostly by making use of the quarter rate, provides a lower average data rates for a given voice quality. The 4GV codec that should be used in CDMA2000 is based on EVRC-B.

To date, the most accomplished source-controlled variable bit rate codec is the variable rate multimode wideband (VMR-WB). This codec was standardized in 2004 by 3GPP2 for wideband telephony and multimedia streaming services over CDMA2000 [11]. Among its numerous benefits, the VMR-WB codec is compatible with rate sets I and II. It supports several operating modes achieving various ADRs. One of its coding modes (mode 3) is interoperable with the ITU-T and 3GPP standard AMR-WB. It can operate on either narrowband or wideband input signals and deliver either narrowband or wideband decoded signals. This is the first speech coding standard for cellular telephony to include an algorithm for acoustic background noise reduction [12] (only the very low bit rate military standard NATO STANAG 4591 which is based on mixed excitation linear predictive – MELPe – also includes noise reduction). At a given bit rate, different

coding strategies are used for different classes of signals. At 6.2 kbits/s, for example, frames are encoded using an unvoiced, a voiced, or a generic coding model. Among other differences, the unvoiced coding model uses a Gaussian innovative codebook while the others use an ACELP innovative codebook. For voiced signals, a signal modification technique equivalent to that of RCELP is applied to improve the coding efficiency [34]. Again, this is the first speech coding standard to include in its bitstream specific FEC (frame erasure concealment) information to improve concealment and recovery in case of frame losses [35]. It is therefore extremely robust to frame erasures (up to 6%). To give an idea of the efficiency of this codec, the quality obtained at 5.5 kbits/s (the codec operating in mode 1 on a wideband speech signal with 60% voice activity) is equal to that of the ITU G.722 standard (subband ADPCM) at 56 kbits/s.

4.2 Robustness to Channel Impairments

The problem of channel impairments is traditionally handled by channel coding. When transmitting general data, two channel-coding approaches are considered: ARQ (automatic repeat request) and FEC (forward error correction). ARQ consists of sending again any data that was not correctly received by the destination. FEC consists of adding redundant information to the source information so that the destination can detect and correct transmission errors. Typical FEC methods include parity checking or cyclic redundancy checking (CRC) to detect bit errors and block convolutional or turbo codes to correct them [36].

Because of the delay it introduces, ARQ is generally not considered when transmitting speech (although it can be used on some high-capacity sections of the network). FEC on the other hand is widely used in voice communication systems. There is a big difference however between data transmission and speech transmission, as speech decoders are generally tolerant, to a certain extent, to bit errors. As a matter of fact, from the perspective of voice quality it is preferable to receive frames containing some bit errors than to receive no frame at all. This is a central paradigm to the way FEC is used with speech coders.

FEC is therefore used in speech communication to minimize the percentage of residual bit errors on the source bitstream. Some techniques that are commonly used in speech coding to minimize the impact of those errors are described in the following subsection. Contingency measures that are taken when a frame is declared totally lost by the system or the channel decoder will then be described in the following subsection (frame erasure concealment).

4.2.1 Improving Robustness to Bit Errors

Channel-Optimized and Robust Vector Quantization

Vector quantization (VQ) [37] is extensively used in speech coding. It consists of replacing an input vector of parameters by another (the "codeword") taken from a finite set of vectors (the "codebook") that is supposed to be representative of the whole input space. In Figure 2, VQ is used to quantize the synthesis filter parameters, the innovative excitation, and the gains of the adaptive and innovative codebooks. In VQ, coding is achieved by only transmitting the position (called the "index") of the selected codeword within the codebook. VQ is known to be sensitive to bit errors, as one single bit error in an index can lead to a codeword that comes from a totally different region of the multidimensional input space.

A number of "joint source-channel coding" techniques have been proposed to improve the robustness of vector quantizers to bit errors. Two main approaches are possible [38]. The first one is referred to as channel-optimized vector quantization (COVQ). In that approach, the distortion measure used when designing the vector quantizer takes into account the possibility of bit errors. The resulting vector quantizer is therefore redundant, as it essentially performs channel coding [39]. The second approach is referred to as "robust vector quantization". In that approach, the impact of bit errors on a given quantizer is minimized by carefully choosing the index assignment (i.e., the pairing between indexes and codewords). Ideally, codewords that are closer in terms of their index (Hamming distance) should also be closer in terms of their value (Euclidian distance, or distortion). Index assignment can be optimized using binary switching [40], for example.

Predictive Quantization

Prediction is another key technique for speech coding. When a parameter is temporally correlated, quantizing the difference between this parameter and a predicted value is more efficient than quantizing the parameter itself. Prediction however causes error propagation in the case of transmission errors. Moving average (MA) predictors are generally preferred to Autoregressive (AR) predictors, as error propagation is limited to a duration on the order of prediction in the MA case while it is potentially

infinitely long (depending on the value of the prediction coefficients) in the AR case [41]. Robustness of both AR- and MA-based quantizers can be further improved by using a safety-net approach [42]. Those considerations apply both to bit errors and frame erasures.

Unequal Error Protection

Another common way to improve robustness of speech communication against bit errors is via the use of unequal error protection (UEP). UEP is based on the fact that not all bits within a coded frame are equally sensitive to bit errors. Therefore, not all bits require the same amount of protection from channel coding. In a CELP coder, for example, and specifically in a CELP coder using the ACELP model for the innovative excitation, it is well known that the adaptive codebook (delay and gain) is much more sensitive than the innovation codebook. The sensitivity of the bits that compose a coded frame can be determined automatically, for example, by systematically inverting a given bit and measuring the output degradation using a signal-to-noise ratio or a more perceptually correlated measure such as PESQ (perceptual evaluation of speech quality) [43]. Bits are then generally grouped into three classes of bits: class A, class B, and class C. Class A contains the bits most sensitive to errors, and any error in these bits typically results in a corrupted speech frame which should not be decoded without applying appropriate frame erasure concealment. Classes B and C contain bits where increasing error rates gradually reduce the speech quality, but decoding of an erroneous frame is usually possible without annoying artifacts. Class B bits are more sensitive to errors than class C bits. In the AMR-WB standard codec at 12.65 kbits/s, for example, 72 bits out of the 253 that comprise one coded speech frame are considered very sensitive and belong to class A [44]. They are the VAD bit, some bits from the multistage split LPC quantizer (specifically from the first quantization stage and from low frequencies), the adaptive codebook delay, and the codebook gains.

4.2.2 Improving Robustness to Lost Frames

The techniques used to mitigate the effects of a lost or erased frame can be divided into two main groups. The first group of techniques, generally referred to as “concealment”, consists of computing a replacement frame for the lost frame. The second group of techniques deals with the recovery of the decoder when good frames are received again. The main techniques are described in the following. Many other solutions can be found in [45–47].

Frame Loss Concealment

For non-predictive or sample-based codecs such as PCM, frame loss concealment can be done either in the temporal domain using prediction techniques [48] or in the frequency domain using sinusoidal modeling [49]. Combining those two approaches provides a packet loss concealment mechanism which adapts to the signal characteristics and is not restricted to pure speech signals [50].

In CELP-based speech codecs, frame loss concealment is facilitated by the presence of a speech model. An input parameter called the bad frame indicator (BFI) is used to indicate to the decoder that the frame is lost or corrupted. The decoder then generally simply extrapolates the coding parameters of the previous correctly received frame [51]. Interpolation requires a one-frame delay at the decoder but gives better results when a single frame is lost [27].

Frame Loss Recovery

Recovery after a lost frame is not a problem for non-predictive or sample-based codecs. For those codecs, a simple fade-in fade-out between the concealed segment and the following one does the job.

Recovery can be highly problematic however for predictive codecs such as CELP. As a matter of fact, interframe dependencies due, for example, to the use of long-term prediction or of predictive quantizers (for parameters such as the gains or the spectral envelope) can cause the error introduced by one single lost frame to propagate through several frames. This is particularly true when a critical segment of signal (an onset for example) is lost [52]. It has been proposed in [53] to completely eliminate interframe dependencies from the codec. As was shown in [54], however, this approach is not necessarily more efficient than using standard CELP-based codecs together with straightforward bitstream-based redundancy. Another solution to limit error propagation in the decoder is to introduce some constraints in the encoder [55]. Those constraints can be further exploited at the decoder to speed up recovery [56]. Yet another solution used in the VMR-WB and G.729.1 standards [27,35] consists of sending specific frame erasure concealment information (voicing classification, energy, and pitch phase) within the bitstream. That information can be used by the decoder to improve both concealment (using the classification and energy) and recovery (using mainly the phase of the signal). As will be shown later, in some other contexts (voice over packet networks), late frames when they are available can also be used to limit error propagation [57,58].

4.2.3 Adaptation to Packet Networks

Packet switching is a very challenging environment for voice codecs. This section describes the main problems encountered when transmitting voice over packet networks and presents some specific processing technologies and adaptations that are typically brought to the codec to improve voice quality over packet networks (VoIP).

Jitter Buffering

VoIP is characterized by variations in the time that packets take to transit through the network. Jitter buffering is generally used by the receiver to control the effects of the jitter (which can be defined as the difference between the actual arrival time of the packets and a reference clock at the normal packet rate). This buffer works by introducing an additional "playout" delay (which is defined with respect to the reference clock that was, for example, started at the reception of the first packet). The playout delay is chosen so as to provide the best compromise between the latency of the communication and the percentage of late packets. This delay allows transforming the uneven flow of arriving packets into a regular flow of packets as seen at the output of the jitter buffer. In this way, the decoder can provide a sustained flow of speech to the listener, provided the application can sustain the additional delay produced by the jitter buffer.

Time Scaling

Several strategies [59] can be used to adapt the playout delay to the network conditions. In the fixed jitter buffer strategy, the playout delay is set at the beginning of the conversation and maintained afterward. In the adaptive jitter buffering strategy, the playout delay is changed only at the beginning of each silence period ("talk-spurt"-based approach). For quickly varying networks, better results are obtained when the playout delay is also adapted during active speech [60].

Changing the playout delay during periods of active speech requires a means to change the rhythm of the speech signal. As explained in Table 1, playing out shorter frames (i.e., speeding up speech) decreases the playout delay, while playing out shorter frames (i.e., slowing down speech) increases it.

Time scaling can be done on the decoded signal (in the speech samples domain) using techniques such as SOLA [61] or its variants, PSOLA [62], or TDHS [63]. When a predictive codec is used, this operation can also be carried out during decoding, for example, in the excitation domain. As shown in [64], this approach has many advantages in terms of complexity and voice quality. In particular, it does not involve any additional complexity, provided some very reasonable constraints are applied with respect to when and how time scaling is applied.

TABLE 1 The algorithm for playout delay adaptation (the playout delay for packet i is the difference between P_i and the reference clock)

1.	Using past jitter values, estimate the "ideal" playout time $\hat{P}_{i+1}$ of frame number i+1;
2.	Send frame number i to the time scaling device, requesting it to generate a frame of length $\hat{T}_i = \hat{P}_{i+1} - P_i$ where P_i is the playout time of frame i.
3.	The actual playout time of frame i+1 is$P_{i+1} = P_i + T_i$ where T_i is the actual frame length achieved by the time scaling device
4.	Play out frame number i when its playout time P_i has arrived and iterate from step 1.

Late Packet Processing

Despite all the efforts done to adapt the playout delay to the jitter, some packets will still arrive too late to be decoded and must be considered as lost. As mentioned before, the concealment procedure used to replace the missing segment of speech is not perfect and introduces an error in the concealed segment. Moreover, in predictive codecs such as CELP, the concealment procedure does not correctly update the internal state of the decoder which becomes desynchronized with the encoder. Thus the error propagates in the following segments. A method to improve the recovery of a speech decoder after the reception of one or several late frames has been proposed in [57,58]. Rather than considering the late frame as "lost", it is used to update the internal state of the decoder. This limits and in some cases stops error propagation.

Other Techniques Used in VoIP Applications

Forward error correction (FEC) is widely used in VoIP applications to minimize the occurrence of lost frames. As mentioned earlier, FEC consists of sending redundant information to the destination. A simple and generic method to get redundant information using parity codes based on the exclusive OR (XOR) is described in [65]. Redundant information can also be obtained by repetition or double encoding [66]. Repetition consists of sending multiple copies of the coded frame in different packets. Double encoding consists of using one primary encoder, and one or several secondary encoders operating at a lower bit rate. FEC can be made channel controlled as proposed in [67,68], or source controlled as proposed in [52]. More elaborate solutions can be found in [69].

Interleaving, which consists of never sending two contiguous frames in the same packet, can be used to avoid bursts of frame erasures. This is useful, as frame erasure concealment procedures perform better for isolated erasures. This technique introduces an

additional delay, however, and does not decrease the percentage of lost frames.

Finally, an IP transport protocol called UDP Lite can be used on error-prone networks [70]. UDP normally discards the whole packets when the checksum indicates the presence of bit errors. UPD Lite, by limiting the scope of the checksum, allows the destination to receive frames with bit errors. This is useful as voice codecs, as mentioned above, are rather tolerant to bit errors.

5 • Current Developments and Future Challenges

The last 15 years has seen the emergence of numerous speech and audio coding standards. These standards have followed the evolution of communication systems and application needs. The exponential growth of cellular systems and the Internet as a multimedia environment is largely due to the existence of efficient speech and audio coding algorithms (MP3 is more widely known than speech coding standards, but the latter are no less important for wireless communication systems as we know them today). In this chapter, we have presented several ways that signal processing can be used to improve service and audio quality in various input and channel conditions. Increasingly, these solutions are applied to existing standards, by either combining them or extending their capability while maintaining interoperability. For example, embedded coders typically use a well-deployed standard as the core. Emerging standards often ensure interoperability with other existing standards in a least part of their operating modes.

What has driven recent algorithmic development and standards adoption is in particular the move toward wideband telephony and multimedia handheld devices. Interconnecting the different heterogeneous networks (cellular and Internet, for example) is certainly a driving force to find innovative solutions to allow access to any information, in any format from anywhere. We can expect significant advances in signal processing and information coding to facilitate these interconnections, and to ensure that the large amounts of multimedia data already available in different formats will remain accessible in these evolving networks.

References

1. A.S. Spanias, “Speech coding: a tutoral review”, Proceedings of the IEEE, vol. 82, no. 10, pp. 1541–1582, October 1994.
2. B. Kleijn and K. Paliwal, eds., *Speech Coding and Synthesis*, Elsevier, 1995.

3. L.R. Rabiner, R.W. Shafer, *Digital Processing of Speech Signals*, Prentice-Hall Signal Processing Series, 1978.
4. R.A., Salami, L., Hanzo, R., Steele, K.H.J. Wong, and I. Wassell, Speech coding, in R., Steele, eds., *Mobile Radio Communications*, chapter 3, pp. 186–346. IEEE Press – Wiley, 1992.
5. B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)", IEEE Transactions on Speech and Audio Processing, vol. 10, no. 8, pp. 620–636, November 2002.
6. R. Salami, C. Laflamme, B. Bessette, and J.-P. Adoul, "ITU-T Recommendation G.729 Annex A: reduced complexity 8 kbit/s CS-ACELP codec for digital simultaneous voice and data", IEEE Communications Magazine, vol. 35, no. 9, pp. 56–63, September 1997.
7. C. Laflamme, J.P. Adoul, H.Y. Su, and S. Morissette, "On reducing computational complexity of codebook search in CELP coder through the use of algebraic codes", IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 177–180, Albuquerque, New Mexico, USA, April 3–6, 1990.
8. K. Järvinen et al. "GSM enhanced full rate codec", IEEE 1997 International Conference on Acoustics, Speech and Signal Processing, pp. 771–774, Munich, Germany, April 20–24, 1997.
9. ITU-T Recommendation P.48, "Specification for an intermediate reference system, volume V of the Blue Book", pp. 81–86, ITU, Geneva, February 1996.
10. J. Thiemann, *Acoustic Noise Suppression for Speech Signals Using Auditory Masking Effects*, Masters Thesis, McGill University, 2001.
11. S. Ahmadi and M. Jelinek, "On the architecture, operation, and applications of VMR-WB: the new cdma2000 wideband speech coding standard", IEEE Communications Magazine, vol. 44, no. 5, pp. 74–81, May 2006.
12. M. Jelinek and R. Salami, "Noise reduction method for wideband speech coding", 12th European Signal Processing Conference (EUSIPCO 2004), pp. 1959–1962, Vienna, Austria, September 6–10, 2004.
13. A.S. Spanias, "Perceptual coding of digital audio", Proceedings of the IEEE, vol. 88, no. 4, pp. 451–513, April 2000.
14. E. Ordentlich and Y. Shoham, "Low-delay code-excited linear-predictive coding of wideband speech at 32 kbps", IEEE 1991 International Conference on Acoustics, Speech, and Signal Processing (ICASSP'91), pp. 9–12, Toronto, Canada, May 14–17, 1991.
15. 3GPP Technical Specification TS26.401, "General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description", June 2006.
16. 3GPP Technical Specification TS26.290, "Audio codec processing functions; Extended adaptive multi-Rate – wideband (AMR-WB+) codec; Transcoding functions", June 2005.
17. M. Schug, A. Groschel, M. Beer, and F. Henn, "Enhancing audio coding efficiency of MPEG Layer-2 with spectral band replication (SBR) for DigitalRadio (EUREKA 147/DAB) in a backwards

compatible way", 114th Audio Engineering Society Convention, preprint no. 5850, Amsterdam, The Netherlands, March 22–25, 2003.
18. R. Salami, R. Lefebvre, and C. Laflamme, "A wideband codec at 16/24 kbit/s with 10 ms frames", 1997 IEEE Workshop on Speech Coding, pp. 103–104, Pocono Manor, Pennsylvania, USA, September 7–10, 1997.
19. ITU-T Rec. G.722.1, "Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss", September 1999.
20. T.B. Minde, S. Bruhn, E. Ekudden, and H. Hermansson, "Requirements on speech coders imposed by speech service solutions in cellular systems", 1997 IEEE Workshop on Speech Coding, pp. 89–90, Pocono Manor, Pennsylvania, USA, September 7–10, 1997.
21. A. Uvliden, S. Bruhn, and R. Hagen, "Adaptive multi-rate. A speech service adapted to cellular radio network quality", 32nd Asilomar Conference on Signals, Systems and Computers, vol. 1, pp. 343–347, Pacific Grove, California, USA, November 1–4, 1998.
22. S. Bruhn, P. Blocher, K. Hellwig, and J. Sjöberg, "Concepts and solutions for link adaptation and inband signalling for the GSM AMR speech coding standard", IEEE Vehicular Technology Conference, pp. 2451–2455, Amsterdam, The Netherlands, September 19–22, 1999.
23. K. Järvinen, "Standardisation of the adaptive multi-rate codec", 10th European Signal Processing Conference (EUSIPCO 2000), pp. 1313–1316, Tampere, Finland, September 4–8, 2000.
24. J. Sjöberg, M. Westerlund, A. Lakaniemi, and Q. Xie, "Real-time transport protocol (RTP) payload format and file storage format for the adaptive multi-rate (AMR) and adaptive multi-rate wideband (AMR-WB) audio codec", IETF RFC 3267, June 2002.
25. D.J. Goodman, "Embedded DPCM for variable bit rate transmission", IEEE Transactions on Communications, vol. 28, no. 7, pp. 1040–1046, July 1980.
26. R.D. De Iacovo and D. Sereno, "Embedded CELP coding for variable bit-rate between 6.4 and 9.6 kbit/s", IEEE 1991 International Conference on Acoustics, Speech, and Signal Processing (ICASSP'91), pp. 681–684, Toronto, Canada, May 14–17, 1991.
27. S.A. Ramprashad, "A two stage hybrid embedded speech/audio coding structure", IEEE 1998 International Conference on Acoustics, Speech, and Signal Processing (ISACCP'98), pp. 337–340, Seattle, Washington, USA, May 12–15, 1998.
28. S. Ragot et al., "ITU-T G.729.1: an 8–32 kbit/s scalable coder interoperable with G.729 for wideband telephony and voice over IP", IEEE 2007 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2007), Honolulu, Hawaii, USA, April 15–20, 2007.
29. A. Gersho, and E. Paksoy, "An overview of variable rate speech coding for cellular networks", International Conference on Selected Topics in Wireless Communications, pp. 172–175, Vancouver, Canada, June 25–26, 1992.

30. E. Paksoy, K. Srinivasan, and A. Gersho, "Variable Bit Rate CELP coding of speech with phonetic classification", European Transactions on Telecommunications and Related Technologies, vol. 5, no. 5, pp. 591–602, September–October 1994.
31. A. DeJaco, W. Gardner, P. Jacobs, and C. Lee, "QCELP: The North American CDMA digital cellular variable rate speech coding standard", 1993 IEEE Workshop on Speech Coding for Telecommunications, pp. 5–6, Sainte-Adèle, Québec, Canada, October 13–15, 1993.
32. W.B. Kleijin, P. Kroon, and D. Nahumi, "The RCELP speech-coding algorithm", European Transactions on Telecommunications and Related Technologies, vol. 5, no. 5, pp. 573–582, September–October, 1994.
33. S.C. Greer, and A. DeJaco, "Standardization of the selectable mode vocoder", IEEE 2001 International Conference on Acoustics, Speech and Signal Processing (ICASSP'01), pp. 953–956, Salt Lake City, Utah, USA, May 7–11, 2001.
34. M. Tammi, M. Jelinek, and V.T. Ruoppila, "A signal modification method for variable bit rate wideband speech coding", IEEE Transactions on Speech and Audio Processing, vol. 13, no. 5, pp. 799–810, September 2005.
35. M. Jelinek, R. Salami, S. Ahmadi, B. Bessette, P. Gournay, C. Laflamme, and R. Lefebvre, "Advances in source-controlled variable bit rate wideband speech coding", Special Workshop in Maui (SWIM), Lectures by Masters in Speech Processing, Maui, Hawaii, January 12–14, 2004.
36. A. Glavieux, *Channel Coding in Communication Networks: From Theory to Turbo Codes*, Iste Publishing Company, 2007.
37. A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1991.
38. M. Skoglund, "On channel-constrained vector quantization and index assignment for discrete memoryless channels", IEEE Transactions on Information Theory, vol. 45, no. 6, pp. 2615–2622, November 1999.
39. H. Kumazawa, M. Kasahara, and T. Namekawa. "A construction of vector quantizers for noisy channels", Electronics and Engineering in Japan, vol. 67-B(1), pp. 39–47, January 1984.
40. K. Zeger, and A. Gersho, "Pseudo-gray coding", IEEE Transactions on Communications, vol. 38, no. 12, pp. 2147–2158, December 1990.
41. J. Skoglund, and J. Linden, "Predictive VQ for noisy channel spectrum coding: AR or MA?", IEEE 1997 International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97), pp. 1351–1354, Munich, Germany, April 21–24, 1997.
42. T. Eriksson, J. Linden, and J. Skoglund, "Interframe LSF quantization for noisy channels", IEEE Transactions on Speech and Audio Processing, vol. 7, no. 5, pp. 495–509, September 1999.
43. J.G. Beerends, A.W. Rix, M.P. Hollier, and A.P. Hekstra, "Perceptual evaluation of speech quality (PESQ): The new ITU standard for end-to-end speech quality assessment, Part I – time-delay compensation;

Part II – psychoacoustic model", Journal of the Audio Engineering Society, vol. 50, no. 10, pp. 765–778, October 2002.
44. "AMR Wideband Speech Codec; Frame Structure", 3GPP Technical Specification 3GPP TS 26.201, March 2001.
45. M. Chibani, "Increasing the robustness of CELP speech codecs against packet losses", Ph.D. Thesis, University of Sherbrooke, Canada, January 2007.
46. C. Perkins, O. Hodson, and V. Hardman, "A survey of packet-loss recovery techniques for streaming audio", IEEE Network, pp. 40–48, September–October 1998.
47. B.W. Wah, X. Su, and D. Lin, "A survey of error-concealment schemes for real-time audio and video transmission over the Internet", 2000 International Symposium on Multimedia Software Engineering, pp. 17–24, Taipei, Taiwan, December 11–13, 2000.
48. E. Gündüzhan and K. Momtahan, "A linear prediction based packet loss concealment algorithm for PCM coded speech", IEEE Transaction on Speech and Audio Processing, vol. 9, no. 8, pp. 778–785, November 2001.
49. J. Lindblom and P. Hedelin, "Packet loss concealment based on sinusoidal modeling", 2002 IEEE Speech Coding Workshop Proceedings, pp. 65–67, Ibaraki, Japan, October 6–9, 2002.
50. V. Vilaysouk and R. Lefebvre, "A hybrid concealment algorithm for non-predictive wideband audio coders", 120th Audio Engineering Society Convention, preprint no.6670, Paris, France, May 20–23, 2006.
51. R. Salami, C. Laflamme, J. Adoul, A. Kataoka, S. Hayashi, T. Moriya, C. Lamblin, D. Massaloux, S. Proust, P. Kroon, and Y. Shoham, "Design and description of CS-ACELP: a toll quality 8 kb/s speech coder", IEEE Transactions on Speech Audio Processing, vol. 6, no. 2, pp. 116–130, March 1998.
52. H. Sanneck and N. Le, "Speech property-based FEC for Internet telephony applications", Proceedings of SPIE vol. 3969, pp. 38–51, Multimedia Computing and Networking 2000, San Jose, California, USA, January 24–26, 2000.
53. S.V. Andersen et al., "ILBC – A linear predictive coder with robustness to packet losses", 2002 IEEE Speech Coding Workshop Proceedings, pp. 23–25, Ibaraki, Japan, October 6–9, 2002.
54. R. Lefebvre, P. Gournay, and R. Salami, "A study of design compromises for speech coders in packet networks", IEEE 2004 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2004), pp. 265–268, Montréal, Canada, May 17–21, 2004.
55. M. Chibani, P. Gournay, and R. Lefebvre, "Increasing the robustness of CELP-based coders by constrained optimization", IEEE 2005 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2005), pp. 785–788, Philadelphia, Pennsylvania, USA, March 19–23, 2005.
56. M. Chibani, R. Lefebvre, and P. Gournay, "Resynchronization of the Adaptive codebook in a constrained CELP codec after a frame erasure", IEEE 2006 International Conference on Acoustics, Speech

and Signal Processing (ICASSP'2006), pp. 13–16, Toulouse, France, March 14–19, 2006.
57. P. Gournay, F. Rousseau, and R. Lefebvre, "Improved packet loss recovery using late frames for prediction-based speech coders", IEEE 2003 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2003), pp. 108–111, Hong Kong, April 6–10, 2003.
58. K.D. Anderson, and P. Gournay, "Pitch resynchronization while recovering from a late frame in a predictive decoder", 9th International Conference on Spoken Language Processing (Interspeech 2006 – ICSLP), pp. 245–248, Pittsburgh, Pennsylvania, USA, September 17–21, 2006.
59. A. Ramjee, J. Kurose, D. Towsley, and H. Schulzrinne, "Adaptive playout mechanisms for packetized audio applications in wide-area networks", The Conference on Computer Communications, 13th Annual Joint Conference of the IEEE Computer and Communications Societies, Networking for Global Communications (INFOCOM'94), pp. 680–688, Toronto, Canada, June 12–16, 1994.
60. Y.J. Liang, N. Färber, and B. Girod, "Adaptive playout scheduling using time-scale modification in packet voice communications", IEEE 2001 International Conference on Acoustics, Speech, and Signal Processing (ICASSP'2001), pp. 1445–1448, Salt Lake City, Utah, USA, May 7–11, 2001.
61. S. Roucos and A.M. Wilgus, "High quality time-scale modification for speech", IEEE 1985 International Conference on Acoustics, Speech, and Signal Processing (ICASSP'85), pp. 493–496, Tampa, Florida, USA, March 26–29, 1985.
62. H. Valbret, E. Moulines, and J.-P. Tubach, "Voice transformation using PSOLA technique", Speech Communication, vol. 11, no. 2–3, pp. 175–187, June 1992.
63. D. Malah, "Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 27, no. 2, pp. 121–133, April 1979.
64. P. Gournay, and K.D. Anderson, "Performance analysis of a decoder-based time scaling algorithm for variable jitter buffering of speech over packet networks", IEEE 2006 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2006), pp. 17–20, Toulouse, France, March 14–19, 2006.
65. J. Rosenberg, and H. Schulzrinne, "An RTP payload format for generic forward Error Correction", IETF RFC 2733, December 1999.
66. C. Perkins, I. Kouvelas, O. Hodson, V. Hardman, M. Handley, J.C. Bolot, A. Vega-Garcia, and S. Fosse-Parisis, "RTP payload for redundant audio data", IETF RFC2198, September 1997.
67. J.C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-based error control for Internet telephony", Proceedings of IEEE INFOCOM'99, pp. 1453–1460, March 1999.
68. C. Padhye, K. Christensen, and W. Moreno, "A new adaptive FEC loss control algorithm for voice Over IP applications", 19th

IEEE International Performance, Computing and Communication Conference (IPCCC 2000), pp. 307–313, Phoenix, Arizona, USA, February 20–22, 2000.
69. I. Johansson, T. Frankkila, and P. Synnergren, "Bandwidth efficient AMR operation for VoIP", 2002 IEEE Speech Coding Workshop Proceedings, pp. 150–152, Ibaraki, Japan, October 6–9, 2002.
70. L.-A. Larzon, M. Degermark, S. Pink, "The lightweight user datagram protocol (UDP-Lite)", L.-E. Jonsson and G. Fairhurst, eds., IETF RFC 3828, July 2004.

PART VI
AUDIO ENGINEERING

Matti Karjalainen

Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland

32 Transducer Models

Juha Reinhold Backman 623

Modelling Fundamentals • Historical Introduction • Current Literature Overview • Overview of the Models • Classical Principles • Fundamentals of Transducers • Dynamic Loudspeakers • Horn Loudspeakers • Electrostatic Loudspeakers • Microphones • Piezoelectric Transducers for Ultrasonic and Underwater Applications

33 Loudspeaker Design and Performance Evaluation

Aki Vihtori Mäkivirta 649

Loudspeaker as a System Design Problem • Crossover Filters • Frequency Response Equalization • Room Response Equalizers • System Delay and Delay Variation • Protection • Understanding Loudspeaker Operation by Modeling • Acoustic Measurement Techniques • Listening Tests

34 PA Systems for Indoor and Outdoor

Jan Felix Krebber 669

Introduction • The PA System • Planning and Setting Up the Loudspeakers • Sound Indoors • Sound Outdoors • Measurement Methods • Special Issues on Signal Processing for PA Systems

35 Beamforming for Speech and Audio Signals

Walter Kellermann 691

The Basic Concept • Data-Independent Beamforming • Differential Microphone Arrays • Statistically Optimum Beamformers • Blind Source Separation as Blind Beamforming • Concluding Remarks

36 Digital Audio Recording Formats and Editing Principles

Francis Rumsey 703

Digital Tape Recording • Digital Recording on Computer Mass Storage Media • Consumer Digital Formats • Digital Audio Editing Using Mass Storage Media

37 Audiovisual Interaction

Riikka Möttönen and Mikko Sams 731

Audiovisual Space • Audiovisual Speech • Neural Basis of Audiovisual Perception

38 Multichannel Sound Reproduction

Ville Pulkki 747

Introduction • Concepts of Spatial Hearing by Humans • Loudspeaker Layouts • Reproduction Methods • Capturing Sound for Multichannel Audio

39 Virtual Acoustics

Tapio Lokki and Lauri Savioja 761

Introduction • Auralization and Signal Processing • Waveguide Meshes in Room Acoustics

40 Audio Restoration

Paulo A.A. Esquef 773

A Brief History of the Recording Technology • Aims and Problems • Overview of Digital Audio Restoration • Future Directions

41 Audio Effects Generation

Udo Zölzer 785

Filters and Delays • Tremolo, Vibrato and Rotary Speaker • Chorus, Flanger and Phaser • Pitch Shifting and Time Stretching • Dynamics Processing • Room Simulation

42 Perceptually Based Audio Coding

Adrianus J.M. Houtsma 797

Introduction • Transparent Coding: General Strategy • Transparent Coding: Specific Examples • Subjective Evaluation of Sound Quality • Objective Evaluation of Sound Quality • Concluding Remarks

32

Transducer Models

Juha Reinhold Backman

Nokia, Espoo, Finland

1 • Modelling Fundamentals

This chapter discusses the basic models with emphasis on audio applications. Loudspeakers are most commonly used as an example of electroacoustic transducers yet, from a modelling point of view, they present the broadest range of challenges to the theoreticians. The fundamental principles are, however, applicable to all transducer problems (microphones, hydrophones, ultrasonics). The reader is assumed to be reasonably familiar with the fundamental concepts of electroacoustics; introductory summaries have been presented by, e.g., Poldy (1994) and Hickson and Busch-Vishniac (1997).

We shall mostly confine ourselves to the classical approach where the description of the full physical system is reduced into a more tractable model. There are different, to some extent independent, means towards achieving this goal:

- Linearization: replacing the system with its small-signal approximation
- Reducing the dimensionality of the problem
- Dividing the system into subsystems with simple coupling

Linearization, the basic step in mathematical treatment of any complex system, renders the solutions independent of amplitude and enables the superposition of solutions at different frequencies. In the discussion below the models are assumed fundamentally linear, except for some aspects of loudspeakers which are discussed separately.

Reducing the dimensionality from a three-dimensional physical description to scalar quantities is a feature more characteristic to the acoustic transducer design. The typical systems are well described through scalar mechanical variables (force, velocity), acoustical variables (pressure, volume velocity), and the electrical quantities.

The ability to divide the system into smaller subsystems better suited for modelling is typical to the audio transducer construction. This is to some extent related to the reduction of the dimensionality, since the well-defined input and output quantities of one-dimensional systems lend themselves quite naturally to the separation (also in the mathematical sense). This is further helped by the fact that often the electromechanical transduction and mechanical subsystems are the only significantly nonlinear parts in a model, and acoustical subsystems can be regarded as linear.

Subsystems common to all electroacoustic systems include

- Electrical, including amplification and driving/loading impedances
- Electromechanical coupling
- Mechanical aspects of the system (moving mass, compliance, damping, resonances, etc.)
- Mechanical–acoustical transformation (sound radiation or reception)
- Enclosure interior sound field
- Exterior sound field, including the effects of immediate transducer housing and acoustical coupling between transducers

In microphones, ultrasonic transducers, underwater transducers, and electrostatic loudspeakers the transducers are used as a single entity, either without an enclosure, etc., or with the enclosure already forming an inseparable part of the total design, whereas in conventional dynamic loudspeakers the same driver is quite often intended to be used in multitude of enclosures, which necessitates independent transducer and enclosure models.

2 • Historical Introduction

The fundamental idea of the correspondence between the electrical impedance and factors affecting the movement of the transducer was discovered during the early stages of audio development (Butterworth 1915; Kennelly and Pierce 1912). The basic ideas grew soon into maturity, and the more complex acoustical systems were discussed by Mason (1942), whose book

contains extensive references to early work in equivalent circuits and mechanical and acoustical signal processing. After Mason, Olson (1943) published his less detailed, but more general-purpose and audio-oriented work on the use of equivalent circuits. The theoretical understanding evolved further in the 1950s: Hunt (1954, pp. 92–127) published his treatment of transducer theory, Hueter and Bolt (1955) discussed transducers from ultrasonics point of view, and Fischer (1955) published his very clear introduction to the physical fundamentals of electroacoustic transduction.

For overviews, the author would rather refer the reader for the earliest loudspeaker theory developments towards McLachlan (1935), for overall development of electroacoustics until mid-1950s towards Hunt (1954, pp. 1–91) and Beranek (1954), and for a critical review of the more recently published loudspeaker model related work towards Gander (1998). Groves (1981) has edited an anthology with emphasis on early material.

3 • Current Literature Overview

In the modern loudspeaker literature the most usable materials on modelling are the compilation edited by Borwick (1994) (especially on electrostatic speakers and general aspects of lumped-parameter models) and the monographs by Eargle (1997) (especially on thermal models and magnetic system design) and Geddes and Lee (2002). Literature written for the amateur loudspeaker builders is typically rather weak on modelling methods, but notable exceptions to this are the books by Dickason (2000) and Bullock and White (1991). Most of the microphone literature concentrates on applications, but modelling aspects are discussed in compilations edited by Gayford (1994) and Wong and Embleton (1995). Key publications in loudspeaker, and to some extent microphone, theory have been collected in the Audio Engineering Society Anthology series (1978a, 1978b, 1978c, 1983, 1991a, 1991b, 1996). General transducer theory, especially for piezoelectric transducers and ultrasonic applications, is discussed in the Physical Acoustics series edited by Mason, Thurston, and Pierce. (Mason 1964a, 1964b; Mason 1973, 1979, 1981; Thurston and Pierce 1990).

4 • Overview of the Models

The various aspects of a typical (audio) electroacoustic transducer model are illustrated by the loudspeaker in Figure 1.

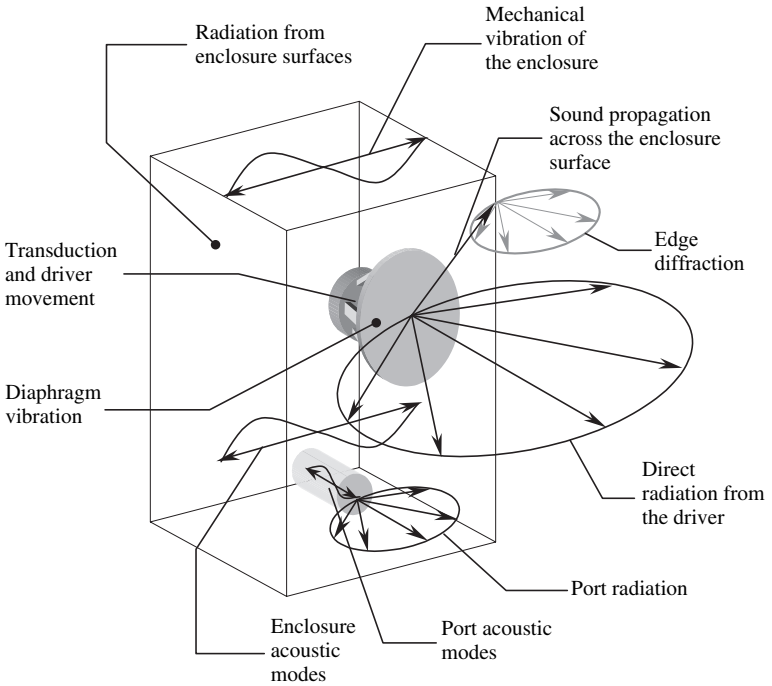


FIGURE 1 Breakdown of the acoustics of a loudspeaker enclosure from modelling point of view.

The classical electroacoustic system models, most often quoted in literature and very widely applicable, are the electrical analogue (or equivalent) circuits represented as lumped-component models. They are well applicable to traditional transducers, where the various subsystems can be clearly identified by physical inspection. The essential assumption for the simplest models is that the system dimensions are smaller than the wavelength, although the lowest drumhead mode of, for example, a microphone diaphragm can be approximated by a lumped-component model. The success of the lumped-component models is largely due to the ability to use the concepts and the tools of electric circuit theory directly in the mechanical–acoustical system modelling. Also, the transducers are always interfaced to an electric circuit, and in some cases the electric circuit has a significant effect on the apparent mechanical behaviour of the transducer (in both loudspeakers and ultrasonic transducers). The components of the equivalent circuit can be frequency-dependent (e.g. flow resistances or radiation impedances).

TABLE 1 A summary of complexity, accuracy, and computational load of some modelling approaches.

Model type	Complexity	Computational load	Accuracy	Remarks
Lumped parameter	Very low	Very low	Usable at low frequencies	Enabler for deriving analytical transducer theory
1-dimensional waveguide	Low	Low	Extended frequency range of validity, horns, etc., included	Strict geometric restrictions, usable also for ultrasonics
Waveguide + exterior diffraction	Moderate	Moderate	Medium to good for conventional loudspeakers	Moderate geometric restrictions
Lossless FEM/FD for enclosure interior or ultrasonic transducers	Moderate to high	High	Good for identifying eigenmodes of low-loss systems	Good for identifying potential problems, not for predicting their magnitude
Lossy (multiphysics) FEM/FD	High	High	Good for identifying realistic modal shapes for lossy systems	Difficulty in formulating the description of material losses
Coupled FEM/FD for acoustics and enclosure vibration	High	Very high	Good	Necessary for thin-walled structures or large enclosures
FEM/BEM/FD for interior (rigid enclosure), FEM/BEM for the exterior	High	Very high	Good, includes also diffraction and coupling through exterior field	Appropriate approximation for typical high-fidelity loudspeakers
Coupled FEM/BEM/FD for interior and enclosure, FEM/BEM/FD for the exterior	High	Very high	Potential for high accuracy	Applicable to all transducer types
Analytical solutions for acoustics	High	High for symbolic computation, low for numerical evaluation	Very high when underlying assumptions are met	Benchmark for numerical calculations, feasible only in limited cases

The electrical equivalent circuit approach can be developed further by introducing one-dimensional waveguide (or transmission line) models. These allow taking longitudinal resonances, propagation delay, reflections from terminations and discontinuities, etc., into account.

The lumped-parameter and one-dimensional waveguide models do not as such address explicitly the sound radiation problem. Sound radiation can be included as a simple source model (monopole or dipole as appropriate) or by some suitable analytical approximation, such as a flat piston in a baffle, at the end of a tube or free, or by more complex shapes, such as a cone or a dome.

Usually the transducer models are represented in the frequency domain. This corresponds to the practices of electric circuit modelling, and if propagation in lossy materials needs to be taken into account, then the theory is commonly formulated in the frequency domain. However, explicit time-domain descriptions can be very useful for more physics-based system description, for instance in describing nonlinear behaviour or in the analysis of an ultrasonic transducer.

The shift in transducer modelling has been towards more common use of direct numerical solution of the partial differential equations using finite element (FE), boundary element (BE), or more recently finite difference (FD) methods. This paradigm shift is quite simply due to increase in computing power and the amount of memory available. The numerical solutions are mostly based on the wave equation, but with the increasing computing capacity the application of the (linearized) Navier–Stokes equation would increase the reliability of the models for cavities, ports, etc. Table 1 summarizes the main properties of different modelling methods.

5 • Classical Principles

An electroacoustic transducer is typically described by two principal transforms: from electrical to mechanical and from mechanical to acoustical. Most types of transducers can be described through a unified description (Woollett 1966), where the nature of the transducer determines the coupling coefficient. The common practice is to use a gyrator to describe the electromechanical (or directly electroacoustic) conversion. A transformer could be equally well applied, with the slight theoretical hitch of having a transformer coefficient that is not dimensionless. The conversion of mechanical vibration to sound is typically represented by a transformer (Figure 2).

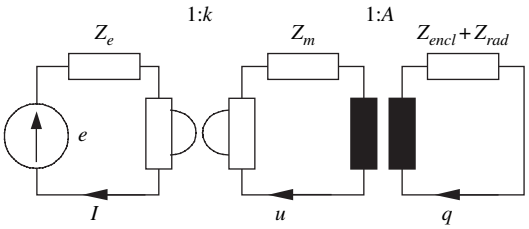


FIGURE 2 Basic equivalent circuit of a simple electroacoustic transducer; Z_e = electrical impedance, Z_m = mechanical impedance, Z_{encl} = enclosure impedance, Z_{rad} = radiation impedance, k = coupling coefficient, A = effective radiating area; e = driving source voltage, I = transducer current, u = transducer velocity, q = transducer volume velocity.

A more widely applicable form of this basic equivalent circuit can be realized by substituting a four-pole for one or more of the components of the circuit above. This allows a broader description of electromechanical transduction, mechano-acoustical conversion, and acoustical loading. Four-pole models are essential for, e.g., nonreciprocal transducers, and on the acoustical side for concise descriptions of any system that can be described as scalar inputs and outputs, such as simple or coupled-cavity loudspeakers, waveguides, or horns. The lumped-parameter and waveguide representations are illustrated by the example presented later for bass reflex loudspeakers.

The generic electroacoustic transducer can be represented by a four-pole illustrated below (Figure 3), where the quantities on the right-hand side can be chosen to be acoustical (volume velocity q and pressure p) or mechanical (velocity u and force F).

The input–output relationships of this four-pole can be described by

$$\begin{cases} e &= Z_I I + M_{ue} u \\ F &= M_{FI} I + z_l u \end{cases}$$

or alternatively in a matrix form as

$$\begin{bmatrix} e \\ F \end{bmatrix} = \begin{bmatrix} Z_I & M_{ue} \\ M_{FI} & z_l \end{bmatrix} \begin{bmatrix} I \\ u \end{bmatrix}.$$



FIGURE 3 Four-pole representation of transducer.

The matrix coefficients can be assigned a physical interpretation as follows: M_{ue} = transduction from velocity to voltage, M_{FI} = transduction from current to force, Z_I = electrical impedance of a transducer locked immobile, z_l = mechanical impedance of a transducer with open electrical circuit.

This representation can be applied to most simple transducers. Furthermore, if the transducer is passive (i.e. does not contain any energy sources) certain constraints can be placed on the relationship between the coefficients, and if the transducer is reciprocal (i.e. the transduction coefficient from electrical to acoustical/mechanical is the same as from mechanical/acoustical to electrical) the matrix is also symmetrical. It is possible to define multi-port representations for transducers with several inputs or outputs.

The importance of the four-pole representation is that the same formalism can be extended to mechanical–acoustical transform (i.e. sound radiation and reception) and purely acoustic subsystems (cavities, ports, horns, etc.; Figure 4). The matrix representation is comparable to the electromechanical four-poles, and also the physical interpretation of the matrix elements can be understood in the analogous manner.

An important special case of the acoustical four-pole is a homogeneous tube loaded with a known impedance (Figure 5). Besides the obvious simple cases, such a tube can be used as a building block for describing systems with either abrupt or gradual cross-section area or material property changes.

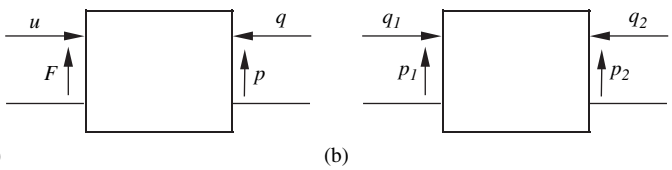


FIGURE 4 Input and output quantities of (**a**) a mechanical-acoustical and (**b**) an acoustical four-pole.

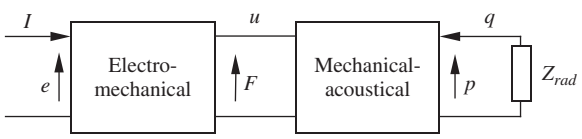


FIGURE 5 Building the simple equivalent circuit using four-poles.

For modelling purposes it is usually more practical to divide the four-pole into two separate entities, one representing the electromechanical transduction (like the voice coil and the magnet assembly of a loudspeaker) and another the mechanical–acoustical transformation (like the cone and the enclosure of a loudspeaker). Furthermore the mechanical and acoustical subsystems can be modelled as combinations of purely mechanical and acoustical four-poles yielding a more manageable detailed description of the system. Several simple systems, like low-frequency models of loudspeakers or microphones, can be quite easily described directly using lumped-component equivalent circuits without reverting to this kind of more generic approach. The four-pole approach, however, allows easy incorporation of both distributed and lumped components.

The choice of the fundamental input and output quantities to be examined depends on the usage situation. In a loudspeaker a suitable input quantity is voltage (although current could be argued to be a more basic input quantity for dynamic loudspeakers), and the suitable output quantity is the volume velocity, from which the far-field sound pressure is readily calculated. In ultrasonic applications the design target is often a certain force (i.e. pressure × area) at the application point. In microphones the usual choice for input quantity is the sound pressure or pressure gradient indirectly specified through pressure, frequency, and the distance from a sound source.

6 • Fundamentals of Transducers

6.1 Dynamic Transducers

The basic dynamic transducer has a conductor moving in a constant magnetic field, usually generated with permanent magnets. This principle appears in loudspeakers, microphones, high-level low-frequency underwater sound sources, and some mechanical actuators. Usually a dynamic transducer is built using a lengthy wire wound into a coil moving in a narrow air gap, but the simplest form, a straight short conductor moving in a transverse magnetic field, is occasionally encountered in ribbon microphones and loudspeakers and in the ultrasonic EMAT (electromagnetic–acoustic) transducers (Thomson 1990), where a wire is mounted directly on the surface of a solid substrate.

The classic description of the use of equivalent circuits for dynamic transducers was presented by Bauer (1976a). Modern dynamic transducer equivalent circuits have been discussed in a concise form by, e.g., Poldy (1994) and, together with other modern modelling tools, by Barnet and Opitz (2001).

6.2 Electromagnetic Transducers

Electromagnetic transducers, as opposed to dynamic transducers, rely on modulating the magnetic flux within a system where the magnetic flux goes through a variable air gap and the coil system is stationary. They were popular in the past in hearing aid microphones (and record player pickups) due to their sensitivity, but now they are found mostly in hearing aid earpieces. Their equivalent circuits resemble those of dynamic transducers; the only difference is in the form of the transduction coefficients.

6.3 Electrostatic Transducers

There are two categories of electrostatic transducers, benefiting from slightly different theoretical approaches. The one-sided transducers, having one fixed electrode and a moving diaphragm (like most condenser microphones, some headphones, and ultrasonic transmitters), rely as sources on the changing force due to the voltage changes and as a receiver on the voltage change of a constant-charge capacitor (carrier wave microphones are deliberately ignored here). The symmetrical transducers, like most electrostatic loudspeakers, where the moving diaphragm is between two fixed perforated electrodes, can be easiest understood as a variable charge moving in a constant electric field. Thus the symmetrical transducer is essentially linear, while the one-sided is essentially nonlinear.

6.4 Piezoelectric Transducers

In piezoelectric transducers both the external electrical field and the deformation are always essentially three-dimensional. A full description of the relationship between the electrical field acting on a piezoelectric device and the deformation or stress of the device therefore requires the use of a tensor description. However, as transducers are manufactured to have one direction of maximum sensitivity, in common applications the models can be simplified towards generic forms used for other transducer types. Characterisation of a piezoelectric transducer needs mechanical, electrical (dielectric), and transduction properties to be taken into account (Mason 1972).

7 • Dynamic Loudspeakers

The conventional dynamic loudspeaker has been one of the great successes of lumped-parameter modelling. Models succeed in predicting the low-frequency response (low frequency implying a frequency not much above system fundamental resonances and with wavelength large compared to dimensions, not a frequency actually regarded as subjectively low), but at higher frequencies they need to be complemented by more precise

enclosure acoustics, diaphragm behaviour, and external sound field models.

7.1 Driver Models

The simple description of a dynamic driver consists of the electric circuit (resistance and inductance), linear electromagnetic coupling, and a mass-spring-damper system for the vibrating system. This was the starting point for the classical analytic theory, and this simple model is a surprisingly strong predictor of loudspeaker low-frequency behaviour.

A critical success factor for models is an exact determination of the driver parameters. This cannot, unfortunately, be achieved with high precision from the mechanical dimensions, weight of the moving parts, etc., as there may be frequency dependence in the parameters, the effective radiating area is difficult to determine, etc. The parameters can be better determined by fitting a model to impedance measurements with a driver in free space and sealed enclosure, measurements in ported enclosures, simultaneous current and displacement measurements, or fitting current measurements to a system model (Seidel and Klippel 2001; Clark 1997).

Since the voice coil produces an eddy current with frequency-dependent penetration depth, a simple resistance–inductance circuit suggested by the coil structure is insufficient to model the voice coil impedance (Vanderkooy 1988). Optimal magnet assembly design requires the use of detailed numerical models (Bie 1992).

7.2 Direct-Radiating Loudspeakers, Sealed and Ported Enclosures

In the 1960s and 1970s the Australian scientists J. E. Benson (1996), A. Neville Thiele (1971a, 1971b), and Richard H. Small (1972, 1973a, 1973b, 1973c, 1973d, 1973e) utilized the topological equivalence between electric filters and the loudspeaker-equivalent circuits to take the desired low-frequency transfer function shape as a design starting point. This led quickly to adoption of key parameters used in driver characterisation (Thiele–Small parameters) and to design nomograms, tables, and calculation formulae essentially removing guesswork from the low-frequency design of conventional loudspeakers. Later the low-frequency modelling theory was extended by several authors towards numerical methods and more generally applicable to analytical models (Keele 1973, 1981) and analytical parameter synthesis approach (Rutt 1985). Since all lumped-parameter models of simple direct-radiating enclosures have a similar fundamental structure, they can be described employing a unified mathematical treatment (Bullock 1989).

7.3 Absorption Material Models

The absorption materials can be included in the models as time constants (Putland 1998) or as well suited for wageguide models, through specific impedance and complex propagation coefficient (describing speed of sound and propagation loss) (Delany and Bazley 1970; Miki 1990), utilizing a model for static flow resistance derived from material parameters (Sides et al. 1971). Also description as equivalent circuit in simple lumped-component enclosure models is possible (Leach 1989). The models mentioned here, unfortunately, lead to slightly inconsistent results due to their differing theoretical and experimental starting points.

7.4 Other Enclosure Types

Coupled-cavity loudspeakers (i.e. loudspeakers where the sound is radiated only through ports of resonators or passive radiators on one or both sides of the driver) exhibit typically strong coupling between the acoustical resonances and the driver resonance, thus necessitating the simultaneous modelling of the acoustics on both sides of the driver (Geddes 1989a; Fincham 1979). Although the equivalent circuits employed for deriving the basic theory usually describe ports as lumped components (inductors), it is highly recommended to use distributed (i.e. waveguide) in actual design models for the ports, since the longitudinal port modes have a significant effect on stop-band response even if the cavity modes would be controlled with the application of acoustic absorbents.

The waveguide approach is quite naturally appropriate for modelling of transmission line loudspeakers, where the additional low-frequency load and resonant boost is provided by the longitudinal modes of the enclosure air space (Backman 1992).

The open-baffle loudspeakers are from the modelling point of view much more challenging than their more enclosed counterparts. Describing their response needs a model for the edge diffraction (e.g. geometric diffraction models discussed below are appropriate), and the acoustical asymmetry of the driver needs to be taken into account in order to obtain the correct mid-range frequency response and polar pattern (Backman 1999a). Acoustical resistance enclosures, where the enclosure has a significant well-damped leak, can be approximated using lumped-parameter models (Backman 1999b).

7.5 Thermal Models

A particular feature of loudspeakers is the need for a dynamic thermal model for the voice coil. It is important for system reliability calculations and in predicting the compression and frequency response shape changes as a function of input level

(Eargle 1997, pp. 190–199; Button 1992; Klippel 2003). The parameters describing the thermal circuit need to be determined experimentally (Behler and Bernhard 1998). For other transducer types ambient temperature influence (in microphones, through thermal expansion and air viscosity changes, (Brüel & Kjær 1996); in piezoelectric devices through temperature dependence of coupling coefficients) can be taken into account as a static correction, but loudspeakers, due to their significant input power, have relatively rapid temperature variations affecting the electrical characteristics of the system. The time constants associated with these variations are, however, typically long as compared to the signal period, and thus the thermal effects are, from electroacoustic model point of view, quasi-static parameter changes (mostly change in voice coil resistance).

The thermal behaviour is described through its own equivalent circuit, but as the temperatures are measures of energy, not amplitude, the amplitude quantities of thermal models actually correspond to the temperature and thermal power. In thermal calculations sound radiation in direct-radiating loudspeakers is usually ignored as a minor correction term amounting only at most to a few percent of the total power (Figure 6).

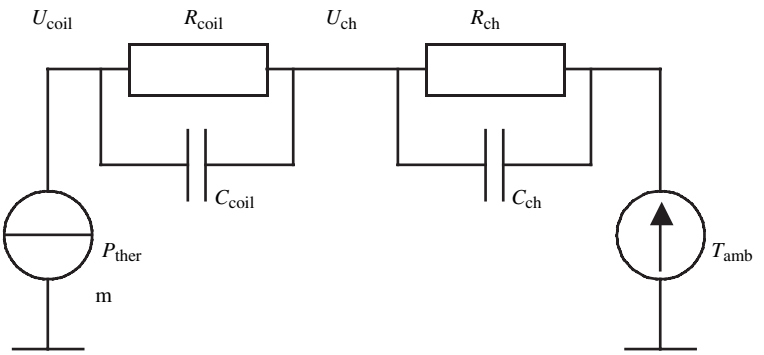


FIGURE 6 Simple thermal model of a loudspeaker describing the effects of the relatively rapid heating of the voice coil, the slower heating of chassis (including the magnet assembly), and the external temperature. P_{therm} = thermal input power (real power to the loudspeaker), R_{coil} = thermal loss of voice coil (radiation, conduction, etc.), C_{coil} = heat capacity of the voice coil, R_{ch} = thermal loss of the chassis (radiation, convection, conduction), C_{ch} = heat capacity of the frame, T_{amb} = ambient temperature inside the enclosure. The U_{coil} represents the voice coil temperature, and the voltage U_{ch} represents the chassis temperature as seen by the voice coil.

7.6 Nonlinear Models

Loudspeakers, especially woofers, are linear for only a very narrow part of their operating range. Thus nonlinear models are essential for both identifying the appropriate targets for transducer development and assessing the full system performance.

Nonlinear models have been developed for all types of transducers (Sherman and Butler 1995). Besides the nonlinearity in the transduction between the input signal and the exciting force the specific feature of loudspeakers is the nonlinearity of the passive mechanical structure also, which can be modelled through a direct numerical solution for the nonlinear differential equation (Olsen and Christiensen 1996). The direct solution is rather tedious, and thus there has been interest in developing computationally more efficient methods, such as the use of Volterra series (Kaizer 1987), which unfortunately suffer from the great difficulty in obtaining the needed kernel functions experimentally. In general, the determination of nonlinearity parameters is the critical factor in model selection, as all the computational methods provide essentially correct descriptions of nonlinearity within their range of validity (in terms of displacement and frequency). When the nonlinearity is described in terms of displacement-dependent coupling coefficient, compliance, and inductance, each of these can be approximated using a low-order polynomial (Geddes and Phillips 1991; Clark and Michelich 1999; Klippel 1999).

The enclosure itself also exhibits some nonlinearity. As the change in the internal volume is not negligible, the compliance (as a simple approximation) changes inversely proportional to the total volume. The phase shift due to damping material losses (and nonlinearity mechanisms of the damping material itself) complicates the practical situation somehow. This nonlinear compliance can be incorporated easily into the nonlinear suspension model. A more problematic aspect, both from modelling and practical engineering point of view, is the behaviour of the reflex port. There are some quasi-empirical formulae and empirical results, but predicting both the onset level and seriousness of the nonlinearity is difficult (Roozen et al. 1998).

7.7 Radiation Models

Although the analysis of sound radiation could be fundamentally regarded as a single computational effort, it is usually beneficial, from both application and computation point of view, to divide the problem into the following subtopics:

- Radiation load on the transducer (seen as Z_{rad} in the equivalent circuits); this can be extended to describe the coupling between multiple sound sources
- Sound radiation from the transducer itself

- Diffraction effects on the transducer housing/enclosure
- Radiation from, and through, enclosure surfaces, necessitating a numerical vibroacoustics model

7.7.1 Radiation Impedance

Analytical descriptions for a simple piston in an infinite baffle have been well discussed in the classic theoretical acoustics literature. Also a circular piston at the end of an infinite tube (Levine and Schwinger 1948) and a rectangular piston in an infinite baffle (Bank and Wright 1990) can be described analytically. The radiation load can often be approximated by a simple two-component equivalent circuit (Bauer 1976b). For direct-radiating loudspeakers and reflex ports it is usually sufficient to consider only the reactive component ("added mass"). Relatively simple approximations can be derived for the interaction of two pistons in an infinite baffle (Jacobsen 1976). A more precise model of the radiation impedance needs also taking into account the backscattering of the sound from the transducer housing, and thus diffraction models must be employed.

7.7.2 Diaphragm Radiation

A simple, but useful, approximation is to describe an actual sound source as a rigid piston. This model yields an estimate for the displacement needed for a certain sound pressure at low frequencies and defines a lower limit for the frequency at which the source (or receiver) becomes directive. The rigid-piston model is an excellent model for longitudinal ultrasonic transducers coupled to a fluid, as the transducer is excited very homogenously; so even if the dimensions of the transducer would allow for higher modes, the piston mode remains dominant.

Due to the finite bending wave propagation speed the rigid-piston model will fail for loudspeaker diaphragms at higher frequencies. Typically, the model inaccuracy will take effect around the same frequency range the driver in general becomes directive. Thus, rather than regarding the loudspeaker diaphragm as a rigid structure, a better conceptual model is to consider the sound radiation to be due to bending waves propagating on the diaphragm surface (Frankfort 1978). The propagating wave concept has been applied with reasonable success to driver and system modelling (Murphy 1993), but for development purposes the use of element models is recommended, as the pseudo-analytical expressions for local bending stiffness needed for the travelling wave model can be derived only for very simple geometries (flat, straight cone), and radial modes are difficult to include.

7.7.3 Enclosure Diffraction

In modern high-quality loudspeakers the major cause for mid-range and high-frequency colouration can be enclosure diffraction. For accuracy and computational efficiency the

best general-purpose approach is the time-domain formulation of geometric diffraction theory (Svensson et al. 1999). This approach, however, has geometric limitations, e.g. curved surfaces necessitate the use of more generic numerical models.

7.8 Distributed-Mode Loudspeakers

As opposed to conventional loudspeakers, where higher vibration modes are avoided, distributed-mode loudspeakers use a rigid plate that is excited to produce a high number of vibration modes, which are designed to have reasonably even distribution as a function of frequency. Despite this complexity, a good qualitative description of their function can be obtained using a combination of statistical and lumped-parameter models (Harris and Hawksford 2000; Panzer and Harris 1998; Panzer and Kavanagh 1999).

7.9 Full Numerical Enclosure Acoustic and Structural Models

Although the models based on analytic theory have the advantage of being computationally exceptionally fast, all of them contain built-in constraints and approximations for the geometry of the system. Comparisons of numerical models, analytical models, and measured results indicate that although simple geometries are described with reasonable accuracy by analytical models (Ih 1991), more realistic systems need numerical approach (Sakai et al. 1984), especially when fluid–structure interaction needs to be taken into account (Karjalainen et al. 2001). The need of analysing the fluid–structure interaction stems from the unfortunate coincidence that the speed of the bending waves in typical enclosures has the same order of magnitude as the sound.

Microphones and miniature loudspeakers (by which in this context we mean audio frequency loudspeakers with maximum dimensions of a few centimetres) would appear excellent candidates for a lumped-parameter description, as they are much smaller than the wavelength, but also their analysis can benefit from finite-element approach. One reason for partial failure of lumped-parameter models is that in small devices distinguishing between clearly defined cavities, ports, etc., is often difficult, and describing losses call for precise geometrical details.

8 • Horn Loudspeakers

Horn models are equally well needed for audio and the solid horns in ultrasonic applications. Classic plane-wave horn theory is quite similar for both fluid-filled (audio) horns and solid horns, but the differences arise in solid transverse waveguides, like those

applied in some ultrasonic tools. They need a fundamentally different approach due to the dispersive propagation, since there is no general approach corresponding to the horn equation for the fourth-order partial differential equation describing the propagation of the bending waves.

8.1 Classic Theory

It can be shown that the wave propagation in any system where the system can be described with one space variable obeys the horn equation (Putland 1993). The classic horn theory stems from the fortunate fact that there are families of closed-form analytical solutions for the horn equation for some of the horn shapes with practical importance (Salmon 1998). The limits of the classic theory have been discussed extensively by Geddes (1989b, 1993), who introduced curvilinear coordinate techniques for a more general class of horns and also proved that a fully generalized analytical solution is not feasible. However, Holland et al. (1991) have experimentally shown that with reasonable limitations a horn with a straight axis can be modelled as consecutive simple blocks. A very useful four-pole representation of the Salmon family of solutions was developed by Leach (1996). Arbitrary profiles not included in the analytical solutions are commonly constructed with connected straight (or conical) sections approximating the actual horn profile. The classic theory, however, is unable to predict horn polar patterns and high-frequency behaviour, for which numerical models (FEM, BEM) are needed.

8.2 Nonlinear Models

Horn drivers share the electromechanical transduction nonlinearity mechanisms with direct-radiating drivers. The nonlinear aspects of horn and driver acoustics, on the other hand, have a more direct link to the fundamental physical concepts. The fundamental theory of horn driver nonlinearity has been discussed in great detail by Voishvillo (2002, 2003), who is to be credited for discussing clearly the fundamental distinction between nonlinear effects (e.g. the inverse relationship between air pressure and density) and parametric effects (e.g. modulation of acoustical parameters due to changes in system dimensions with the movement of loudspeaker diaphragm).

In large audio systems at high power levels also the propagation of sound over large distances contributes to the overall distortion. This is mainly due to the nonlinear asymmetric steepening of the propagating wave (Czerwinski et al. 2000). (The nonlinear propagation can also be utilized to produce very directional audio frequency sound from modulated ultrasound in both air and underwater applications.)

9 • Electrostatic Loudspeakers

The analysis of electrostatic loudspeaker behaviour needs attention to fundamentally different aspects as compared to dynamic loudspeakers. Besides the transduction principle the fundamental differences are

- Diaphragm moving mass is close to negligible as compared to the radiation load.
- Listener is often in the near field.
- Diaphragm is in an open baffle.

Due to the fundamentally linear nature of transduction and the small mechanical impedance it is relatively easy to come up with a quasi-analytical description covering the basic behaviour even in greater detail than is possible for dynamic loudspeakers (Baxandall 1994). The near-field radiation aspect can be analysed using generally applicable approaches derived for conventional transducers.

10 • Microphones

10.1 General Features of Microphone Models

Modelling of small-scale acoustics needs closer attention to viscosity and heat conduction than the relatively large structures found in loudspeakers (Maisano 1994; Stinson 1991). The critical parameter determining the smallness of the structure is not here the ratio of the dimensions to the wavelength, but to the thickness of the boundary layer close to the surface, within which both viscosity and heat conduction play significant part.

Microphone models need an accurate description of acoustical resistance. The acoustical resistance can be in the form of a thin air layer (as in condenser microphones) or separate acoustical resistance elements (as often in dynamic microphones). These structures are the greatest challenge for accurate microphone theory. The difficulty of small-scale acoustics comes partially from the coupled nature of the problem and partially from the difficulty of characterizing the acoustical resistance elements in a satisfying manner. The geometry of even the simplest woven meshes is very complex for an accurate model of the flow, let alone more random porous materials, especially if the structural compliance of the materials needs to be taken into account. Thus, acoustical resistance materials are usually best described through quasi-experimental numerical models, although simplified theoretical models lend very usable insight into their design. The model of a thin flat air layer between the diaphragm and the condenser microphone back plate is

more feasible due to the precisely definable geometry. Even then the solution needs some simplifying approximations, but the accuracy can already be very high.

The noise sources characteristic to microphones include thermal acoustical noise in the air and the mechanical damping of the diaphragm structure itself, thermal electrical noise in the resistive parts of the microphone electrical circuit, and the noise of the possible semiconductor devices. The acoustically and mechanically generated noise does not need any addition to the model, but it is calculated simply by evaluating the thermal noise in the resistances of the equivalent circuit (Olson 1972). The effect of external air load can be taken into account using the simple equivalent circuit mentioned earlier for loudspeaker radiation or a more detailed model (Leach 1990).

10.2 Condenser Microphones

The relatively accurate theory that can be derived for condenser microphones has enabled their development into trustworthy precision measurement instruments. The main features of condenser microphone operation can be described by a relatively simple model found in most electroacoustics textbooks, but describing the behaviour of the microphone diaphragm with the accuracy needed for development purposes requires several factors to be taken into account: diaphragm tension, electrostatic forces acting on the diaphragm, compliance of the air space behind the diaphragm, and the damping caused by the air layer between the diaphragm and the back electrode, and in the holes of the back electrode. The computational models have been summarized (with extensive references) by Hawley et al. (1995) and Zuckerwar (1995).

It should be remembered that one of the key achievements of transducer theory was the early correct model for condenser microphones (Wente 1922). This had very profound implication on the overall development of experimental acoustics, since it enabled the independent determination of the microphone performance parameters from basic physical quantities, in turn enabling the early development of precision acoustical measurements. Another fruitful relationship between the fundamental transducer theory and precision measurements led to the theory of reciprocal calibration, which still is the primary precision transducer calibration method for microphones and hydrophones.

10.3 Dynamic Microphones

Conventional dynamic microphones appear to some extent similar to dynamic loudspeakers. The essential difference is that the models are intended to cover the full operating frequency

range instead of just describing the low-frequency behaviour. Thus relatively complex acoustical structures are needed to maintain a flat frequency response across the entire operating range (Rossi 1988). The problem of modelling dynamic miniature loudspeakers and earpieces is to some extent similar to the microphones.

11 • Piezoelectric Transducers for Ultrasonic and Underwater Applications

The feature that is common to ultrasonic and underwater transducers, and which separates them from conventional loudspeakers and microphones, is the focus on impedance matching between the transducer and the acoustical load. As opposed to audio transducers, for which there are a few well-defined typical equivalent circuit descriptions, the models for the ultrasonic transducers need to be selected according to the application. The transducers (especially transmitters) are quite often employed as narrow-band resonant devices, which is reflected in the choice of the equivalent circuit representation and allows a significant simplification of the description around the principal (usually the lowest) resonance, but wide-band models require the thickness of the transducers to be taken into account (Hickson and Busch-Vishniac 1997).

References

AES Anthology: Loudspeakers vol. 1, Audio Engineering Society, New York, 1978a, 448 pp.

AES Anthology: Sound Reinforcement, Audio Engineering Society, New York, 1978b, 339 pp.

AES Anthology: Microphones, Audio Engineering Society, New York, 1978c, 392 pp.

AES Anthology: Loudspeakers vol. 2, Audio Engineering Society, New York, 1983, 464 pp.

AES Anthology: Loudspeakers vol. 3, Audio Engineering Society, New York, 1991a, 456 pp.

AES Anthology: Loudspeakers vol. 4, Audio Engineering Society, New York, 1991b, 496 pp.

AES Anthology: Sound Reinforcement, vol. 2, Audio Engineering Society, New York, 1996, 496 pp.

Backman, J.: A Computational Model of Transmission Line Loudspeakers. 92nd Convention of the Audio Eng. Soc., March 1992, Preprint 3326.

Backman, J.: A Model of Open-Baffle Loudspeakers. 107th Convention of the Audio Eng. Soc., September 1999a, Preprint 5025.

Backman, J.: Theory of Acoustical Resistance Enclosures. 106th Convention of the Audio Eng. Soc., May 1999b, Preprint 4979.

Bank, G.; Wright, J. R.: Radiation Impedance Calculations for a Rectangular Piston. J. Audio Eng. Soc., vol. 38, no. 5, May 1990, pp. 350–354.

Barnet, R.; Opitz, M.: Modern Development Tools for Dynamic Transducers. 111th Convention of the Audio Eng. Soc., December 2001, Preprint 5438.

Bauer, B. B.: Equivalent Circuit Analysis of Mechano-acoustic Structures (Reprint). J. Audio Eng. Soc., vol. 24, no. 8, October 1976a, pp. 643–652 (originally in Transactions of IRE, vol. AU-2, July–August 1954, pp. 112–120).

Bauer, B. B.: On the Equivalent Circuit of a Plane Wave Confronting an Acoustical Device. J. Audio Eng. Soc., vol. 24, no. 8, October 1976b, pp. 653–655.

Baxandall, P. J.: Electrostatic Loudspeakers, in Borwick, J. (ed.): Loudspeaker and Headphone Handbook, 2nd ed., Focal Press, Oxford, 1994.

Behler, G. K.; Bernhard, A.: Measuring Methods to Derive the Lumped Elements of the Loudspeaker Thermal Equivalent Circuit. 104th Convention of the Audio Eng. Soc., May 1998, Preprint 4744.

Benson, J. E.: Theory & Design of Loudspeaker Enclosures, Prompt Publications (Howard W. Sams & Company), Indianapolis, 1996, 244 pp. (reprints of Benson's articles from 1960s).

Beranek, L. L.: Loudspeakers and Microphones. J. Acoust. Soc. Am., vol. 26, no. 5, September 1954 (Twenty-Fifth Anniversary Celebration Issue), pp. 618–629.

Bie, D.: Design and Theory of a New Midrange Horn Driver. 93rd Convention of the Audio Eng. Soc., October 1992, Preprint 3429.

Borwick, J. (ed.): Loudspeaker and Headphone Handbook, 2nd ed., Focal Press, Oxford, 1994, 601 pp.

Brüel & Kjær: Microphone Handbook, Brüel & Kjær, Naerum, Denmark, 1996, pp. 3-18–3-19.

Bullock III, R. M.: A Unified Model for Closed-Boxes, Vented-Boxes, and Passive-Radiators with Losses. 87th Convention of the Audio Eng. Soc., October 1989, Preprint 2841.

Bullock, R. M.; White, R.: Bullock on Boxes. Old Colony Sound Laboratory, 1991, 72 p.

Butterworth, S.: On Electrically Maintained Vibrations. Proc. Phys. Soc. (London), vol. 26, 1915, p. 264; vol. 27, 1915, p. 410.

Button, D. J.: Heat Dissipation and Power Compression in Loudspeakers. J. Audio Eng. Soc., vol. 40, no. 1/2, January 1992, pp. 32–41.

Clark, D.: Precision Measurement of Loudspeaker Parameters. J. Audio Eng. Soc., vol. 45, no. 3, March 1997, pp. 129–141.

Clark, D.; Michelich, R. J.: Modeling and Controlling Excursion-Related Distortion in Loudspeakers. 106th Convention of the Audio Eng. Soc., May 1999, Preprint 4862.

Czerwinski, E.; Voishvillo, A.; Alexandrov, S.; Terekhov, A.: Propagation Distortion in Sound Systems – Can We Avoid It? J. Audio Eng. Soc., vol. 48, no. 1/2, January/February 2000, pp. 30–48.

Delany, M. E.; Bazley, E. N.: Acoustical Properties of Fibrous Absorbent Materials. Appl. Acoust., vol. 3, 1970, pp. 105–116.
Dickason, V.: The Loudspeaker Design Cookbook, 6th ed., Audio Amateur Press, Peterborough, 2000, 233 pp.
Earlgle, J. M.: Loudspeaker Handbook, Chapman & Hall, New York, 1997, 325 pp.
Fincham, L. R.: A Bandpass Loudspeaker Enclosure. 63rd Convention of the Audio Eng. Soc., May 1979, Preprint 1512.
Fischer, F. A.: Fundamentals of Electroacoustics, Interscience, New York, 1955, 186 pp. (Originally Grundzüge der Elektroakustik, Fachverlag Scheile & Schön, Berlin, 1950).
Frankfort, F. J. M.: Vibration Patterns and Radiation Behavior of Loudspeaker Cones. J. Audio Eng. Soc., vol. 26, no. 9, September 1978, pp. 609–622.
Gander, M. R.: Fifty Years of Loudspeaker Developments as Viewed Through the Perspective of the Audio Engineering Society. J. Audio Eng. Soc., vol. 46, no. 1/2, January 1998, pp. 43–58.
Gayford, M. (ed.): Microphone Engineering Handbook, Focal Press, Oxford, 1994, 449 pp.
Geddes, E. R.: An Introduction to Band-Pass Loudspeaker Systems. J. Audio Eng. Soc., vol. 37, no. 5, May 1989a, pp. 308–342.
Geddes, E. R.: Acoustic Waveguide Theory. J. Audio Eng. Soc., vol. 37, no. 7/8, July 1989b, pp. 554–569.
Geddes, E. R.: Acoustic Waveguide Theory Revisited. J. Audio Eng. Soc., vol. 41, no. 6, June 1993, pp. 452–461.
Geddes, E.; Lee, L.: Audio Transducers, GedLee Associates, 2002, 304 pp.
Geddes, E.; Phillips, A.: Efficient Loudspeaker Linear and Nonlinear Parameter Estimation. 91st Convention of the Audio Eng. Soc., October 1991, Preprint 3164.
Groves Jr., I. D.: Acoustic Transducers, Benchmark Papers in Acoustics, vol. 14, Hutchinsorn Ross Publishing Company, Stoudsburg, 1981, 391 pp.
Harris, N. J.; Hawksford, M. O.: Introduction to Distributed Mode Loudspeakers (DML) with First-Order Behavioural Modelling. IEE Proc. Circuits Devices Syst., vol. 147, no. 3, June 2000, pp. 153–157.
Hawley, M. S.; Romanov, F. F.; Warren, J. E.: The Western Electric 640AA Capacitance Microphone: Its History and Theory of Operation, in Wong, G. S. K.; Embleton, T. F. W. (eds.): AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1995, pp. 8–34.
Hickson, E. L.; Busch-Vishniac, I. J.: Transducer Principles, in Crocker, M. J. (ed.): Encyclopedia of Acoustics, vol. 4, Wiley, New York, 1997, pp. 1889–1902.
Holland, K. R.; Fahy, F. J.; Morfey, C. L.: Prediction and Measurement of the One-Parameter Behavior of Horns. J. Audio Eng. Soc., vol. 39, no. 5, May 1991, pp. 315–337.
Hueter, T. F.; Bolt, R. H.: Sonics, Wiley, New York, 1955, pp. 109–118, 180–186.

Hunt, F. V: Electroacoustics, 1954 (reprinted by the Acoustical Society of America, 1982).

Ih, J.-G.: Acoustic Wave Action Inside Rectangular Loudspeaker Cabinets. J. Audio Eng. Soc., vol. 39, no. 12, December 1991, pp. 945–955.

Jacobsen, O.: Some Aspects of the Self and Mutual Radiation Impedance Concept with Respect to Loudspeakers. J. Audio Eng. Soc., vol. 24, no. 2, March 1976, pp. 82–92.

Kaizer, A. J. M.: Modeling of the Nonlinear Response of an Electrodynamic Loudspeaker by a Volterra Series Expansion. J. Audio Eng. Soc., vol. 35, no. 6, June 1987, pp. 421–433.

Karjalainen, M.; Ikonen, V.; Antsalo, P.; Maijala, P.; Savioja, L.; Suutala, A.; Pohjolainen, S.: Comparison of Numerical Simulation Models and Measured Low-Frequency Behavior of Loudspeaker Enclosures. J. Audio Eng. Soc., vol. 49, no. 12, December 2001, pp. 1148–1166.

Keele Jr., D. B.: Sensitivity of Thiele's Vented Loudspeaker Enclosure Alignments to Parameter Variations. J. Audio Eng. Soc., vol. 21, no. 4, May 1973, pp. 246–255.

Keele Jr., D. B.: Direct Low-Frequency Driver Synthesis from System Specifications. 69th Convention of the Audio Eng. Soc., May 1981, Preprint 1797.

Kennelly, A. E.; Pierce, G. W.: The Impedance of Telephone Receivers as Affected by the Motion of Their Diaphragms. Proc. Am. Acad. Arts Sci., vol. 48, 1912, pp. 113–151. Reprinted in Groves, 1981, pp. 328–352.

Klippel, W.: Measurement of Large-Signal Parameters of Electrodynamic Transducer. 107th Convention of the Audio Eng. Soc., September 1999, Preprint 5008.

Klippel, W.: Nonlinear Modeling of the Heat Transfer in Loudspeakers. 114th Convention of the Audio Eng. Soc., March 2003, Preprint 5733.

Leach, W. M.: Electroacoustic-Analogous Circuit Models for Filled Enclosures. J. Audio Eng. Soc., vol. 37, no. 7/8, July 1989, pp. 586–592.

Leach, M. W.: On the Electroacoustic-Analogous Circuit for a Plane Wave Incident on the Diaphragm of a Free-Field Pressure Microphone. J. Audio Eng. Soc., vol. 38, no. 7/8, July 1990, pp. 566–568.

Leach Jr., M. W.: A Two-Port Analogous Circuit and SPICE Model for Salmon's Family of Acoustic Horns. J. Acoust. Soc. Am., vol. 99, no. 3, March 1996, pp. 1459–1464.

Levine, H.; Schwinger, J.: On the Radiation of Sound from an Unflanged Circular Pipe. Phys. Rev., vol. 73, no. 4, February 1948, pp. 383–406.

Maisano, J.: An Improved Transmisson Line Model for Visco-thermal Lossy Sound Propagation. 96th Convention of the Audio Eng. Soc., February 1994, Preprint 3828,

Mason, W. P.: Electromechanical Transducers and Wave Filters, D. Van Nostrand Company, New York, 1942, 333 pp.

Mason, W. P. (ed.): Physical Acoustics, vol. IA, Academic Press, New York, 1964a, 515 pp.

Mason, W. P. (ed): Physical Acoustics, vol. IB, Academic Press, New York, 1964b, 376 pp.

Mason, W. P.: Properties of Transducer Materials, in American Institute of Physics Handbook, 3rd ed., McGraw-Hill, New York, 1972, pp. 3-118–3-129.
Mason, W. P.; Thurston, R. N. (eds.): Physical Acoustics, vol. X, Academic Press, New York, 1973, 403 pp.
Mason, W. P.; Thurston, R. N. (eds.): Physical Acoustics, vol. XIV, Academic Press, New York, 1979, 561 pp.
Mason, W. P.; Thurston, R. N. (eds.): Physical Acoustics, vol. XV, Academic Press, New York, 1981, 382 pp.
McLachlan, N. W.: Elements of Loud Speaker Practice, Oxford University Press, London, 1935, 160 pp.
Miki, Y.: Acoustical Properties of Porous Materials – Modifications of Delany-Bazley Models. J. Acoust. Soc. Jpn (E), vol. 11, no. 1, 1990, pp. 19–24.
Murphy, D. J.: Axisymmetric Model of a Moving-Coil Loudspeaker. J. Audio Eng. Soc., vol. 41, no. 9, September 1993, pp. 679–690.
Olsen, E. S.; Christiensen, K. B.: Nonlinear Modeling of Low-Frequency Loudspeakers: A More Complete Model. 100th Convention of the Audio Eng. Soc., May 1996, Preprint 4205.
Olson, H. F.: Dynamic Analogies, D. Van Nostrand Company, New York, 1943 (2nd ed., 1958, 278 pp.).
Olson, H. F.: Microphone Thermal Agitation Noise. J. Acoust. Soc. Am., vol. 51, no. 2 (pt 1), 1972, pp. 425–432.
Panzer, J. W.; Harris, N.: Distributed-Mode Loudspeaker Simulation Model. 104th Convention of the Audio Eng. Soc., May 1998, Preprint 4739.
Panzer, J.; Kavanagh, S.: Modal Network Solver for the Simulation of Complex Mechanoacoustical Systems. 107th Convention of the Audio Eng. Soc., September 1999, Preprint 5022.
Poldy, C. A.: Headphones, in Borwick, J. (ed.): Loudspeaker and Headphone Handbook, 2nd ed., Focal Press, Oxford, 1994, pp. 495–500.
Putland, G. R.: Every One-Parameter Acoustic Field Obeys Webster's Horn Equation. J. Audio Eng. Soc., vol. 41, no. 6, June 1993, pp. 435–451.
Putland, G. R.: Thermal Time Constants and Dynamic Compressibility of Air in Fiber-Filled Loudspeaker Enclosures. J. Audio Eng. Soc., vol. 46, no. 3, March 1998, pp. 139–151.
Roozen, N. B.; Bockhots, M.; van Eck, P.; Hirschberg, A.: Vortex Sound in Bass-Reflex Ports of Loudspeakers. J. Acoust. Soc. Am., vol. 104, no. 4, October 1998, pp. 1914–1918 (Part I), 1919–1928 (Part II).
Rossi, M.: Acoustics and Electroacoustics, Artech House, Norwood, 1988, pp. 518–536.
Rutt, T. E.: Root-Locus Technique for Vented-Box Loudspeaker Design. J. Audio Eng. Soc., vol. 33, September 1985, pp. 659–668.
Sakai, S.; Kagawa, Y.; Yamabuchi, T.: Acoustic Field in an Enclosure and Its Effect on Sound-Pressure Responses of a Loudspeaker. J. Audio Eng. Soc., vol. 32, no. 4, April 1984, pp. 218–227.

Salmon, V.: Horns, in Crocker, M. J. (ed.): Encyclopedia of Acoustics, vol. 4 Wiley Interscience, 1998, pp. 1925–1931.
Seidel, U.; Klippel, W.: Fast and Accurate Measurement of the Linear Transducer Parameters. 110th Convention of the Audio Eng. Soc., May 2001, Preprint 5308.
Sherman, C. H.; Butler, J. L.: Analysis of Harmonic Distortion in Electroacoustic Transducers. J. Acoust. Soc. Am., vol. 98, no. 3, September 1995, pp. 1596–1611.
Sides, D. J.; Attenborough, K.; Mulholland, K. A.: Application of a Generalized Acoustic Propagation Theory to Fibrous Absorbents. J. Sound Vibr., vol. 19, no. 1, 1971, pp. 49–64.
Small, R. H.: Closed-Box Loudspeaker Systems Part I: Analysis. J. Audio Eng. Soc., vol. 20, no. 10, December 1972, pp. 798–808.
Small, R. H.: Closed-Box Loudspeaker Systems Part II: Synthesis. J. Audio Eng. Soc., vol. 21, no. 1, January/February 1973a, pp. 11–18.
Small, R. H.: Vented-Box Loudspeaker Systems, Part I: Small-Signal Analysis. J. Audio Eng. Soc., vol. 21, no. 5, June 1973b, pp. 363–372.
Small, R. H.: Vented-Box Loudspeaker Systems, Part II: Large-Signal Analysis. J. Audio Eng. Soc., vol. 21, no. 6, July/August 1973c, pp. 438–444.
Small, R. H.: Vented-Box Loudspeaker Systems, Part III: Synthesis. J. Audio Eng. Soc., vol. 21, no. 7, September 1973d, pp. 549–554.
Small, R. H.: Vented-Box Loudspeaker Systems, Part IV: Appendices. J. Audio Eng. Soc., vol. 21, no. 8, October 1973e, pp. 635–639.
Stinson, M. R.: The Propagation of Plane Sound Waves in Narrow and Wide Circular Tubes, and Generalization to Uniform Tubes of Arbitrary Cross-Sectional Shape. J. Acoust. Soc. Am., vol. 89, no. 2, 1991, pp. 550–558.
Svensson, U. P.; Fred, R. I.; Vanderkooy, J.: An analytic secondary source model of edge diffraction impulse responses. J. Acoust. Soc. Am., vol. 106, no. 5, November 1999, pp. 2331–2344.
Thiele, A. N.: Loudspeakers in Vented Boxes, Part I. J. Audio Eng. Soc., vol. 19, no. 5, May 1971a, pp. 382–392.
Thiele, A. N.: Loudspeakers in Vented Boxes, Part II. J. Audio Eng. Soc., vol. 19, no. 6, June 1971b, pp. 471–483.
Thomson, R. B.: Physical Principles of Measurements with EMAT Transducers, in Thurston and Pierce 1990, pp. 157–200.
Thurston, R. N., Pierce, A. D. (eds.): Physical Acoustics, vol. XIX, Ultrasonic Measurement Methods, Academic Press, New York, 1990.
Vanderkooy, J.: A Model of Loudspeaker Driver Impedance Incorporating Eddy Currents in the Pole Structure. 84th Convention of the Audio Eng. Soc., Paris, March 1988, Preprint 2619.
Voishvillo, A.: Nonlinearity in Horn Drivers – Where the Distortion Comes From? 113th Convention of the Audio Eng. Soc., October 2002, Preprint 5641.
Voishvillo, A.: Nonlinear Versus Parametric Effects in Compression Drivers. 115th Convention of the Audio Eng. Soc., October 2003, Preprint 5912.

Wente, E. C.: The Sensitivity and Precision of the Electrostatic Transmitter for Measuring Sound Intensities. Phys. Rev., vol. 19, 1922, pp. 498–503.

Wong, G. S. K.; Embleton, T. F. W. (eds.): AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1995, 321 pp.

Woollett, R. S.: Effective Coupling Factor of Single-Degree-of-Freedom Transducer. J. Acoust. Soc. Am., vol. 40. 1966, pp. 1112–1123.

Zuckerwar, A. J.: Principles of Operation of Condenser Microphones, in Wong, G. S. K.; Embleton, T. F. W. (eds.): AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1995, pp. 37–69.

33

Loudspeaker Design and Performance Evaluation

Aki Vihtori Mäkivirta

Genelec Oy, Iisalmi, Finland

1 • Loudspeaker as a System Design Problem

A loudspeaker comprises transducers converting an electrical driving signal into sound pressure, an enclosure working as a holder for transducers, front baffle and box to contain and eliminate the rear-radiating audio signal, and electronic components. Modeling of transducers as well as enclosures is treated in Chap. 32 of this handbook. The purpose of the present chapter is to shed light on the design choices and options for the electronic circuits conditioning the electrical signal fed into loudspeaker transducers in order to optimize the acoustic performance of the loudspeaker.

Loudspeakers with wide bandwidth, flat and smooth magnitude response, and uniformly wide directivity are generally rated high by experienced listeners (Olive, 2003). This type of response has become the widely accepted design target in the industry, although there are also examples of highly directive designs mainly for consumer use. To achieve such high quality, the audio signal is split in two, three, or more frequency bands by *crossover filters*, and separate transducers are used for low frequencies (called "woofer"), high frequencies ("tweeter"), and possibly for mid-frequencies ("midrange").

Loudspeakers come in two flavors – passive and active. A *passive loudspeaker* contains transducers, an enclosure, and

crossover filters. Crossover filters are built by passive electrical components to divide the signal frequency components from a power amplifier to transducers (see Figure 1a).

In professional applications the passive principle is used less and less because the active loudspeaker concept offers many significant benefits. The response of a passive loudspeaker is influenced by the output impedance of the driving amplifier, modified by the feeding cable. These factors are not a concern with active loudspeakers.

An *active loudspeaker* contains power amplifiers directly connected to transducers with filtering and signal conditioning electronics placed before the power amplifiers (see Figure 1b). The signal conditioning electronics include crossover filtering, system equalization, and protection. Modern active

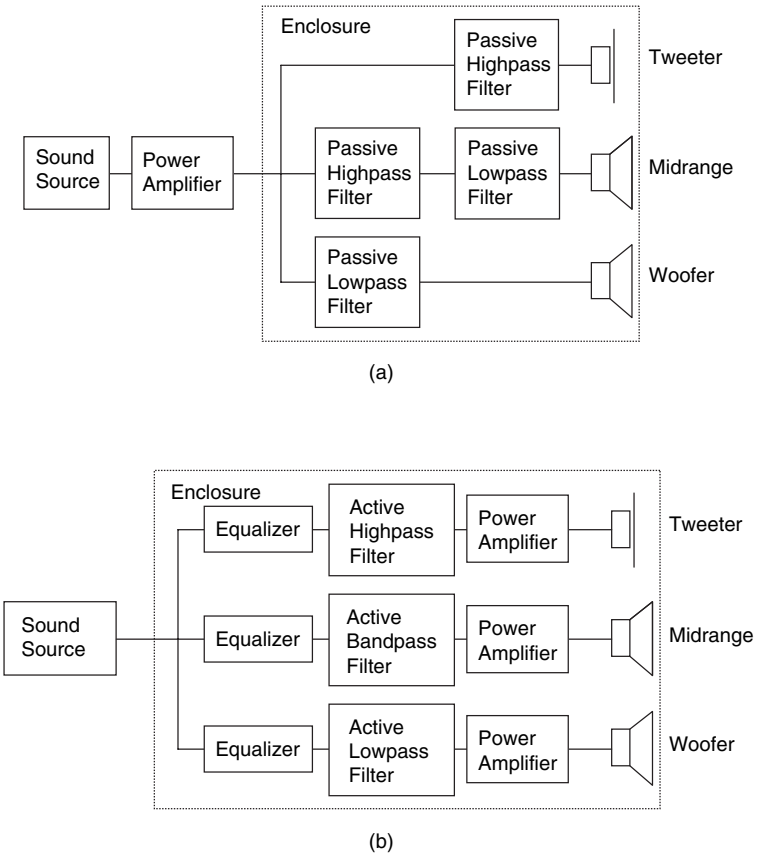


FIGURE 1 Fundamental functional blocks of a loudspeaker: (**a**) passive loudspeaker and (**b**) active loudspeaker.

loudspeakers also provide adjustable room equalization to optimize the frequency response in the final room installation. Active loudspeakers are built using both analogue and digital signal processing techniques. Because power amplifiers determine the source impedances seen by transducers, the source impedances can be carefully designed and maintained in an active loudspeaker. Precise circuitry to protect transducers can be implemented. Active loudspeakers can use high-order crossover filtering and precise equalization without impacting the source impedance seen by transducers.

2 • Crossover Filters

The basic crossover filter consists of a complementary filter pair of a lowpass filter $H_{LP}(s)$ and a highpass filter $H_{HP}(s)$. The frequency band where the crossing over between the two outputs occurs can be loosely defined as the frequency range where the two outputs are both contributing to form the acoustic output signal. Both outputs are contributing at equal level at the *crossover frequency*. In a modern crossover filter, the complex sum $H(s)$ of the filter outputs results in a flat magnitude response on the *acoustic axis* of the loudspeaker. The actual complex summation occurs in the air, at the position of the listener's ear. Other types of crossover alignments are used when the power output of a loudspeaker is optimized to be flat (see, e.g., Garde, 1980 for an overview of traditional alignments).

The requirements for crossover filter are characterized by (a) a defined flatness of the response (complex sum output) across the audio band, (b) complementary magnitude responses across the crossover transition band and (c) matching phase responses for the two filter outputs:

$$| H(s) | = | H_{LP}(s) + H_{HP}(s) | = 1.0 \quad \text{(requirements } a \text{ and } b\text{)}, \tag{1}$$

$$\arg(H_{LP}(s)) = \arg(H_{HP}(s)) \quad \text{(requirement } c\text{)}. \tag{2}$$

Our basic two-output crossover filter could be feeding two transducers. If requirements b or c are violated, output phases are no longer the same, and the magnitude of the complex sum of outputs is no longer exactly unity. Gain variations across the crossover region result.

Regarding requirement c, there are several factors that can contribute to produce a change in acoustic output phases. The transducers are mechanical bandpass devices with several resonant mechanisms, resulting in variations in the rate of phase angle change. The phase for each transducer must be determined

at the listening position. Excluding the listening room effects (assuming measurement in an anechoic chamber), the differences in transducer characteristics, the mechanical layout of transducers in a loudspeaker enclosure, and diffractions created by the loudspeaker mechanics all contribute to change the phases of audio signals from those of the electronic signals fed into transducers. If the relative phase between transducers does not remain constant across the crossover, the acoustic axis can shift in direction. When optimizing a crossover filter, all these factors must be accounted for. It is generally not a good idea to design a crossover filter in isolation as an academic exercise, without measurements of transducers in the actual loudspeaker enclosure.

The now ubiquitous *Linkwitz–Riley (LR) crossover filter* (Linkwitz, 1976, 1978; Lipshitz and Vanderkooy, 1986) was an improvement over earlier crossover networks that showed directional shifts and an increase in the main lobe magnitude. Cascading two identical Butterworth filters forms an even-order LR crossover network. The highpass and lowpass LR filters have identical phase characteristics except for a 180-degree phase difference for certain filter orders. The LR filter has a flat passband and monotonously increasing stopband attenuation. The combined magnitude response of LR high and lowpass filters is flat. The attenuation of the low and highpass filters separately is 6.0 dB at the crossover point (see Figure 2), corresponding to the 3.0-dB passband edges of the prototype Butterworth filters used as building blocks of an even-order LR filter.

As other conventional crossover filter types, the LR filter is a minimum-phase filter showing peaking of the throughput delay near the crossover frequency. Already at the time of its introduction, Linkwitz was concerned about this local change in delay in the LR filter because certain signals (he gives the example of a square wave) can show strong changes in the waveform shape due to shifting in time of harmonics constituting a signal waveform. Linkwitz did some listening tests to conclude that the increase of delay did not cause audible changes to such signals for a then-realistic fourth-order LR filter (Linkwitz, 1976).

A faster crossover transition can be obtained by increasing the filter order for designs using filters with monotonously increasing stopband attenuation, such as LR filters. When the order of a crossover filter is increased, an analysis of coefficient sensitivity is needed to understand how realizable the filter is with the chosen technology. It is possible to reduce the sensitivity of the design by selecting favorable circuit topologies (Cochenour et al., 2003).

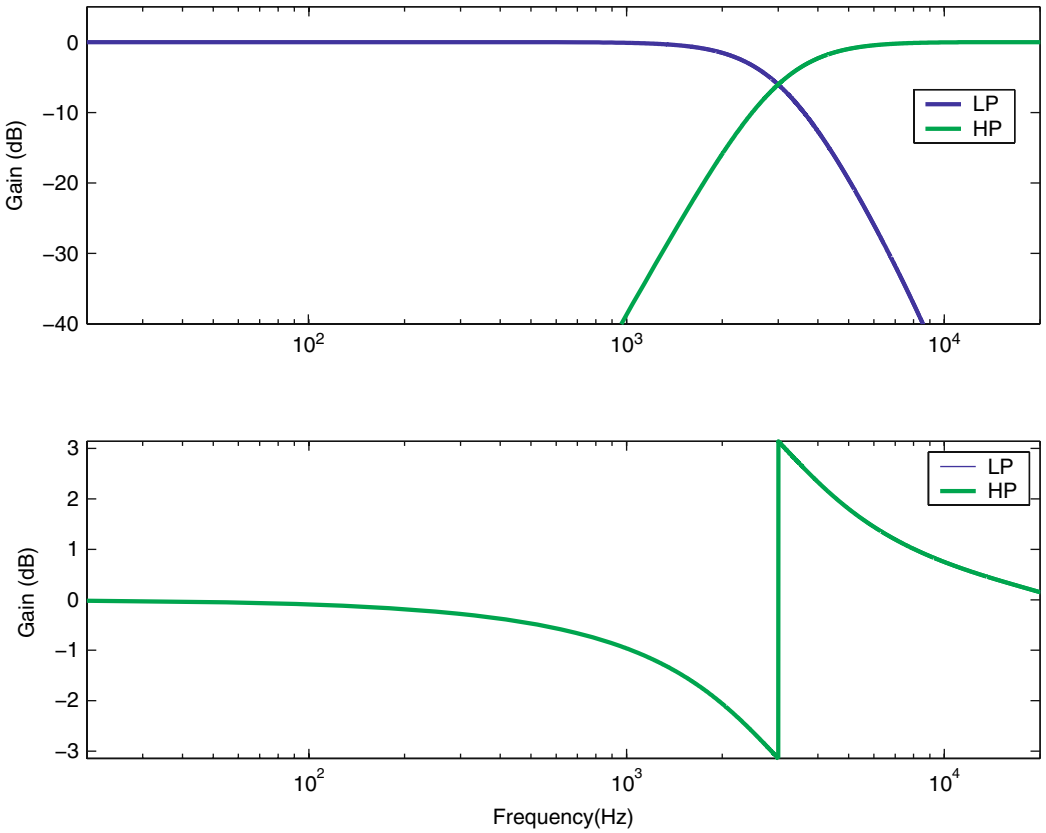


FIGURE 2 Example of the Linkwitz–Riley crossover filter response: complementary magnitude responses (*top*) and phase responses (*bottom*). LP = lowpass filter, HP = highpass filter.

A faster transition, and hence a narrower transition bandwidth, can also be obtained with filters having transfer function zeros in their stopband. Thiele (2000) describes a class of crossover filters having a null in their highpass and lowpass transfer functions at frequencies close to the crossover transition. The initial attenuation rate is higher than with LR filters, while the sum of the high and lowpass outputs shows an allpass response. When nulls are moved to very high and low frequencies, these transfer functions degenerate into Butterworth functions for the odd order and into Linkwitz–Riley functions for the even order. Thiele presents both active and passive realizations. Compared to LR filters, with monotonously increasing stopband attenuation, the faster transition of these filters results in a larger delay variation close to the crossover frequency.

Certain minimum-phase filter types such as Bessel and Gaussian filters show constant group delay at low frequencies. A crossover filter with reduced local increase of delay at the crossover point can be constructed using these filters (e.g., Fink, 1980). The main drawback is that these filters do not show a gain of −6 dB at the crossover point for designs that have equal phase response for both the high and lowpass filters. The complex sum of outputs will show level deviations at the crossover point.

By subtracting the filter output from the delayed input signal we can conveniently obtain two complementary outputs:

$$H_{\mathrm{HP}}(\mathrm{s}) = 1 - H_{\mathrm{LP}}(\mathrm{s}) \tag{3}$$

This principle has been applied to obtain *high-slope analogue* crossover filters (Lipshitz and Vanderkooy, 1983; Wilson et al., 1989). In the digital domain it is possible to obtain an exactly linear phase filter using a *finite impulse response* (FIR) filter. Such a filter can be designed to have a delay of an exact multiple of the sampling intervals, making the delay needed by the complementary filter design very easy to implement in the digital domain. *Infinite impulse response* (IIR) filters with similar behavior have also been suggested (Reviriego et al., 1998).

The transition bandwidth of crossover filters can in principle be made arbitrarily narrow, but narrowing the transition band also results in the increase of filter delay, particularly for linear phase FIR filters. Depending on the application, there will be a practical limit to how large the real-time delay of a loudspeaker can be. Another potential problem with linear phase filters is related to the fact that the impulse response of such a filter is symmetric so that the amount of energy in a linear phase system impulse response is equal before and after the maximum energy peak. FIR systems with fast transition bands also show long impulse responses before the main lobe, also known as pre-echo. Moving zeros in the FIR response inside the unit circle makes an FIR response increasingly minimum phase, and this can circumvent the problem of pre-echo. As a consequence, the FIR begins to show increasing local delay close to the crossover point, very much in the same way as with traditional designs.

In a conventional crossover design, where the phase is changing systematically and fairly slowly across the crossover region, one can always find a listening angle at a given distance where the two radiators are either in phase or in antiphase. An interesting idea due to Hawksford (1997) suggests a high-order crossover filter with random phase variation in both the highpass and lowpass

outputs across the crossover frequency. This *stochastic interleave* crossover could in theory minimize angle-dependent co-radiation effects.

3 • Frequency Response Equalization

Equalization techniques are needed to produce a flat and smooth frequency response. Rudimentary equalization is possible by modifying passive crossover filters based on anechoic acoustic measurements, which is a standard practice in loudspeaker design. Precise implementation of system equalization becomes easier with the active loudspeaker concept where component values and circuit properties can be chosen more freely than in passive designs.

Equalization is typically implemented as modifications of standard crossover filter responses. Optimization to a flat overall target response can be performed manually when the number of variables or electronic components to optimize is fairly small. There is also the possibility for automated optimization using nonlinear optimization methods. Optimization of circuitry and component values can be done off-line based on anechoic response measurements using standard system simulation and modeling tools, such as Spice (Spice, 2005) and Matlab (Matlab, 2005), or software intended for loudspeaker design such as Calsod (Audiosoft, 2005), or Leap (LinearX, 2005). A particular challenge in loudspeaker design is to combine the equalization with the mechanical design of the loudspeaker so that the frequency response, when moving away from the acoustic axis, remains flat (for an example, see Figure 3).

Digital signal processing offers further advantages in system equalization. Digital filters can have negligible transfer function deviation from the ideal target because coefficients can be presented very accurately, and coefficient sensitivity can be made small compared to their analogue counterparts. Once designed, an equalizer transfer function can therefore be implemented very precisely. Infinite impulse response filters can be used in a similar fashion to analogue systems (White, 1986; Orfanidis, 1997). Finite impulse response (FIR) filters offer new possibilities to use easily designable and unconditionally stable high-order equalizers in practical loudspeaker systems (Hawksford, 1999). FIR filters can also be used to equalize frequency-dependent delay (Clarkson et al., 1985; Murray, 1994), a feature seldom practical in most analogue implementations. Special filter implementations such as warped FIR filters and Kautz filters allow filter resolution to be frequency specific, for example with increasing zero-pole

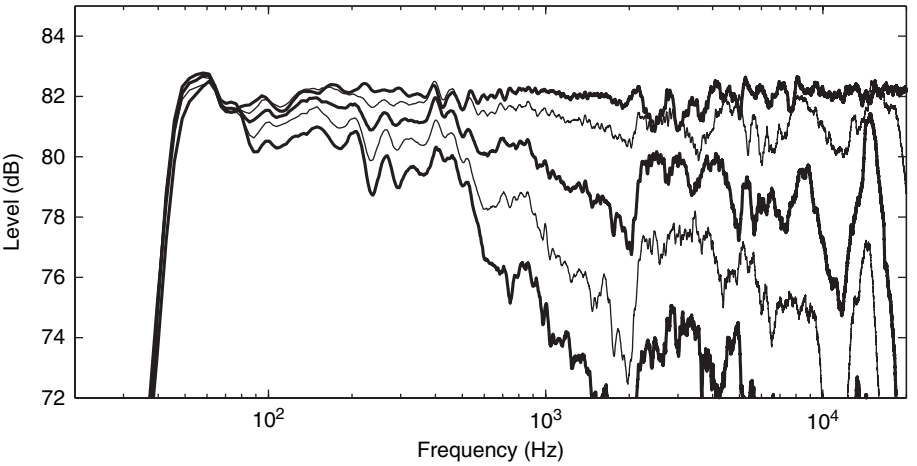


FIGURE 3 Example of the change in a loudspeaker frequency response when moving horizontally off-axis; *traces from top to bottom*: on-axis, 15 degrees, 30 degrees, 45 degrees and 60 degrees off-axis on the horizontal plane.

density toward low frequencies (Karjalainen et al., 1999; Wang et al., 2000; Paatero and Karjalainen, 2003).

Precise equalization can be effected by filtering the system H_S feed signal with the transfer function inverse H_E:

$$H_E(s) = 1/H_S(s), \tag{4a}$$

$$H(s) = H_S(s)\, H_E(s) = H_S(s)/H_S(s) = 1.0. \tag{4b}$$

Direct inverse of a loudspeaker response is usually not desirable. Theoretically, an inverse exists if the magnitude response at all frequencies of interest is larger than zero. Straightforward application of an inverse response, resulting in spectral whitening, leads to widening of the loudspeaker passband (for example, Clarkson et al., 1985). Typically there are certain frequencies, such as passband edges or transition bands, where equalization should not change the inherent response of the loudspeaker. Normally only the range of frequencies considered as the passband should be equalized. This leads to special requirements for the equalizing filter design. Several solutions for this problem have been proposed (Greenfield and Hawksford, 1991; Karachalios et al., 1995; Potchinkov 1998a, b; Kirkeby and Nelson, 1999; Karjalainen et al., 1999; Paatero and Karjalainen, 2003).

The ultimate precision of equalization depends on three factors – the capability of the equalizing filter to correct the

acoustic response, the precision of the acoustic measurements used as the data on which the design is based, and the capacity of the human hearing system. The practical limit to how flat a loudspeaker must be rendered by equalization is shifting from being a filter design and implementation problem to being a problem of precision and relevance of measurements used as the basis for equalizer design as well as the lack of capacity in the human hearing system to recognize small imprecision in the acoustic response flatness.

A loudspeaker can in principle be equalized to be perfectly flat in anechoic conditions for all practical purposes (e.g., Wilson et al., 1989; Karjalainen et al., 1999), but flatness only exists relative to the acoustic measurements used as the basis for equalizer design. Acoustic measurements show *variability* and *measurement uncertainty,* limiting how precisely the true acoustic response can be known. Variability is related to measurement noise. Measurement uncertainty is related to the instrumentation used, notably the microphone and anechoic chamber. These two should be calibrated relative to some (traceable) reference. With all this done, the true response is only known to a certain tolerance. Equalizer design can only be as good as the measurement upon which its design is based.

Also the human auditory system has limits beyond which further improvement in audio system flatness is no longer readily recognized. Broadband deviations from a flat response are heard more sensitively than narrow-band ($Q > 10$) peaks or notches, and narrow-band peaks are more readily heard than notches (Toole and Olive, 1988; Olive et al., 1997). Broadband continuous signals, such as pink noise, are the best to reveal wideband flatness deviations because the stationary spectral content in pink noise allows the spectral balance to be recognized. Narrow-band resonances and antiresonances become audible with transient signals (such as clicks). These work as short duration exciters for narrow-band resonances eliciting their natural impulse responses that become audible after the masking related to the exciting signal has ceased. We do not yet know in full detail the ability of the human auditory system to hear small impairments in the loudspeaker frequency response, yet practical equalizer design must strike a balance between complexity and subjective benefit.

4 • Room Response Equalizers

A loudspeaker with flat response in anechoic conditions will not appear flat when taken into a listening room. This is because of *early reflections,* or reflections producing a large number of

delayed replicas of the output signal, arriving at the listener very shortly after the direct sound, and because the *late reverberant field* in a room may be colored due to the sound absorption varying as a function of frequency. The aim of a loudspeaker installation is to set the loudspeaker up such that the frequency response at the primary listening area is sufficiently wide, with flat and smooth frequency response, so that there is a uniform directivity of sound over the whole intended listening area.

Equalization of the loudspeaker so that the frequency response at the listening area is flat can greatly improve perceived audio quality. Modern active loudspeakers typically provide special *room response adjustments* for this purpose. Low-order controls correcting the room response have a gentle shaping effect on the loudspeaker response. Typically tilt and roll-off controls are provided for the low frequencies of the woofer and for the high frequencies of the tweeter. In multi-way loudspeakers also transducer signal levels may be adjustable.

A loudspeaker is usually designed to produce a flat response in the free field. At low frequencies a loudspeaker is omnidirectional (radiating to 4π space), but changes to forward directing (radiating to 2π space) at a frequency determined approximately by the size of its front baffle. The woofer response controls are needed to optimize the low-frequency response for the 2π–4π transition when the loudspeaker is installed near or at a wall. The wall loading will limit the low-frequency radiation space, and boost the sound level below the 4π–2π transition frequency. When the loudspeaker is used behind a screen or if the acoustic axis of the speaker is not properly aligned, the tweeter response adjustment can be used to calibrate the very high-frequency response at the listening area.

Digital signal processing allows for easy implementation of more sophisticated and higher-order room response corrections. While it is possible to greatly improve reproduction quality, careless use of sophisticated equalizations can also result in audio quality worse than had no room equalizers been available. Typically a room response is optimized not at a single point in the listening space but within a listening area, typically with a diameter of 0.5–0.8 m. The response measured in a room can show strong *comb filtering effects* (transfer function zeros at a series of frequencies) due to strong reflections. To alleviate this problem, the response may be averaged in frequency either before filter design or indirectly as a part of it (Hatziantoniou and Mourjopoulos, 2000). When multiple responses are measured within the listening area, a weighted average of them may be calculated as a basis for equalizer design (Wilson, 1991;

Mourjopoulos, 1994; Bharitkar et al., 2002). This weight may also be frequency dependent. Particularly at low frequencies it is more likely that there are certain frequencies where a single microphone location is inadvertently placed in a room mode null. Moving the microphone around will enable this to be detected and accounted for (Mäkivirta et al., 2003).

It is also possible to reduce the perceived reverberation of a listening space to some degree by careful design of room equalization, such as by *complex smoothing* of the measured response before equalizer filter design (Hatziantoniou and Mourjopoulos, 2004). The so-called *modal equalization* can be applied to shorten the too long decay time of low-frequency modes of a room (Mäkivirta et al., 2003).

5 • System Delay and Delay Variation

Traditional loudspeaker designs have used fairly low-order minimum-phase filtering, and the system throughput delay or its variation as a function of frequency has not been an issue (e.g., Linkwitz, 1976: 7–8). Steep minimum-phase crossover filters, such as the Linkwitz–Riley crossovers, can produce excessive amounts of delay variation in frequency. As mechanical systems, loudspeakers are bandpass systems, and therefore will not have uniform delay characteristics. In fact, the system delay may show marked increase at a mechanical resonance of the transducer-enclosure system or the transducer itself.

The woofer displacement increases rapidly toward the low-frequency cut-off in a closed box design. Bass reflex and passive radiator designs can limit this increase, and are therefore frequently used to maintain low-frequency distortion at a low level, particularly in small enclosure designs. Both bass reflexes and passive radiators are mechanical resonators and will therefore increase loudspeaker delay close to their mechanical resonance. Despite a long-standing debate over the subjective importance of this added system delay, and demands in the popular press to make the phase response linear also at very low audio frequencies, there does not seem to be strong evidence suggesting that the additional delay generated in such designs would be detrimental to audio fidelity.

Frequency-dependent *delay variation* can shift in time the frequency components constituting a sound waveform, and can therefore change time-domain waveforms. In order to understand whether this has audible consequences we must know how much delay variation there can be before the human auditory system can notice a change. Audibility of delay variation is

frequency dependent. Early investigations report subtle audible differences for delay variation caused by high-frequency roll-off (Meyer, 1984), mid-frequencies (4 kHz Preis and Bloom, 1984; 2 kHz Fincham, 1985), and low-frequency roll-off (Preis and Bloom, 1984; Deer et al., 1985). At very low frequencies and at very high frequencies significant increases in the delay can remain inaudible. Variation of several milliseconds appears to be inaudible at frequencies below a few hundred Hertz while in the mid-frequencies (1–5 kHz) variation of more than 1 ms may be audible (Preis, 1982).

Linear phase systems are interesting because they do not cause any frequency-dependent delay variation. However, they can produce large system *latency*. Latency produces a problem similar to a large physical distance – the visual event and sound event happen at different times. Tolerance to latency depends on the type of use of an audio system. If performance depends on audiovisual coordination, for example in dubbing or ensemble performance mediated over loudspeakers, a large latency deteriorates performance. In the case of audiovisual monitoring, for example in film and television, the sound cannot be delayed much in relation to the picture. For these types of applications the loudspeaker designer must balance between minimizing the latency while keeping the frequency-dependent delay variation in the loudspeaker acceptably small, and linear phase filtering can be used only with limitations. For audio-only playback applications much larger system delay is tolerable, which allows more efficient use of linear phase filters.

6 • Protection

A loudspeaker can be destroyed by several mechanisms. The most obvious is *thermal overload*. The other central destruction mechanism is *displacement overload*. Most of the input power fed into a dynamic transducer is turned into heat. There are also several mechanisms of voice coil instability and diaphragm movement that can potentially lead to destruction of a transducer (Klippel, 1990, 1992).

In large woofers, excess heat is conducted away from the voice coil efficiently. In tweeters protection levels and time constants must be set more stringent than for woofers. A typical overheat protection uses a double integral of the output voltage (voltage integrated into power, power integrated into energy). When the heat energy estimated in this way exceeds a design limit, the output level is reduced to maintain safe long-term power level and to allow cooling. Short-term instantaneous input power can

be made very large compared to continuous power, particularly for tweeters. A more precise protection is possible if the voice coil and magnet structure temperatures can be known. In most practical systems such methods have not been deemed necessary because the incremental benefit in terms of additional output capacity usually is rather modest.

Several methods have been devised for displacement protection. The basic idea is to measure the voice coil location and to use this information to activate protection. It is also possible to calculate the voice coil position by using the physical force–mass relations. By double integrating the transducer driving current, related in electro-dynamic transducers to the force exerted by the voice coil, we can obtain an estimate of the voice coil location. This principle works when the nonlinearities of the suspension and variation of the magnetic flux in the air gap are also considered in calculations (Klippel, 1998, 2003).

7 • Understanding Loudspeaker Operation by Modeling

Lumped parameter modeling can be considered the principal method for developing understanding about the interaction of quantities governing loudspeaker system behavior (see Section 33.3). It may be beneficial to describe the whole electro-mechanic-acoustic system as a lumped parameter model. Computer-aided design techniques then allow various system scenarios to be examined early in the design phase without expensive and time-consuming prototype building. Different transducer parameters can also be inserted into simulations to determine how overall performance of a system might be affected. This enables the designer to identify important system design variables and to try different signal processing options to see their effects on the loudspeaker response.

Lumped parameter modeling using the Thiele–Small parameters was initially used to enable designers to predict the performance of a woofer in an enclosure and to optimize the low-frequency performance of the loudspeaker with a given transducer. Today, computer programs automate the use of lumped parameter calculation. Extensions to the original Thiele–Small model for large displacement conditions have been developed. Commercial measurement equipment exists for determining nonlinear changes to the lumped model parameters under large signal conditions (e.g., Klippel, 2005).

Directionality of a sound radiator depends primarily on its size (Kinsler et al., 1982). Design of a multi-way loudspeaker seeks a

good compromise between the transducer size and displacement-related distortion at the desired maximum sound level. The uniformity of energy dispersion across the listening area can be improved by mechanical design of the front baffle where the transducers are mounted. Shallow horns (Holland et al., 1996), pioneered by the Genelec DCW (Genelec, 2005), have become popular as a means to tailor directivity and to match transducer directivities across the crossover transition. Advanced modeling techniques based on *finite and boundary element methods* can be used to extend modeling to these structures, allowing simulations starting from fundamental physical properties, and therefore able to yield deeper understanding of mechanical phenomena of coupling of mechanical vibration of the transducer to air and propagation of the sound in air; and interaction with mechanical boundaries of this propagating sound wave suffers while it is traveling from the transducer to the listener.

A loudspeaker is typically designed as a stiff structure to eliminate secondary or parasitic vibrations. The intention is that only the transducers radiate sound. A notable exception from this principle is the distributed mode loudspeaker (DML, Prokofieva et al., 2002) where the radiation is produced by shaking a large surface area, evoking randomly distributed modal vibrations, and breaking up any pistonic operation. Such a radiator has fewer tendencies to be directive than a similar size pistonic radiator, but cannot be easily designed to produce a well-defined on-axis frequency response in the same way a conventional radiator can. Element modeling methods and *statistical modeling* have been used to study the properties of such radiators.

8 • Acoustic Measurement Techniques

Acoustic measurements reveal the actual behavior of a loudspeaker. Acoustic measurements can be categorized in free field and in situ measurements. *Free field measurements*, typically performed in anechoic rooms, are central to loudspeaker development work and in principle enable the study of a loudspeaker in isolation. Methods have been developed to simulate free field conditions in a real room by signal windowing (Struck and Temme, 1994).

In situ measurements study the combination of a loudspeaker and a listening room. The measurement methodology is in principle similar to that used in anechoic measurements. Additionally, smoothing in frequency is often used to obtain average response level in the presence of dense notching produced by the complex acoustic reflection field present in a listening

room. Time windowing of the impulse response data is also frequently employed to dissociate direct sound from reverberation.

Multitone techniques, where a signal is a composite of multiple suitably phased sinusoids, allow very fast approximate measurement of system magnitude and phase responses at all constituent frequencies (Czerwinski et al., 2001). Contemporary methods to estimate system impulse responses include the use of *maximum length sequences* (MLS; see, e.g., Rife and Vanderkooy, 1989; Vanderkooy, 1994) and *log-swept sinusoidal* signals (e.g. Müller and Massarani, 2001; Stan et al., 2002). The MLS measurement appears robust against environmental noise and can produce accurate results in noisy conditions (Stan et al., 2002). Also log-frequency sinusoidal sweep techniques can capture impulse responses with high signal-to-noise ratio (SNR) due to the possibility of rejecting simultaneous harmonic distortions and allowing analysis of linear responses and harmonic distortions simultaneously (Müller and Massarani, 2001). Measurement quality can degrade in the presence of *time variance*, for example swaying of the microphone stand (Svensson and Nielsen, 1999) or when a system is strongly nonlinear (Vanderkooy, 1994).

The loudspeaker response is optimized on the *acoustic axis*, whereby the acoustic axis defines the assumed direction of the listener. The *directivity* characteristics of a loudspeaker are usually studied by measuring frequency responses in certain angles (Seidel and Staffeldt, 1996). A typical selection of angles is from zero (on the acoustic axis) to 60 degrees with 15-degree increments. Rotating the loudspeaker around its horizontal and vertical planes on a turntable and measuring the frequency responses can obtain a more accurate understanding of directivity. It is important to rotate the loudspeaker around its acoustic and not the mechanical center. These two may not coincide. If the loudspeaker is rotated around its mechanical center, there will be a direction-related error in the acoustic distance and hence on the recorded sound level.

The list of commercially available acoustic measurement systems is rather volatile, but some examples include (primary technologies are given in parenthesis) MLSSA (2005, MLS sequence), WinMLS (Morset, 2005, log-sine sweep and MLS), Clio (Audiomatica, 2005, log-sine sweep and MLS), and LMS (LinearX, 2005, sine sweep). This list could be a lot longer and may not be understood to be an endorsement for any of the products mentioned. There are several other quite capable measurement systems available, and the reader should use the list merely as a starting point for evaluating products for his purposes.

9 • Listening Tests

Modern measurement and modeling methods produce detailed data on loudspeaker performance. Yet we lack a method to map this information to the listener's experience. The subjective quality judgments relate to the main categories of flatness of the frequency response, quality and accuracy of virtual imaging, and lack of distortion or extraneous noises during operation. A loudspeaker is expected not to color the sound even when used in a reverberant room. Virtual sound images constructed between loudspeakers are expected to be precise and well defined (not stretched or blurred). The loudspeaker is expected to be able to reproduce the whole dynamic range in the sound recording without detectable increase in distortion at high output levels and without producing extraneous sounds. Although discussion about listening tests goes beyond the scope of this article, the importance of first-hand listening testing can never be overemphasized.

References

Audiomatica. Clio Measurement System. Audiomatica Srl. http://www.mclink.it/com /audiomatica/home.htm, 2005

Audiosoft. http://members.optusnet.com.au/~ audiosoft/, 2005

Bharitkar S., Hilmes P., Kyriakakis C., "Robustness of Spatial Averaging Equalization Methods: A Statistical Approach." IEEE 36th Asilomar C. Signals, Systems & Computers, Pacific Grove, CA, Nov. 3–6, 2002

Clarkson P. M., Mourjopoulos J., Hammond J. K., "Spectral, Phase, and Transient Equalization for Audio Systems", J. Audio Eng. Soc., Volume 33, Number 3, pp. 127–132. March 1985

Cochenour B., Chai C., Rich D. A., "Sensitivity of High-Order Loudspeaker Crossover Networks with All-Pass Response", J. Audio Eng. Soc., Volume 51, Number 10, pp. 898–911. October 2003

Czerwinski E., Voishvillo A., Alexandrov S., Terekhov, A., "Multitone Testing of Sound System Components, Some Results and Conclusions, Part 2: Modeling and Application", J. Audio Eng. Soc., Volume 49, Number 12, pp. 1181–1192. December 2001

Deer J. A., Bloom P. J., Preis, D., "Perception of Phase Distortion in All-Pass Filters", J. Audio Eng. Soc., Volume 33, Number 10, pp. 782–786. October 1985

Fincham L. R., "The Subjective Importance of Uniform Group Delay at Low Frequencies", J. Audio Eng. Soc., Volume 33, Number 6, pp. 436–439. June 1985

Fink D. G., "Time Offset and Crossover Design", J. Audio Eng. Soc., Volume 28, Number 9, pp. 601–611. September 1980

Garde P., "All-Pass Crossover Systems", J. Audio Eng. Soc., Volume 28 Number 9, pp. 575–584. 1980

Genelec. Genelec DCW Technology. http://www.genelec.com/support/dcw_tech.php, 2005

Greenfield R., Hawksford M. O., "Efficient Filter Design for Loudspeaker Equalization", J. Audio Eng. Soc., Volume 39, Number 10, pp. 739–751. October 1991

Hatziantoniou P. D., Mourjopoulos J. N., "Generalized Fractional-Octave Smoothing of Audio and Acoustic Responses", J. Audio Eng. Soc., Volume 48, Number 4, pp. 259–280. April 2000

Hatziantoniou P. D., Mourjopoulos J. N., "Errors in Real-Time Room Acoustics Dereverberation", J. Audio Eng. Soc., Volume 52, Number 9, pp. 883–899. 2004

Hawksford M. O., "Digital Signal Processing Tools for Loudspeaker Evaluation and Discrete-Time Crossover Design", J. Audio Eng. Soc., Volume 45, Number 1/2, pp. 37–62. January/February 1997

Hawksford M. O., "MATLAB Program for Loudspeaker Equalization and Crossover Design", J. Audio Eng. Soc., Volume 47, Number 9, pp. 706–719. September 1999

Holland K. R., Fahy F. J., Newell P. R., "The Sound of Midrange Horns for Studio Monitors", J. Audio Eng. Soc., Volume 44, Number 1/2, pp. 23–36. January/February 1996

Karachalios G., Tsoukalas D., Mourjopoulos J., "Multiband Analysis and Equalisation of Loudspeaker Responses", AES 98th Convention, 1995

Karjalainen M., Piirilä E., Järvinen A., Huopaniemi, J., "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques", J. Audio Eng. Soc., Volume 47, Number 1/2, pp. 14–31. January/February 1999

Kinsler L., Frey A., Coppens A., Sanders J., Fundamentals of Acoustics. Third edition. Wiley, New York. 1982

Kirkeby O., Nelson P. A., "Digital Filter Design for Inversion Problems in Sound Reproduction", J. Audio Eng. Soc., Volume 47, Number 7/8, pp. 583–595. July/August 1999

Klippel W., "Dynamic Measurement and Interpretation of the Nonlinear Parameters of Electrodynamic Loudspeakers", J. Audio Eng. Soc., Volume 38, Number 12, pp. 944–955. December 1990

Klippel W., "Nonlinear Large-Signal Behavior of Electrodynamic Loudspeakers at Low Frequencies", J. Audio Eng. Soc., Volume 40, Number 6, pp. 483–496. June 1992

Klippel W., "Direct Feedback Linearization of Nonlinear Loudspeaker Systems", J. Audio Eng. Soc., Volume 46, Number 6, pp. 499–507. June 1998

Klippel W., "Assessment of Voice-Coil Peak Displacement Xmax", J. Audio Eng. Soc., Volume 51, Number 5, pp. 307–324. May 2003

Klippel W., http://www.klippel.de/pubs/default.asp, 2005

LinearX. http://www.linearx.com/, 2005

Linkwitz S. H., "Active Crossover Networks for Noncoincident Drivers", J. Audio Eng. Soc., Volume 24, Number 1, pp. 2–8. January/February 1976

Linkwitz S. H., "Passive Crossover Networks for Noncoincident Drivers", J. Audio Eng. Soc., Volume 26, Number 3, pp. 149–150. March 1978

Lipshitz S. P., Vanderkooy J., "A Family of Linear-Phase Crossover Networks of High Slope Derived by Time Delay", J. Audio Eng. Soc., Volume 31, Number 1/2, pp. 2–20. January/February 1983

Lipshitz S. P., Vanderkooy, J., "In-Phase Crossover Network Design", J. Audio Eng. Soc., Volume 34, Number 11, pp. 889–894. November 1986

Mäkivirta A., Antsalo P., Karjalainen M., Välimäki, V., "Modal Equalization of Loudspeaker – Room Responses at Low Frequencies", J. Audio Eng. Soc., Volume 51, Number 5, pp. 324–343. May 2003

Matlab. http://www.mathworks.com/, 2005

Meyer J., "Time Correction of Anti-aliasing Filters Used in Digital Audio Systems", J. Audio Eng. Soc., Volume 32, Number 3, pp. 132–137. March 1984

MLSSA. MLSSA Measurement Software. DRA Laboratories. http://www.mlssa.com/, 2005

Morset. WinMLS Software. Morset Sound Development. http://www.winmls.com/, 2005

Mourjopoulos J., "Digital Equalization of Room Acoustics", J. Audio Eng. Soc., Volume 42, Number 11, pp. 884–900, 1994

Müller S., Massarani, P., "Transfer-Function Measurement with Sweeps", J. Audio Eng. Soc., Volume 49, Number 6, pp. 443–471. June 2001

Murray J. A., "Microalignment of Drivers via Digital Technology", J. Audio Eng. Soc., Volume 42, Number 4, pp. 254–264. April 1994

Olive S. E., "Differences in Performance and Preference of Trained Versus Untrained Listeners in Loudspeaker Tests: A Case Study", J. Audio Eng. Soc., Volume 51, Number 9, pp. 806–825. September 2003

Olive S. E., Schuck P. L., Ryan, J. G., Sally, S. L., Bonneville, M. E., "The Detection Thresholds of Resonances at Low Frequencies", J. Audio Eng. Soc., Volume 45, Number 3, pp. 116–128. March 1997

Orfanidis S. J., "Digital Parametric Equalizer Design with Prescribed Nyquist-Frequency Gain", J. Audio Eng. Soc., Volume 45, Number 6, pp. 444–455. June 1997

Paatero T., Karjalainen, M., "Kautz Filters and Generalized Frequency Resolution: Theory and Audio Applications", J. Audio Eng. Soc., Volume 51, Number 1/2, pp. 27–44. January/February 2003

Potchinkov A., "Frequency-Domain Equalization of Audio Systems Using Digital Filters, Part 1: Basics of Filter Design", J. Audio Eng. Soc., Volume 46, Number 11, pp. 977–987. November 1998a

Potchinkov A., "Frequency-Domain Equalization of Audio Systems Using Digital Filters, Part 2: Examples of Equalization", J. Audio Eng. Soc., Volume 46, Number 12, pp. 1092–1108. December 1998b

Preis D., "Phase Distortion and Phase Equalization in Audio Signal Processing-A Tutorial Review", J. Audio Eng. Soc., Volume 30, Number 11, pp. 774–794. November 1982

Preis D., Bloom P. J., "Perception of Phase Distortion in Anti-alias Filters", J. Audio Eng. Soc., Volume 32, Number 11, pp. 842–848. November 1984

Prokofieva E., Horoshenkov K. V., Harris N., "Intensity Measurements of the Acoustic Emission from a DML Panel", Preprint Number 5609, Proc. 112 Convention of AES, April 2002

Reviriego P., Parera J., García R. "Linear-Phase Crossover Design Using Digital IIR Filters", J. Audio Eng. Soc., Volume 46, Number 5, pp. 406–411. May 1998

Rife D., Vanderkooy, J. "Transfer-Function Measurement with Maximum-Length Sequences", J. Audio Eng. Soc., Volume 37, Number 6, pp. 419–444. June 1989

Seidel F., Staffeldt, H., "Frequency and Angular Resolution for Measuring, Presenting, and Predicting Loudspeaker Polar Data", J. Audio Eng. Soc., Volume 44, Number 7/8, pp. 555–568. July/August 1996

Spice. http://bwrc.eecs.berkeley.edu/Classes/IcBook/SPICE/, 2005

Stan G.-B., Embrechts J.-J., Archambeau, D., "Comparison of Different Impulse Response Measurement Techniques", J. Audio Eng. Soc., Volume 50, Number 4, pp. 249–262. April 2002

Struck C. J., Temme, S. F., "Simulated Free Field Measurements", J. Audio Eng. Soc., Volume 42, Number 6, pp. 467–482. June 1994

Svensson U. P., Nielsen, J. L., "Errors in MLS Measurements Caused by Time Variance in Acoustic Systems", J. Audio Eng. Soc., Volume 47, Number 11, pp. 907–927. November 1999

Thiele N., "Loudspeaker Crossovers with Notched Responses", J. Audio Eng. Soc., Volume 48, Number 9, pp. 786–799. September 2000

Toole F. E., Olive, S. E., "The Modification of Timbre by Resonances: Perception and Measurement", J. Audio Eng. Soc., Volume 36, Number 3, pp. 122–142. March 1988

Vanderkooy J., "Aspects of MLS Measuring Systems", J. Audio Eng. Soc., Volume 42, Number 4, pp. 219–231. April 1994

Wang P., Ser W., Zhang, M., "A Dual-Band Equalizer for Loudspeakers", J. Audio Eng. Soc., Volume 48, Number 10, pp. 917–921. October 2000

White S. A., "Design of a Digital Biquadratic Peaking or Notch Filter for Digital Audio Equalization", J. Audio Eng. Soc., Volume 34, Number 6, pp. 479–483. June 1986

Wilson R., "Equalization of Loudspeaker Drive Units Considering Both On- and Off-Axis Response", J. Audio Eng. Soc., Volume 39, Number 10, pp. 127–139, 1991

Wilson R., Adams, G., Scott, J., "Application of Digital Filters to Loudspeaker Crossover Networks", J. Audio Eng. Soc., Volume 37, Number 6, pp. 455–464. June 1989

34

PA Systems for Indoor and Outdoor

Jan Felix Krebber

Institute of Communication Acoustics, Ruhr-University Bochum, Bochum, Germany, e-mail: jan.krebber@rub.de

1 • Introduction

Sound reinforcement (SR) of the voice has allowed to build bigger halls and to address large gatherings. Between 1915 and 1960, public address (PA) systems were mainly used to address gatherings. The first recorded use of an outside PA system or SR system goes back to December 30th 1915 (Yaxley, 2003). In San Francisco, Governor Hiram Jones was intended to deliver a speech in the new Civic Auditorium. However, unfortunately he fell ill and engineers Jensen and Pridham thought of a solution of how to transmit Hiram Jones' voice to the public. The auditorium was equipped with "Magnavox" equipment to reinforce the talker's voice. Jensen and Pridham installed a special line between the Governor's home and the sound reinforcement system in the auditorium. Eventually, Hiram addressed the public by means of a microphone while sitting in front of his fireplace.

SR systems were capable of distributing sound for more than 90,000 people (e.g., opening ceremony of the Wembley Exhibition 1925 by King George). In the 1960s, Rock'n' Roll music started to need amplification, especially for electric guitars and vocals. A new phenomenon occurred: The audience was not silent anymore during the performance, and PA systems had to be louder than

the audience. In 1965, the Beatles were the first rock band to play in a stadium of more than 55,000 visitors (Shea Stadium, New York). People screamed at such a volume, that even the musicians on stage could not hear themselves anymore. This was one of the reasons, why the Beatles stopped performing live concerts in 1966 (Cross, 2005). The band "The Who" was the first band known to the author using a stage monitor in 1968, which partly transmitted the vocals back to the stage. In the early years of monitoring, sidefill stage loudspeakers provided the same signal as the PA system.

Over the years, the monitoring strategy changed from onstage self-monitoring by the musicians' own amplifiers to individual monitor mixes and monitor wedges for each member of the band. The most important step in monitoring was to move from ear-damaging sound pressure levels (SPL) on stage to ear-friendly in-ear monitoring.

The latest deep impact on PA systems was made by L-ACOUSTICS by introducing line arrays in the early 1990s and the research results of "Wavefront Sculpture Technology"[1] (Urban et al., 2003). Line arrays deliver a higher and more directive sound pressure level than any other system known so far. Nowadays, it is a must for every PA system manufacturer to offer at least one line array system from their own product range.

Another interesting "back to the roots" PA solution is introduced by Bose's "Personalized Amplification System"[2] for smaller venues. The system provides again own loudspeakers for every musician. These loudspeakers serve as monitor and as PA at the same time.

This chapter presents a first insight into PA systems. It will shortly discuss all components of a PA system and give some ideas how to plan and how to set up the sound system regarding indoor and outdoor effects and measurements. After all, there will be a short discussion about digital vs. analog equipment and about things to be taken care besides signal processing. However, the design of sound systems is some kind of art, and it requires much experience, gained from practical and theoretical knowledge.

For further reading, Davis and Jones (1989), Davis and Davis (1997), and Stark (1989) will discuss the field of sound reinforcement from a practical point of view, whereas Ahnert and Steffen (1999) and Ballou (2002) provide very detailed

[1]Wavefront Sculpture Technology and WST are trademarks of L-ACOUSTICS.
[2]Personal Amplification System is a trademark of Bose.

and theoretical information on sound systems. For the first steps using mixing desks, the reader may find helpful advice in Trubit (2000), and for microphone techniques, Pawera (2003), Huber and Williams (1922) or Alldrin (1997) will give advice.

2 • The PA System

The core of a PA system consists of one or more input transducers, one or more signal processing stages, and one or more output transducers. A conceptual model of a sound system is shown in Figure 1.

Input transducers convert either sound waves or mechanical, optical, or magnetic changes over time from storage media into electrical signals. The types of input transducers normally found in sound reinforcement systems are as follows:

Air pressure and velocity microphones are used to convert sound waves traveling through air into electrical signals. Usually dynamic, condenser, or electret microphones are used. Other types of microphones like ribbon microphones or carbon microphones are not of interest for sound reinforcement applications due to the lack of suitable frequency range (carbon microphone) or stability (ribbon and piezoelectric microphone) for hard hiring company environment.

Contact pickups convert sound waves traveling through a solid medium into electrical signals.

Magnetic pickups convert sound waves of induced magnetism, e.g., metal strings into electrical signals.

Phono pickups convert mechanical movement of a stylus into electrical signals. There exist two different systems: MM = moving magnet and MC = moving coil.

Tape heads, optical pickups, and *laser pickups* convert recorded material from a certain medium (magnetic tape, photographic film, and imprinted patterns on a CD via digital-analog conversion) into electrical signals. The signal shape of all the electrical signals after conversion (and equalizing frequencies if

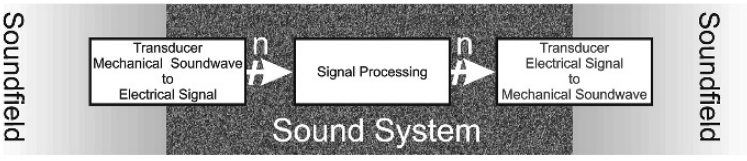


FIGURE 1 Conceptual model of a sound system.

necessary, e.g., RIAA[3]) look very similar (regarding the use of high-quality equipment), but the overall level of the converted electrical signals may be very different depending on the type of conversion.

The signal processing is divided into several functional units. The same units may occur several times at different positions in the signal paths. These units are pre-amplification, delay, equalization, dynamic level processing, room and delay effects, mixing and distributing, and power amplification.

Pre-amplification brings the weak microphone signal to line level. The microphone pre-amplification stage may allow gains up to 70 dB.

Delays allow a simple phase reverse or the time alignment for different microphone positions targeting the same source. This technique requires much knowledge as it might lead to severe comb-filter effects. In the later signal processing path, delays may also be used for the time alignment of loudspeakers.

Equalization or *filtering* allows to adjust the tonal balance of the audio signal. There are several types of filters: highpass and lowpass filters, shelving filters, and peaking filters. Highpass and lowpass filters may allow a variable cutoff frequency. Shelving filters can be controlled in cut or boost and depending on the implementation also in cutoff frequency. Fixed peaking filters or graphic EQs can cut or boost the audio signal at a fixed frequency. The semi-parametric or quasi-parametric EQs can cut or boost the audio signal with filters at a variable frequency, whereas the fully parametric EQ allows also controlling of the filter width. A special type of filter is the notch filter. It cancels a certain frequency with a very narrow bandwidth. Variable notch filters are usually not found in analog equipment because circuit design for proper control of the frequency is too complicated. Instead, peaking EQs with a fixed gain cut and very narrow filter width (1/10 octave) are implemented. All the relevant filter types are shown in Figure 2.

Dynamic Level Processing units are compressors, limiters, de-essers, and noise. These units reduce the dynamic range. For linear dynamic level processing, the ratio output level/input level is a constant of value 1. A *compressor* allows to change this ratio (compression ratio) from a certain input level (threshold) upward.

A special case of a compressor is the *limiter*. Above a set threshold, the ratio is $\infty : 1$, which means that the input signal

[3]Record Industry Association of America. The RIAA response curve is used for records to overcome technical problems while cutting and playing records.

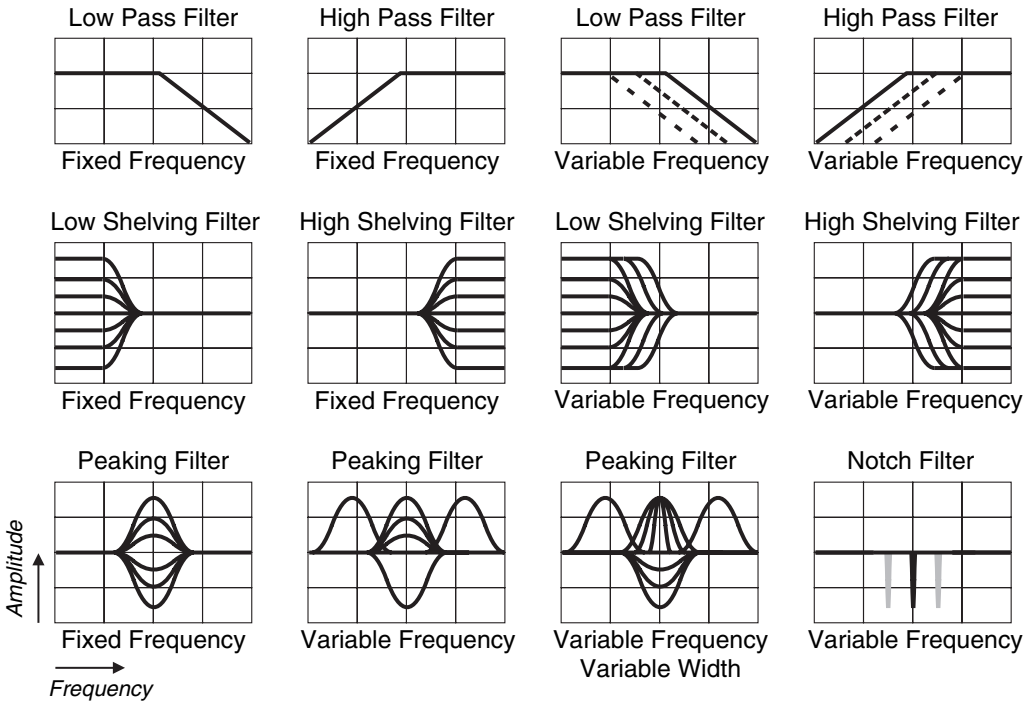


FIGURE 2 Lowpass and highpass, shelving, peaking, and notch filters with fixed or variable frequency.

could theoretically reach infinity, but the output signal would stay at the threshold level.

Another member of the compressor family is the *de-esser*. There is an additional bandpass filter in the control path of the compressor, which allows for frequency-dependent compression of the audio signal. De-essers are mainly used to reduce the loudness of "s"-sounds (which may occur to be louder than the rest of the vocal signal) leaving the rest of the material untouched. Without a de-esser, one might reduce the high-mid frequencies to avoid too loud "s"-sounds, but this might result in a dull sound. The "s"-phoneme frequencies are mainly found around 4.5 and 6.7 kHz and the "sh"-phoneme frequencies around 3.4 and 5.1 kHz for male and female voices, respectively.[4]

Gates cut off a signal for which power is below a certain threshold. Usually, not all microphones are used at the same time and they can be "turned off" for the duration of their source

[4]Fricative resonance frequencies vary according to vowel context.

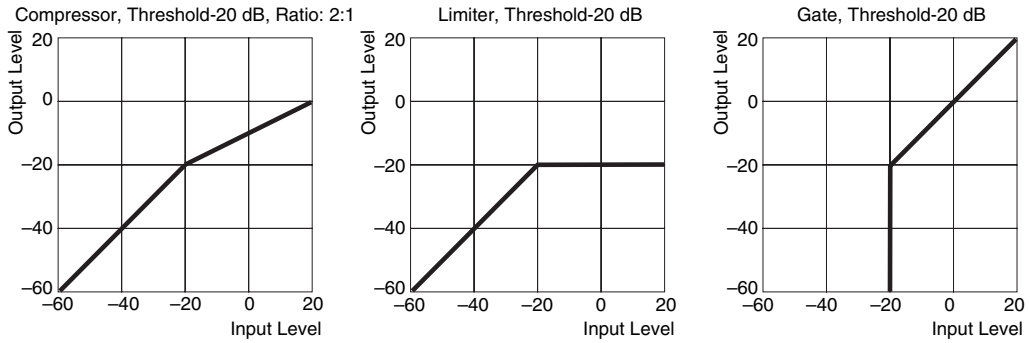


FIGURE 3 Level dependency of input and output signals for different dynamic level processors.

being "silent". In this way, gates allow to reduce the amount of microphones "being open" and finally increase the gain-before-feedback.[5] The level dependency of input and output signals for dynamic level processors is depicted in Figure 3.

Delay and reverb effects are based on simulating the idea of reflecting a sound source when it hits a boundary. The delay repeats the sound source several times with constant time interval and more and more reduced intensity. Usually, the delay effect is implemented as a single delay stage with an intensity adjustable feedback loop. The delay effect is more artificial than realistic, and it is mainly used to thicken the sound with echoes in a controlled, rhythmic way. The reverb effect aims at modeling a room with some early reflections and subsequent dense reverberation. The parameter settings allow either more or less realistic simulations of certain rooms, but also some extreme and artificial room effects which would be impossible in the real world. The reverb effect is mainly used to add liveliness and spaciousness to a dry signal.

Mixing and distributing the sound from several sources to different destinations is the main task of a mixer, although in the first step all sources are processed individually (e.g., by the preamplifier, delay, dynamic processing, and EQ). After that, sources are mixed for different purposes, e.g., monitoring, effects for a group of sources, and for the main PA system. A practical guidance how to use mixers can be found in Trubit (2000). Besides the mixing capability, mixing desks also provide EQs and sometimes dynamic effects like gates or compressors.

[5]Gain-before-feedback is the maximum gain allowed between microphone and loudspeaker just before the sound system starts to ring. Doubling the number of microphone being "open" decreases the gain-before-feedback theoretically by 3 dB.

Power amplification is the last stage of the signal processing chain before the electrical signal is sent to the output transducers. There are different topologies of amplifiers available on the market. Nowadays, switch-mode technology can be found in the power supply, or in the amplification stage itself. The advantages are reduced weight and less loss of energy due to a better degree of efficiency. The limitation of power being available from an amplifier is usually caused by the limits of the power supply (overload or clipping) or heatsink capabilities (most times resulting in the minimum load impedance allowed at the output[6]).

So far, only headphones are capable of propagating nearly the full audio spectrum by a single driver unit. However, single output transducers are not able to convert the electrical signal in the full audio frequency spectrum between 20 Hz and 20 kHz (4 decades or 10 octaves) with high sound pressure level. Therefore, transducers are optimized for several frequency ranges. In PA systems, transducers are mainly loudspeakers from different frequency ranges which are combined to cover the full spectrum. The types of loudspeakers most commonly encountered in SR systems are listed in Table 1. The types of enclosures found in PA systems are discussed in the following:

TABLE 1 Speakers, their frequency range, and the possible types of transducers

Speaker	Frequency range	Type of transducer (diameter)
Subwoofer	20–300 Hz, usually 30–100 Hz	15″–21″ cone-type drivers
Woofer	Usually < 500 Hz for LF,[a] < 1.5 kHz for LF + MF[b]	10″–18″ cone-type drivers
Mid-range	Typically > 500 Hz, < 6 kHz	5″–12″ cone-type drivers, 2″–4″ compression drivers
Tweeter	> 1.5 kHz, usually > 6 kHz	2″–5″ cone-type drivers, 1″–4″ compression drivers, electrostatic drivers
Super tweeter	Above 10 kHz	compression drivers, piezoelectric drivers

[a] Low frequencies.
[b] Mid-frequencies.

[6]Often amplifiers are capable of driving lower impedance loads than specified, but they might switch off after a couple of minutes as they are getting too hot.

Sub-lows are usually used for frequencies below 100 Hz down to subsonic.

Lows do not reach as low a frequency as sub-lows, but may go up to 250 Hz.

Full-range enclosures contain loudspeakers which cover the frequency range from 100 Hz or below and up to 13 kHz or above. In case full-range enclosures are used with lows, their low cut-off will shift to higher frequencies to gain more power for the low-mid frequency range.

Monitor Wedges are full-range loudspeakers that are built in a special shape to be laid on the floor and to point to the performer. They are used to return a "monitoring signal" to the performer to help him or her to stay in tune or to keep the timing by listening some certain instruments from the band.

3 • Planning and Setting Up the Loudspeakers

The typical area for sound distribution range is between 10 m for close distances up to 200 m for far distances from a single source. For a rough estimation of the sound pressure level required for a PA system, one may take the following items into consideration:

- The *noise floor* of the venue. A babbling crowd will be around 86 dB.
- The *sound level* should be remarkably higher than the noise floor, e.g., perceived with double loudness than the noise floor. The perceived loudness doubles with 10 dB SPL.
- The *longest distance* of your audience area. The loss will be 6 dB SPL by doubling the distance, e.g., in the far distance of 200 m, the SPL will be 46 dB less than at 1 m distance.
- Allow the system a *headroom*; 10 dB is a reasonable value for normal applications.

The speaker should provide $86 + 10 + 46 + 10\,\text{dB} = 152\,\text{dB}$ if the atmospheric effects are not taken into account. This high SPL cannot be distributed by a single speaker and would cause too high SPL in the near field. Therefore, large area sound distribution requires different types of loudspeakers for short- (up to 20 m), intermediate- (up to 100 m), and long-throw (up to 200 m) sound distribution. However, correctly installed line arrays can handle all distances with one type of loudspeaker except very close distance to the stage (less than 10 m).

Large venues or fixed installations are planned by the help of computer programs. Usually these programs require a computer-aided design (CAD) drawing of the room with some additional parameters concerning the boundaries, as there is

the degree of scattering, the degree of transmission, and the degree of absorption. Loudspeaker parameters are provided by the loudspeaker vendors and are available in databases. Good programs should deliver parameters like SPL,[7] RT,[8] STI,[9] RASTI,[10] C_{80},[11] D_{50},[12] echograms, and auralization.[13] These parameters are necessary for estimating the quality of sound related to the installed PA system.

However, most venues do not need, or the budget does not allow, a computer-aided planning of the loudspeaker positions. In the following, some common setups are discussed and depicted in Figure 4.

1. Central public address: The sound is propagated from a central point. The sound waves from the PA system cover mainly the back area of the listeners area, where the direct sound from the podium would be too weak. To avoid additional reverberation, boundaries of the first reflections

---

[7]Calculation of Sound Pressure Level can be found in IEC61672-1-1:2002.

[8]Reverberation Time is the time the reverberation takes to die away. For example, RT_{60} is the time it takes for a sound's reverberation field to attenuate by 60 dB.

[9]Further information about the Speech Transmission Index can be found in IEC60268-16:1998-16:1998.

[10]The RApid Speech Transmission Index is a faster version of the STI. Further information can be found in IEC60268-16:1998-16:1998.

[11]C_{80} is usually named clarity and is related to music performances. It is an objective value which estimates the perceptibility of tones spread in time or differentiation of concurrent sounds made by different instruments. C_{80} is the ratio of energy within the first 80 ms vs. the energy of the subsequent sound.

$$C_{80} = \frac{\int_{0\,\mathrm{ms}}^{80\,\mathrm{ms}} p^2(t)\mathrm{dt}}{\int_{80\,\mathrm{ms}}^{\infty} p^2(t)\mathrm{dt}}. \tag{1}$$

Further information can be found in ISO3382:1997.

[12]D_{50} "Definition" or "Deutlichkeit" is sometimes used for speech conditions. It is the ratio of energy within the first 50 ms vs. overall energy of the sound. Together with C_{80}, it gives the balance between early and late arriving energy.

$$D_{50} = \frac{\int_{0\,\mathrm{ms}}^{50\,\mathrm{ms}} p^2(t)\mathrm{dt}}{\int_{0\,\mathrm{ms}}^{\infty} p^2(t)\mathrm{dt}}. \tag{2}$$

Further information can be found in ISO3382:1997.

[13]Auralization is the convolution of the estimated room impulse response with a source file at predefined places and predefined directions for playback by headphones.

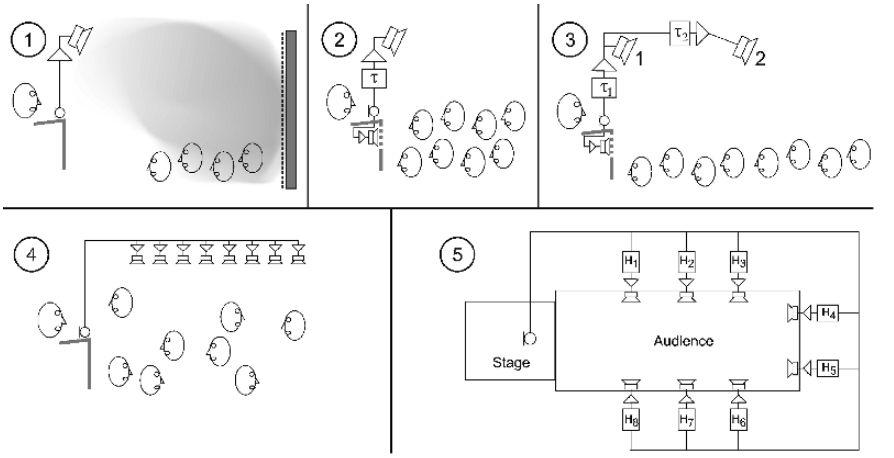


FIGURE 4 Different setups of loudspeakers, adopted from Blauert (2000).

are built in a sound absorbing manner. Usually this is the rear panel and the audience area.

2. The next approach makes use of the law of the first wave front. In case the amplified sound wave is delayed 10–20 ms (related to the original sound wave from the talker), the auditory event is localized at the position of the talker. The amplified sound wave may be up to 10 dB louder at the listener's position without being localized as such, as loudness and intelligibility are related to the amplified sound. For enhanced PA systems for medium-sized events, a loudspeaker is placed in or next to the podium, in addition to main loudspeakers. For live concerts, a short delay of the main PA (5–20 ms) may bring sound from stage more directly to the audience.
3. For large rooms or places or rooms with niches or protrusions covering the listening area, a column setup will be applied. The delay unit τ_2 will compensate the delay of sound between the first and second column. It will cancel out echoes from the first column at the place of the second row.
4. In shopping malls or train stations, distributed PA systems can be found. They include many synchronously driven full-range enclosures which emit a diffuse sound field. The talker will not be localized at his original position, but this does not disturb the listener. The propagated SPL of every loudspeaker has to be set very carefully.
5. Artificial reverberation is added to originally dry rooms which show a short reverberation time, e.g., multipurpose venue places or which should sound larger than they are,

e.g., theaters in cruiser ships. The signal processing is either a set of parallel delay units or an impulse response. These systems are uniquely tuned for rooms they are installed in.

4 • Sound Indoors

Important physical phenomena for indoor sound reinforcement are refraction, reflection, diffusion, diffraction, and absorption.

Refraction bends the sound wave due to changes in velocity of sound through a medium. This will be explained as temperature gradient in Section 5.

Reflection is the ability of the surface material to mirror an incoming sound wave with a certain amount of loss. It is the counterpart of absorption. The surface material acts as a filter to the incoming sound; some energy is turned into heat but other fractions of the energy are reflected back into the room.

Diffusion is a special form of reflection. Diffusor materials have desired surface irregularities which should spread the incoming sound wave equally in all directions.

Diffraction is the bending of a sound wave around an obstacle, in case the obstacle is small with respect to the wavelength. Otherwise the obstacle shades the sound, which means that diffraction and shading are frequency dependent as shown in Figure 5. For further information, Ballou (2002) may provide helpful information.

Absorption is the loss of energy when a sound wave strikes a given material. It is specified by the absorption coefficient (examples given in Table 2), developed by Sabine (1922)

$$n \text{ dB} = 10 \log \frac{1}{1-\bar{a}}, \tag{3}$$

where

n dB is the reduction in sound pressure level and
a the absorption coefficient.

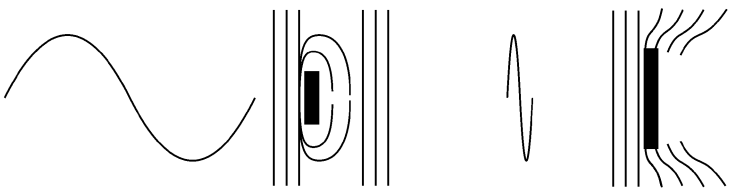


FIGURE 5 Diffraction and shading of sound waves.

TABLE 2 Absorption coefficient chart

Material	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Carpet	0.01	0.02	0.06	0.15	0.25	0.45
Concrete (sealed or painted)	0.01	0.01	0.02	0.02	0.02	0.02
Theater seats (wooden, empty)	0.03	0.04	0.05	0.07	0.08	0.08
Theater seats (wooden, fully occupied)	0.50	0.30	0.40	0.76	0.80	0.76
Glass (6 mm plate, large pane)	0.18	0.06	0.04	0.03	0.02	0.02
Fiberglass board (25 mm thick)	0.06	0.20	0.65	0.90	0.95	0.98
Fiberglass board (100 mm thick)	0.99	0.99	0.99	0.99	0.99	0.99
Plasterboard (12 mm in suspended ceiling grill)	0.15	0.11	0.04	0.04	0.07	0.08
Metal deck (perforated channels, 75 mm batts)	0.73	0.99	0.99	0.89	0.52	0.31
People, adults (per 1/10 person)	0.25	0.35	0.42	0.46	0.50	0.50
Ventilating grilles	0.30	0.40	0.50	0.50	0.50	0.40
Water or ice surface	0.01	0.01	0.01	0.02	0.02	0.03

In an area fully occupied by audience on wooden theater seats, the absorption coefficient is around 0.76 for high and mid frequencies, whereas an area with wooden and empty theater seats has an absorption coefficient below 0.1. The attendance of audience may dramatically change the absorption coefficient for at least one boundary (floor or seating areas) and therefore the impulse response of the venue will vary. The absorption coefficient of the various portions of boundaries are averaged arithmetically by using the respective weighting factors.

$$\bar{\alpha} = \frac{1}{S} \sum S_i \alpha_i, \tag{4}$$

where

S_i is the area of the portion of the boundary, and
α_i is the absorption coefficient of the boundary portion.

All boundary reflections of a room will result in reverberation. There are several equations for calculating the reverb time based on the absorptiveness of a given room. The Sabine formula (5) gives the best correspondence with $\bar{\alpha} < 0.2$, the Eyring formula (6) (Eyring, 1930) should be used for well-behaved rooms having $\bar{\alpha} > 0.2$, and the Millington–Sette formula (7) (Millington, 1932; Sette, 1932) should be used for rooms with a wide variety of absorption coefficients.

$$T = \frac{0.16V}{S\bar{\alpha} + 4mV}, \tag{5}$$

$$T = \frac{0.16V}{-S\ln(1-\bar{\alpha}) + 4mV}, \tag{6}$$

$$T = \frac{0.16V}{\sum -S_i \ln(1-\bar{\alpha}_i) + 4mV}, \tag{7}$$

where

T is the decay time in seconds for 60 dB level reduction,
S_i are the surface areas in m^2,
V is the volume in m^3,
$\bar{\alpha}_i$ are the average absorption coefficients of the individual surfaces,
m is the attenuation constant of air which is responsible for the attenuation of sound traveling through air. For small rooms, the term $4mV$ related to air absorption can be neglected.

When a continuous sound hits a reflective boundary (hard boundary surface) the reflected sound wave combines with subsequent incoming waves. In case the wave crests, i.e., maximum pressure coincide, they combine and reinforce each other. The troughs, i.e., minimum pressure also combine. This results in a stationary pattern in air. It will consist of nodes (low pressure, but high velocity of air) and antinodes (high pressure, but low velocity of air), resulting in very soft and very loud sound areas, respectively. The same effect will occur for integral multiples of the frequency. However, the effect of standing waves is most prominent for low frequencies. Absorption materials have to be placed at the nodes (lambda 1/4 distance from the boundary) for best absorption results. Further ideas for room design for minimizing stationary patterns can be found in Cox et al. 2004.

A good sound system shall deliver a sound as clear as possible and it should allow as much control over the sound as possible. Therefore, the critical distance has to be maximized. It is the distance from the loudspeaker, where the sound pressure level of the direct sound and the sound pressure level of the reverberant sound are equal. When moving further away from the source toward the respective boundary, the reverberant sound field will predominate. The critical distance can be enlarged when using directional loudspeakers instead of omnidirectional ones. Further techniques to maximize the critical distance are discussed in Davis and Jones 1989.

5 • Sound Outdoors

Atmospheric factors of interest especially for outdoor venues are wind, temperature, and humidity. Wind effects are divided into two classes: velocity effects and gradient effects.

Wind velocity: Crosswind can change the direction of sound as it adds a velocity vector to the propagated sound wave. Sound will appear to come from a different location. Tail or head wind (related to the PA) will propagate greater or less distances respectively.

Wind velocity gradient: This effect occurs in case there are two different layers of air next to each other and air layers are moving with different speeds. This might happen in case the audience is shielded by a wall or a barrier of trees. In case wind passes through a velocity gradient, velocity will add a vector to the propagated sound wave. As shown in Figure 6, let's assume that the upper layer is with higher velocity wind than the lower layer. Then sound which is propagated against the wind will be refracted upward, whereas sound propagated with the wind, will be refracted downward. The actual effect of wind is minimal as the wind speed is negligible compared to the speed of sound.

Temperature gradients: The speed of sound is dependent on the temperature of the air[14]

$$c = (331.4 + 0.607T), \tag{8}$$

where

c is the sound speed in meters per second,
T is the temperature in degrees Celsius.

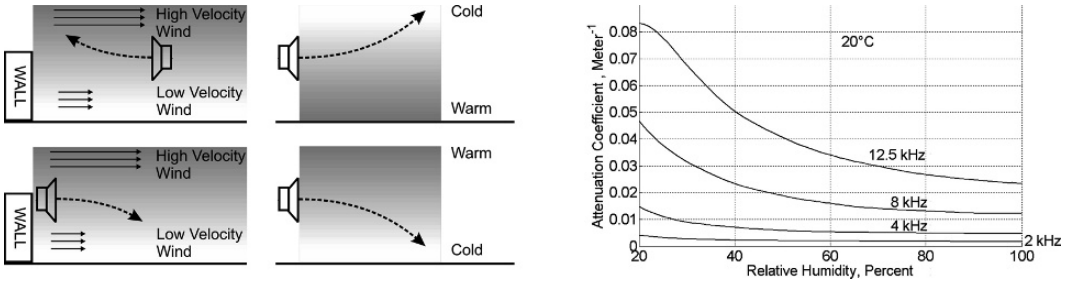


FIGURE 6 Effect of wind and temperature gradient, and absorption of sound in air, values taken from Harris (1966).

[14]Linearized formula shows an error of ±0.2% within +40° and −20°C .

When an acoustic wave passes through a temperature gradient, it will bend toward the cooler air, as shown in Figure 6.

This phenomenon occurs mainly during the hot daytime (sound will bend upward) and in the evening hours (sound will bend downward). This fact has to be noticed when tuning a PA during the afternoon for an evening concert.

Humidity: The medium which sound waves travel through will absorb the energy from the sound and attenuate it. The attenuation of sound in air is affected by the relative humidity. However, the effect will occur only for frequencies above 2 kHz.

6 • Measurement Methods

The measurement of a PA system should always be an essential part of the initial operation, whether it is just a simple check with well trained ears or a set of measured impulse responses. The goal of these measurements is always the same:

- Provide a homogenous sound field for the whole listening area
- Optimize the frequency response at all positions in the listening area
- Reach the maximum gain-before-feedback ratio
- Compensate delays
- Avoid reflections

Some of the commonly used measurement software are MLSSA, WINMLS, SMAART, ETF5, EASERA, Aurora, or the WaveCapture series. Besides measuring impulse responses, these softwares are capable of estimating most of the following parameters: SPL, RT, RASTI, STI, C80, and D50. The measurement systems use special measurement signals like maximum length sequences (MLS), swept sines, pink noise, impulses, or wobbles. It has to be made sure that the PA system and the ears of persons attending the measurement procedure can handle these signals without taking any damage.

An optimal measurement and tuning procedure may look as follows:

- Measure and compensate the delays
- Estimate the mean frequency response out of many measurement positions
- Transfer the filter function gained from the inverse mean frequency response into the sound system
- Measure again the frequency response, especially at former measurement positions with unwanted results, and tune the sound system again if necessary
- Measure the feedback path of the relevant microphones

- Tune the PA system with the results obtained by the feedback measurement
- In case of commercial applications, document the complete procedure, measurement results and steps taken to optimize the sound system

The whole idea of technical measurement has a major drawback: Ears do not correlate with the specs. Measurements may help finding faults or bad frequency shapes, but it is essential to do a final check of the sound system with an expert listener with trained ears.

7 • Special Issues on Signal Processing for PA Systems

There are some signal processing topics which are very closely related to PA systems and which are rarely found within other fields of acoustics:

- Loudspeaker time alignment
- Feedback reduction or gain-before-feedback boost
- Automatic maintenance of room equalization

The *loudspeaker time alignment* is necessary to avoid comb-filter effect sound. As shown in Figure 7, in case there are two sources with same signal and same amplitude but different delay time, the reinforcement and cancelation frequencies depend on the delay time between the arrival of the sound of the first source and the second source.

The cancelation occurs at the frequencies $f = 2x + 1/2t$, where $x = 0, 1, 2, 3, ..$ and t is the delay time in seconds. The reinforced frequencies occur at the frequencies $f = x/t$, where $x = 0, 1, 2, 3, \ldots$ and t is the delay time in seconds. The reinforced frequencies

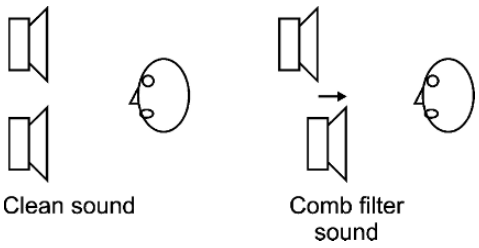


FIGURE 7 Clean sound and comb-filter sound caused by not-time-aligned loudspeakers.

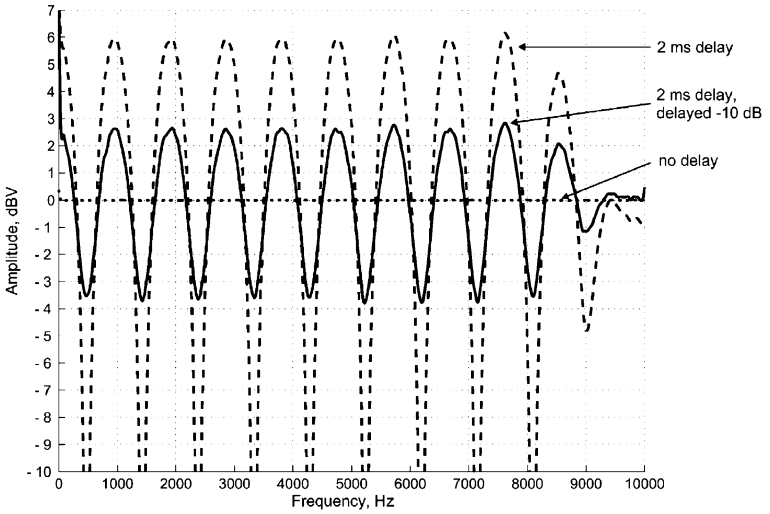


FIGURE 8 Frequency response of two loudspeakers without delay, with a delay of 2 ms and equal SPL, and with a delay of 2 ms, when the delayed loudspeaker is attenuated by 10 dB SPL.

increase in amplitude by 6 dB, while the out-of-phase frequencies cancel to $-\infty$ dB. In case the delayed loudspeaker is reduced in SPL compared to first loudspeaker, e.g., 10 dB as shown in Figure 8, the comb-filter effect will be reduced.

Comb-filter effects are unavoidable to some degree in most sound systems, and they cannot be corrected with equalization. Fortunately, most comb-filter problems can be minimized by carefully synchronizing the signals and reducing the amplitude of the delayed signal. Comb-filter effects occur mostly in speaker clusters, when different types of speakers are combined, e.g., longthrow and downfill or in case the speakers are not fixed properly within the cluster.

The delay time between main PA and delay line speaker as shown in Figure 9 will always be a compromise as the geometry will not allow an exact synchronization of signals for every location behind the delay line. For spoken word programs, the signal from the delay line speaker should arrive within 10 ms of the signal from the center cluster. For programs that are mostly music, one can allow a little more reverberation which means a little bit higher τ. Finally, comb-filter effects have to be reduced by time and level alignment.

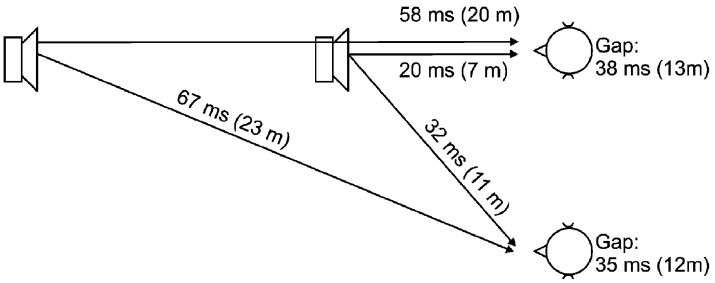


FIGURE 9 Problem of different delay times of two distributed loudspeakers related to the seating position.

For a first estimation one may use

$$\text{delay [ms]} = 1,000 \frac{\text{distance [m]}}{344[\frac{\text{m}}{\text{s}}]}. \tag{9}$$

The problem of *feedback reduction or gain-before-feedback boost* occurs for the house systems as well as for the monitoring systems. In case the gain of the feedback loop is ≥ 1 and the phasing of the system matches, the system gets instable. The positive feedback of the sound system limits the maximum gain. It is the intention to raise the point of self-excitation to higher gains without colorizing the sound too much. There are several techniques to battle this problem.

- The use of *directional sound sources* and directional microphones placed in geometric favorable positions;
- The use of *boundaries with high absorption coefficient* which will be hit directly by the amplified sound;
- A *slow-phase modulation* of the audio signal sent to the amplifier avoids a match of phases while the feedback loop is ≈ 1. This technique is only suitable for speech, not for music;
- The *frequency shifting* goes back to Schroeder (1961). After each run through the feedback loop, the frequency will be shifted by 1–3 Hz. Here as well, the phases do not match while the feedback loop is ≈ 1. The peaks of the feedback loop transfer function are canceled out by averaging. Also this approach is not suitable for musical signals as the instruments sound detuned for each run through the feedback loop;
- For onstage personal monitoring applications, one may *invert the signals* (constant phase shift of 180°) between two monitor wedges which point to a single performer. The ear is not able to detect a frequency shift between left and right ear, but the sound field distributed from the two monitor wedges around

the microphone cancels itself up to a certain degree. The comb-filter effect is negligible, as long as the actor stays within the sweet spot;

- Another approach for onstage feedback reduction is the *insertion of short delays* (e.g., 5–20 ms) for each monitor-send path. This reduces the danger of "hot spots" on stage but may cause timing problems for professional musicians, especially within the rhythm section. Here as well, the effect of comb filters is negligible within most setups.
- By *equalizing,* one may cut the peaks of the feedback loop transfer function. This can be done by manual setting of graphic or parametric peaking EQs, however it will colorize the sound. The automatic feedback reduction algorithms usually use notch filters between 1/5 and 1/80 octave. The general problems of equalizing the signal to increase the gain before feedback are discussed by McPeek (2000). There are two main approaches for feedback detection:

    a. *Level comparison*. If a certain frequency band (1/10–1/100 octave bandwidth) raises in volume to a more-or-less consistent level or disproportionate level, the algorithm assumes that there is a feedback within this frequency band. This algorithm fails for long constant pure tones, like a constant string sound.
    b. *Waveform analysis*. This approach is based on the fact that the waveform signature of a feedback in almost all instances differs quite much from the rest of the audio signal. The algorithm compares the signal with different feedback signatures trying to find correlations. In case the correlation is above a certain threshold, the algorithm assumes that a feedback related to the waveform is stored in the feedback signature database. This algorithm requires more processing power than the level comparison algorithm.

A good feedback reduction algorithm can provide an additional 6–9 dB of overall system gain before feedback.

The *automatic maintenance of room equalization* may be necessary for venues which change their frequency response remarkably with the attendance of audience or during the performance due to the change of humidity. Therefore, the automated EQ is connected to a measurement microphone, trying to keep the actual frequency response close to the predefined frequency response between the PA system and the measurement microphone. Therefore, the signal is bandpass-filtered related to the frequency to be analyzed. The EQ inserts

an inaudible measurement signal related to the frequency to be measured, which is typically 60 dB below the average program level. Meanwhile, a reciprocal filter is placed in the measurement microphone path, removing all program material, so that the reference signal is remaining (Sabine 2001).

References

Ahnert, W. and Steffen, F. (1999). *Sound reinforcement engineering, fundamentals and practice*. E & FN Spon, GB-London.

Alldrin, L. (1997). *The home studio guide to microphones*. Mix Books, CA-Emeryville.

Ballou, G. M. (2002). *Handbook for sound engineers, Third Edition*. Focal Press, MA-Boston.

Blauert, J. (2000). *Lecture notes, communication acoustics, only in German language, (original title: "Akustik 2, Kommunikationsakustik")*. Ruhr-University Bochum, D-Bochum.

Cox, T. J., Antonio, P. D., and Avis. M. R. (2004). Room sizing optimization at low frequencies. *Journal of the Audio Engineering Society*, 52(6):640–651.

Cross, C. (2005). www.beatles-discography.com.

Davis, D. and Davis, C. (1997). *Sound system engineering, Second Edition*. Focal Press, MA-Boston.

Davis, G. and Jones, R. (1989). *The sound reinforcement handbook*. Hal Leonard Corporation, WI-Milwaukee.

Eyring, C. F. (1930). Reverberation time in dead rooms. *Journal of the Acoustical Society in America*, 1:217–249.

Harris, C. M. (1966). Absorption of sound in air versus humidity and temperature. *Journal of the Acoustical Society in America*, 40(1): 148–159.

Huber, D. M. and Williams, P. (1922). *Professional microphone techniques*. Primedia Intertec Publishing Corporation, CA-Emeryville.

IEC60268-16:1998 (1998). Sound system equipment – part 16: Objective rating of speech intelligibility.

IEC61672-1:2002 (2002). Electroacoustic – sound level meters, part 1.

ISO3382:1997 (1997). Acoustics – measurement of the reverberation time of rooms with reference to other acoustical parameters.

McPeek, R. (2000). www.svconline.com.

Millington, G. (1932). A modified formula for reverberation. *Journal of the Acoustical Society in America*, 4:69–82.

Pawera, N. (2003). *Microphones in practical use, only in German language (original title: "Mikrofonpraxis, Tipps und Tricks für Bühne und Studio")*. PPV Medien GmbH, D-Bergkirchen.

Sabine (2001). *Real Q2 operating guide, Version 7.5*. Sabine, Inc., FL-Alachua.

Sabine, W. C. (1922). *Collected papers in acoustics*. Harvard University Press, MA-Cambridge.

Schroeder, M. R. (1961). Improvement of feedback stability of public address systems by frequency shifting. *In audio engineering society preprint, 13th annual meeting, New York*.

Sette, W. J. (1932). A new reverberation time formula. *Journal of the Acoustical Society in America*, 4:193–210.

Stark, S. H. (1989). *The sound reinforcement handbook, Second Edition*. Mix Books, CA-Vallejo.

Trubit, R. (2000). *Compact mixers, Second Edition*. Hal Leonard Corporation, WI-Milwaukee.

Urban, M. Heil, C. and Bauman, P. (2003). Wavefront sculpture technology. *Journal of the Acoustical Society in America*, 51(10): 912–932.

Yaxley, K. (2003). www.historyofpa.co.uk.

35

Beamforming for Speech and Audio Signals

Walter Kellermann

Multimedia Communications and Signal Processing, University Erlangen-Nuremberg, Erlangen, Germany, e-mail: wk@lnt.de

If microphone arrays instead of a single microphone are employed for sampling acoustic wavefields, signal processing of the sensor data can exploit the spatial diversity to better detect or extract desired source signals and to suppress unwanted interference. Beamforming represents a class of such multichannel signal processing algorithms and suggests a spatial filtering which points a beam of increased sensitivity to desired source locations while suppressing signals originating from all other locations. While beamforming techniques are also extensively used in other areas, e.g. in underwater acoustics, ultrasound diagnostics, and radio communications [1–3], the treatment is concentrating here on wideband acoustic signals in the audio frequency range.

In the following, first the basic concept is introduced and then several classes of beamforming techniques are outlined: Common data-independent beamformers form the most popular class for all areas of beamforming applications, whereas differential microphone arrays are especially attractive for wideband audio signals. Statistically optimum beamformers are able to better match given noise and interference scenarios and to follow its changes over time. Finally, blind source separation of convolutive mixtures is linked to beamforming. Please note that in view of the long tradition and the wealth of literature, relatively few

original references are cited, and often survey papers are referenced instead, which provide the links to the original sources.

1 • The Basic Concept

In a general scenario as shown in Figure 1, an acoustic wavefield is sensed by M microphones at positions $\mathbf{r}_\mu$, $\mu = 1, \ldots, M$. The continuous-time microphone signals $x_{\mathrm{c},\mu}(\mathbf{r}_\mu, t)$ will in general capture components from a point source of interest, $s(t)$, at position $\mathbf{r}_\mathrm{S}$ radiating a wavefield $s(\mathbf{r}, t)$, and additive noise components $n_\mu(t)$, which may represent coherent interferers, a diffuse noise field, or sensor noise. After A/D conversion, M sampled microphone output signals $x_\mu(k)$ are available as inputs for the beamformer, with k as the discrete time in integer multiples of the sampling interval $T_\mathrm{s} = 1/f_\mathrm{s}$ (sampling frequency f_s). Considering a generic beamformer for real-valued acoustic signals, each digitized microphone signal $x_\mu(k)$ is filtered by a possibly time-variant linear filter with real-valued impulse response $w_\mu(i,k)$:

$$y(k) = \sum_{\mu=1}^{M} \sum_{i=-\infty}^{\infty} w_\mu(i,k) x_\mu(k-i). \tag{1}$$

As a typical characterization, the response of the beamformer to a harmonic plane wave [4,5]

$$s(\mathbf{r}, t) = S_0 \cdot \mathrm{e}^{\mathrm{j}(\omega t - \mathbf{k}\cdot\mathbf{r})} \tag{2}$$

is considered, where S_0 is the scalar pressure amplitude, ω is the temporal frequency, and the wave vector $\mathbf{k}$ describes the direction of propagation of the wave and its spatial frequency $|\mathbf{k}| = 2\pi/\lambda = \omega/c$ (with λ representing the wavelength and c the sound velocity,

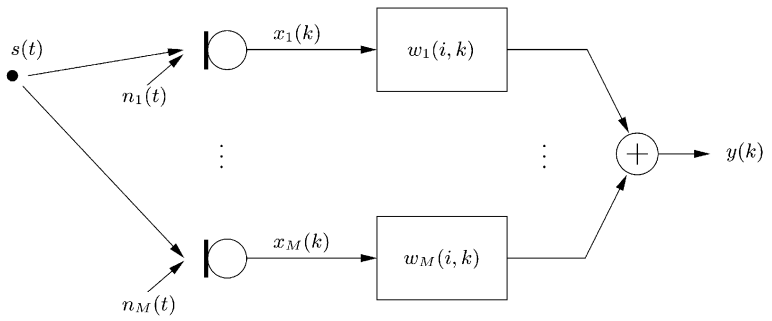


FIGURE 1 Scenario for acoustic beamforming.

respectively). Inserting Eq. (2) for the microphone positions $\mathbf{r}_\mu$, the array output reads

$$y(k) = S_0 \cdot \mathrm{e}^{\mathrm{j}\omega k T_\mathrm{s}} \cdot \sum_{\mu=1}^{M} \sum_{i=-\infty}^{\infty} w_\mu(i,k) \mathrm{e}^{-\mathrm{j}(\omega i T_\mathrm{s} + \mathbf{k}\cdot\mathbf{r}_\mu)}$$
$$:= S_0 \cdot \mathrm{e}^{\mathrm{j}\omega k T_\mathrm{s}} \cdot W(\mathbf{k}, \omega, k), \tag{3}$$

with $W(\mathbf{k}, \omega, k)$ representing the – possibly time-varying – *array response* to a harmonic plane wave with wave vector $\mathbf{k}$.

For the wavefield of a point source S arriving at the microphone array, the plane wave assumption (*far-field assumption*) is said to be valid if the aperture width, i.e. the geometric extension of the array as seen by the impinging wavefront, R, and the distance from the source to the array centre, r_S, fulfill $|r_\mathrm{S}| > 2R^2/\lambda$.

Essentially, most beamforming algorithms can be viewed as techniques that try to extract a sampled and delayed but undistorted version of the signal of interest $s(t)$ from $y(k)$ while at the same time suppressing the undesired noise contributions $n_\mu(t)$.

2 • Data-Independent Beamforming

The most intuitive and simple approach to beamforming is to coherently add the desired components in the observations expecting that thereby the undesired noise components cancel out to a certain extent. Assuming free-field conditions and that only the desired source S located at $\mathbf{r}_\mathrm{S}$ is active, the continuous-time sensor signals can be written as $x_{\mathrm{c},\mu}(\mathbf{r}_\mu, t) = a_\mu s(t - \tau_\mu)$, where a_μ is a constant gain and τ_μ is a delay, both depending on the Euclidean distance between the source and the μth microphone $||\mathbf{r}_\mu - \mathbf{r}_\mathrm{S}||$. For coherent superposition of the desired source components, the beamforming filters are time-invariant and only have to equalize the delays $\tau_\mu = ||\mathbf{r}_\mu - \mathbf{r}_\mathrm{S}||/c$ (*delay-and-sum beamformer*). If all the relative delays between the various sensors are integer multiples of the sampling interval T_s, then all filter impulse responses reduce to time-shifted unit impulses $w_\mu(i) = \delta(i - i_\mu)$. Noninteger delays, however, require interpolation sequences as impulse responses, which are typically approximations of $w_\mu(i) = \mathrm{sinc}\left[\pi(i + (\tau_\mu - \tau')/T_\mathrm{s})\right]$ [6], where $\mathrm{sinc}(x) = (\sin\, x)/x$ and $\tau' \geq \max_\mu \{\tau_\mu\}$ is a necessary common delay if causality of all $w_\mu(i)$ must be assured.

If all the desired source components are aligned perfectly, if the gains a_μ are equal, and if the noise components $n_\mu(t)$ have the same level at each microphone and are mutually orthogonal, then the *array gain*, i.e. the ratio of desired source power relative to

the noise power compared to the ratio at a single microphone, of the *delay-and-sum beamformer* is equal to the number of microphones M. If it is assumed that the noise is due to a spherically isotropic noise field, then the logarithmic array gain, $10\log_{10} M$, is termed *directivity index* and corresponds to the logarithmic ratio of the source power from the steering direction relative to the noise power integrated over all directions [1–3].

As a widely-used special case, a linear equispaced array is considered as shown in Figure 2, where the microphone positions are given by multiples of the spacing d along the x axis. For a planar wave according to Eq. (2), the $M = 2K+1$ sampled microphone signals $x_\kappa(k)$, $\kappa = -K, \ldots, K$, are given by

$$x_\kappa(k) = s(\mathbf{r}_\kappa, kT_\mathrm{s}) = S_0 \cdot \exp\left[\mathrm{j}\omega\left(kT_\mathrm{s} + \kappa\frac{d}{c}\cos\Theta\right)\right]. \tag{4}$$

Assuming now that the time-invariant beamforming filters realize constant gains w_κ and causal delays $\tau_\kappa = \tau' + \kappa\frac{d}{c}\cos\Theta_0$, the output of the beamformer reads

$$y(k) = S_0 \cdot \mathrm{e}^{\mathrm{j}\omega(kT_\mathrm{s}-\tau')} \sum_{\kappa=-K}^{K} w_\kappa \exp\left[\mathrm{j}\kappa d\frac{\omega}{c}(\cos\Theta - \cos\Theta_0)\right]. \tag{5}$$

If for simplicity, $\tau' = 0$ (allowing noncausal delays), the *array response* reads with $u = \cos\Theta$

$$W(u,\omega) = \sum_{\kappa=-K}^{K} w_\kappa \exp\left[\mathrm{j}\kappa d\frac{\omega}{c}(u - u_0)\right] \tag{6}$$

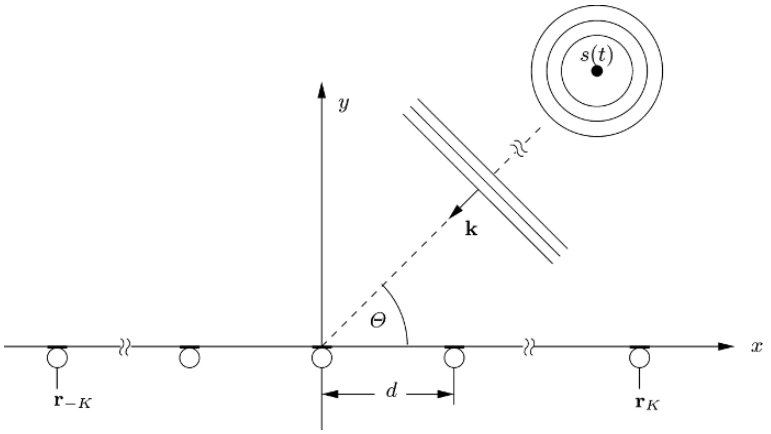


FIGURE 2 Point source in the far field and a linear equispaced array with $M = 2K+1$ sensors.

and shows that frequency ω and spacing d can be interchanged: Doubling the frequency ω and simultaneously halving the spacing d yield the same array response. In Figure 3, the frequency-independent gain can be seen for the mainlobe at $u = 0$ ($\Theta_0 = \pi/2$), but periodic repetitions of the mainlobe appear at $u = \pm 1$ for frequencies above 5,000 Hz. This is due to the periodicity of the array response in u with $W(u, \omega) = W(u + \nu \cdot \frac{2\pi c}{\omega d}, \omega)$ ($\nu \in \mathbb{Z}$) and is called *spatial aliasing*. To avoid it for all steering angles Θ_0, $u \pm \frac{2\pi c}{\omega d}$ must be outside the *visible range* of u, $|u| \leq 1$, which requires $\frac{2\pi c}{\omega d} > 2$, i.e. $d < \lambda/2$.

With regard to the choice of w_κ, it is interesting to note that the array response Eq. (6) can be equated to an FIR filter with frequency response

$$H(\omega) = \sum_{\kappa=-K}^{K} h_\kappa \mathrm{e}^{-\mathrm{j}\omega\kappa T_\mathrm{s}}. \tag{7}$$

While Eq. (7) implies a filtering along a uniformly sampled time axis (at multiples of T_s), Eq. (6) implies FIR filtering along a uniformly sampled spatial coordinate axis (at multiples of d). This means that for a given frequency, the magnitude of the array response (*beampattern*) as a function of u directly corresponds to the magnitude response of the corresponding FIR filter. As a consequence, besides the popular choices of a rectangular

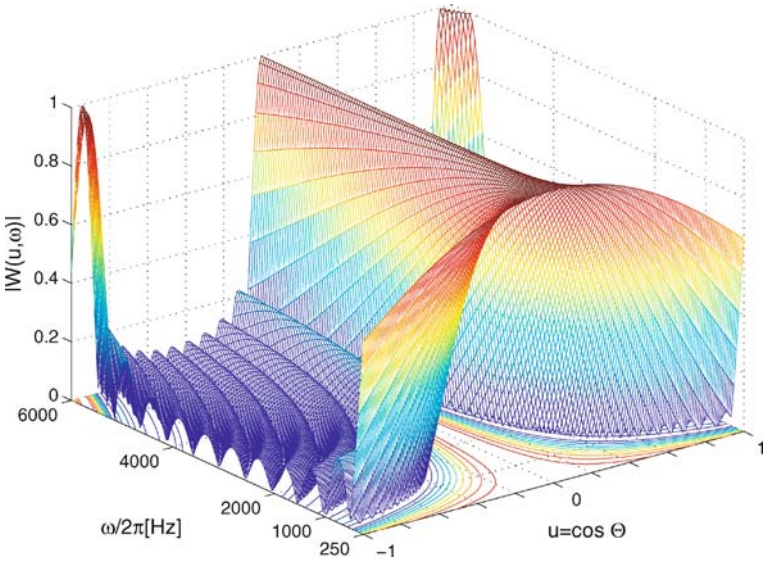


FIGURE 3 Array response (*beam pattern*) for a linear equispaced array with $M = 11$ sensors, $d = 6\,\mathrm{cm}$, $w_\kappa = 1/M$.

window or Dolph-Chebyshev windows [7], many other window types known from FIR filtering or spectrum analysis [8] are directly applicable to beampattern design. For beamforming, the various windows define different tradeoffs between mainlobe width and sidelobe suppression in the angular domain, just as they do in the frequency domain for spectrum analysis.

Obviously, the filters w_κ are not limited to realizing constant gains and delays only, but may be general linear filters (*filter-and-sum beamforming*), so that the response of a time-invariant beamformer for a linear equispaced array to a plane wave reads in accordance with Eq. (1):

$$W(u,\omega) = \sum_{\kappa=-K}^{K} \sum_{i=-\infty}^{\infty} w_\kappa(i) \exp\left[\mathrm{j}(\omega\kappa\frac{d}{c}(u-u_0) - \omega i T_\mathrm{s})\right]. \quad (8)$$

From this it is obvious that the design of a filter-and-sum beamformer for plane waves impinging on linear equispaced arrays is equivalent to a 2D-FIR filter design. As a typical design goal with wideband acoustic signals, filter-and-sum beamformers will aim at a frequency-independent beamwidth for the mainlobe steered towards the desired source. The according design methods [9–11] can well overcome the narrowing of the mainlobe at high frequencies (cf. Figure 3). However, regarding low frequencies, the minimum beamwidth for any given frequency, as, e.g. measured by the first zero in the beampattern off the mainlobe, is always obtained if, at this frequency, the gain for all microphone signals is equal. Therefore, the widening of the beam towards low frequencies can only be overcome if the aperture size is increased, that is, the spacing d is increased, or more microphones are added. For audio and speech applications, this led to the concept of nested arrays, which foresee one microphone array per octave while doubling the spacing d from one octave to the next [4,12].

Although linear arrays are still prevalent for practical applications, the concept of *delay-and-sum* beamforming has been extended from linear arrays to two-dimensional (2D) planar arrays very early [4], and multiple 2D arrays were proposed to realize 3D beamforming [4,13].

3 • Differential Microphone Arrays

The idea of differential microphone arrays is to approximately measure spatial derivatives of the sound pressure by pressure differences between closely spaced locations in the sound field instead of the pressure itself, and thereby obtain better directivity with small aperture arrays even for low frequencies, which is

not possible for ordinary *filter-and-sum* beamformers with larger microphone spacings. For illustration of the principle, the linear array as given by Eq. (5) is considered for only two sensors, $\kappa = 0, -1$, and opposite weights $w_0 = 1, w_{-1} = -1$. Then, the output for a plane wave is given by

$$y(k) = S_0 \cdot \mathrm{e}^{\mathrm{j}\omega(kT_s - \tau')} \left(1 - \exp\left[-\mathrm{j}d\frac{\omega}{c}(\cos\Theta - \cos\Theta_0)\right]\right). \quad (9)$$

If the spacing d is now much smaller than the wavelength, $d \ll \lambda$, the truncated Taylor series $e^x \approx 1 + x$ can be used and the array response reads for $\tau' = 0$:

$$W(u, \omega) = \mathrm{j}\omega\frac{d}{c}(\cos\Theta - \cos\Theta_0) = \mathrm{j}\omega\frac{d}{c}(u - u_0). \quad (10)$$

Obviously, this *first-order differential array* exhibits the same first-order highpass characteristic for all angles of arrival Θ, and has frequency-independent zeroes at $\Theta = \pm\Theta_0$. For any given frequency, the magnitude of the array response is maximum for $\Theta = 0$ or $\Theta = \pi$ depending on the sign of $\cos\Theta_0$, that is, the maximum sensitivity coincides always with the array axis (*endfire array*). The directivity index assumes its maximum of 6.02 dB for $u_0 = \cos 109^0 \approx 1/3$ (hypercardioid, see [14]) and is greater than for a delay-and-sum beamformer for all values of u_0 in the visible range, which leads to the notion of *superdirective beamforming*. The highpass characteristic can be easily equalized by a matching lowpass filter applied to $y(k)$. However, it should be noted that, thereby, spatially uncorrelated low-frequency noise is amplified accordingly. Higher-order differential arrays using $M > 2$ microphones for approximating higher-order derivatives have been investigated in [5,14] where it is shown that the maximum possible directivity index is $10\log_{10} M^2$ for a differential array of order $M - 1$. However, the order of the equalizing lowpass filters has to increase accordingly, which aggravates the *noise susceptibility* of the array and thereby limits the practical use of this concept currently to orders less than four, especially as any mismatch between the individual spatio-temporal transfer functions of the microphones has an equivalent effect as additive noise [14].

4 • Statistically Optimum Beamformers

In order to optimally suppress all interferers and noise sources, it is highly desirable that, especially for broadband signals and complex noise scenarios as typical for audio and speech applications, beamformers match the statistics of the involved signals. Two classes are in common use: One is based on minimizing a

squared error criterion based on the difference $e(k)$ between the beamformer output $y(k)$ and a reference signal $y_{\mathrm{ref}}(k)$. Minimizing the expectation of the mean squared error $\mathcal{E}\{e^2(k)\}$ produces a *minimum mean squared error* (MMSE) estimate for the optimum beamforming filters [1,3,15,16]. In its deterministic form, it is often called *least-squares error (LSE) beamforming* if a weighted sum of squared error samples is minimized instead of an estimated mean squared error. This concept can be viewed as a multichannel Wiener Filter which not only accounts for the second-order statistics within each channel but also optimizes with regard to the coherence between the channels. In practice, the main problem in this context is to obtain a useful reference signal $y_{\mathrm{ref}}(k)$ which may represent either the desired signal or the noise component in the beamformer output [15–17].

The second class minimizes the squared output of the beamformer subject to a linear constraint which essentially tries to preserve the integrity of the desired signal. This approach is called *linearly constrained minimum variance (LCMV) beamforming* [18]. A special case is the *minimum variance distortionless response* (MVDR) *beamformer* which assures a distortionless response for a predefined steering direction Θ_0 of the array and is especially important in acoustic signal processing [15,17]. In Figure 4, an example for an MVDR beamformer is shown: A broadband interferer impinging from $\Theta = 0.37\pi$ is strongly rejected except for the very low frequencies, whereas the desired source assumed at $\Theta = 0.5\pi$ will be left undistorted.

For implementing the MVDR beamformer, the *generalized sidelobe canceller (GSC)* [19] represents an efficient choice, as here, the constraint is assured by a generally simple data-independent structure (a fixed beamformer and a so-called blocking matrix), so that the data-dependent statistical optimization reduces to an unconstrained interference canceller. This also gives way to very efficient adaptive implementations that are able to continuously follow nonstationary signals and time-varying acoustic environments. As, however, the GSC is very vulnerable to movements of the desired sources, robust versions with adaptive blocking matrices have been developed for its application to speech and audio [17,20].

The filter characteristics $w_\mu(i,k)$ of the above statistically optimum beamformers are determined by the power spectral densities of the desired source and the noise sources, respectively, and by the spatial and temporal coherence of the noise. Typically, the performance in terms of noise and interference suppression will exceed that of a fixed filter-and-sum beamformer especially for coherent interferers, as its flexibility allows for

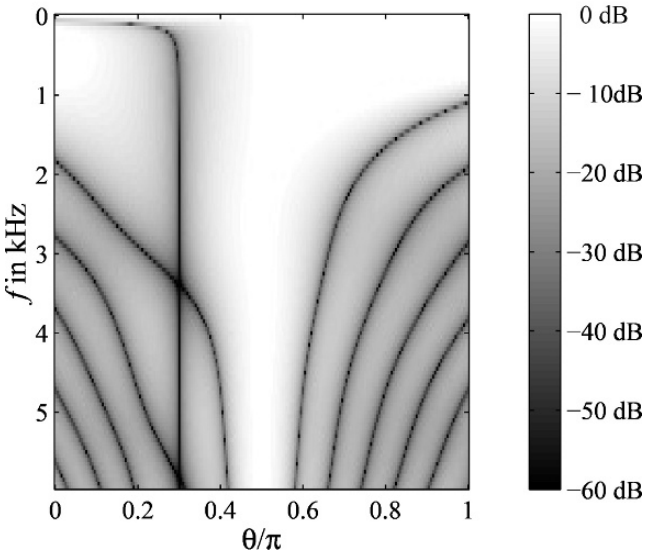


FIGURE 4 Array response (*beam pattern*) of an MVDR beamformer for a linear equispaced array with $M = 9$ sensors, $d = 4\,\text{cm}$ to an interferer emitting white noise from $\Theta = 0.3\pi$ (with permission from [17]).

signal-dependent suppression of certain spatial regions. For low frequencies, where the sensor spacing is relatively small compared to the wavelength, the optimum solutions will often be similar to those for differential arrays, so that constraints on the noise susceptibility may be advisable, whereas for high frequencies the filters will approach *filter-and-sum* characteristics [15,17].

5 • Blind Source Separation as Blind Beamforming

As a relatively recent development, blind source separation (BSS) has been introduced to acoustic applications, where it can be understood as a kind of blind beamforming [21,22]. As opposed to the techniques described in the previous paragraphs, it does not presume any knowledge about the spatial positions of desired sources and interferers (*blindness*), but only assumes statistical independence or at least uncorrelatedness between several sources. In a typical quadratic configuration, with M acoustic sources $s_\nu(t)$ and M microphones, a so-called demixing system will produce M outputs $y_\nu(k)$, and each output should contain only one of the source signals. For each output, the demixing is a linear time-variant multichannel filtering operation as described

by Eq. (1). This filtering can be interpreted as an adaptive beamformer, which treats one of the sources as desired source and tries to suppress all other competing sources as unwanted interferers. For identifying the optimum demixing filters, cost functions are minimized, which capture statistical dependencies between all M outputs of the separation system, so that the optimum demixing matrix produces M mutually independent or at least uncorrelated outputs [23–25].

6 • Concluding Remarks

In this brief treatment of beamforming techniques for acoustic wideband signals such as speech and audio, fundamental concepts have been emphasized and recent advanced developments have only been outlined. The proportions are meant to reflect current practical importance; however, the author is convinced that statistically optimum adaptive beamformers and blind beamforming algorithms will gain more importance in the near future.

References

1. B.D. Van Veen and K.M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–24, April 1988.
2. D.H. Johnson and D.E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Prentice Hall, Englewood Cliffs, NJ, 1993.
3. H.L. van Trees. *Optimum Array Processing*. Wiley, New York, NY, 2002.
4. J.L. Flanagan, J.D. Johnston, R. Zahn, and G.W. Elko. Computer-steered microphone arrays for sound transduction in large rooms. *Journal of Acoustical Society of America*, 78(5):1508–1518, November 1985.
5. G. Elko. Microphone array systems for hands-free telecommunication. *Speech Communication*, 20:229–240, 1996.
6. T. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine. Splitting the unit delay. *IEEE Signal Processing Magazine*, 13(1):34–60, January 1996.
7. C.L. Dolph. A current distribution for broadside arrays which optimizes the relationship between beamwidth and sidelobe level. *Proceedings of the IRE*, 34:335–348, June 1946.
8. F.J. Harris. Use of windows for harmonic analysis. *Proceedings of the IEEE*, 66(1):51–83, January 1978.
9. M. Goodwin. Constant beamwidth beamforming. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, I-169 to I-172, Minneapolis, MN. IEEE, April 1993.

10. T. Chou. Frequency-independent beamformer with low response error. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 2995–2998, Detroit, MI. IEEE, May 1995.
11. D.B. Ward, R.A. Kennedy, and R.C. Williamson. Constant directivity beamforming. In M.S. Brandstein and D. Ward, editors, *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 13, pp. 3–18. Springer, Berlin, May 2001.
12. J.L. Flanagan, D.A. Berkley, G.W. Elko, J.E. West, and M.M. Sondhi. Autodirective microphone systems. *Acustica*, 73:58–71, 1991.
13. H. Silverman, W.R. Patterson, J.L. Flanagan, and D. Rabinkin. A digital processing system for source location and sound capture by large microphone arrays. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 251–254, Munich, Germany. IEEE, April 1997.
14. G. Elko. Superdirectional microphone arrays. In S. Gay and J. Benesty, editors, *Acoustic Signal Processing for Telecommunication*, pp. 181–238. Kluwer, Dordrecht, 2000.
15. J. Bitzer and K.U. Simmer. Superdirective microphone arrays. In M.S. Brandstein and D. Ward, editors, *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 13, pp. 19–38. Springer, Berlin, May 2001.
16. S. Doclo and M. Moonen. GSVD-based optimal filtering for multi-microphone speech enhancement. In M.S. Brandstein and D. Ward, editors, *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 6, pp. 111–132. Springer, Berlin, May 2001.
17. W. Herbordt and W. Kellermann. Adaptive beamforming for audio signal acquisition. In J. Benesty and Y. Huang, editors, *Adaptive Signal Processing: Application to Real-World Problems*, chapter 6, pp. 155–194. Springer, Berlin, January 2003.
18. O.L. Frost. An algorithm for linearly constrained adaptive array processing. *Proceedings of the IEEE*, 60(8):926–935, August 1972.
19. L.J. Griffiths and C.W. Jim. An alternative approach to linear constrained adaptive beamforming. *IEEE Transactions on Antennas and Propagation*, 30(1):27–34, January 1982.
20. O. Hoshuyama and A. Sugiyama. A robust adaptive beamformer with a blocking matrix using constrained adaptive filters. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pp. 925–928. IEEE, 1996.
21. J.-F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *IEE Proceedings-F*, 140(6):362–370, December 1993.
22. L. Parra and C. Fancourt. An adaptive beamforming perspective on convolutive blind source separation. In G. Davis, editor, *Noise Reduction in Speech Applications*. CRC Press LLC, 2002.
23. J.-F. Cardoso. Blind signal separation: Statistical principles. *Proceedings of the IEEE*, 86(10):2009–2025, October 1998.
24. L. Parra and C. Spence. Separation of non-stationary natural signals. In S. Roberts and R. Everson, editors, *Independent Components*

Analysis, Principles and Practice, pp. 135–157. Cambridge University Press, 2001.

25. H. Buchner, R. Aichner, and W. Kellermann. Blind source separation for convolutive mixtures: A unified treatment. In Y. Huang and J. Benesty, editors, *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, chapter 10, pp. 255–296. Kluwer, Dordrecht, 2004.

36

Digital Audio Recording Formats and Editing Principles

Francis Rumsey

Institute of Sound Recording, University of Surrey, Guildford, UK

This chapter summarises digital audio recording formats and the principles of digital audio editing. A brief overview of the most common storage formats is given followed by an introduction to audio editing technology and techniques. Digital audio conversion and encoding principles are covered elsewhere in this book and will not be described here.

1 • Digital Tape Recording

There are still a number of dedicated digital tape recording formats in existence, although they are being superseded by computer-based products that use removable disks or other mass storage media. Tape has a relatively slow access time because it is a linear storage medium. However, a dedicated tape format can easily be interchanged between recorders, provided another machine operating to the same standard can be found. Disks, on the other hand, come in a very wide variety of sizes and formats. Even if the disk fits a particular drive it may not be possible to access the audio files thereon, owing to the multiplicity of levels at which compatibility must exist between systems before interchange can take place.

1.1 Background to Digital Tape Recording

When commercial digital audio recording systems were first introduced in the 1970s and early 1980s, it was necessary to employ recorders with sufficient bandwidth for the high data rates involved (a machine capable of handling bandwidths of a few megahertz was required). Analogue audio tape recorders were out of the question because their bandwidths extended only up to around 35 kHz at best, so video tape recorders (VTRs) were often utilised because of their wide recording bandwidth. PCM (pulse code modulation) adaptors (Doi et al. 1978) converted digital audio data into a waveform which resembled a television waveform, suitable for recording on to a VTR. The Denon Company of Japan developed such a system in partnership with the NHK broadcasting organisation, and they released the world's first PCM recording on to LP (long-playing vinyl disk) in 1971. In the early 1980s, devices such as Sony's PCM-F1 became available at modest prices, allowing 16-bit, 44.1-kHz digital audio to be recorded on to a consumer VTR, resulting in widespread proliferation of stereo digital recording.

Dedicated open-reel digital recorders using stationary heads were also developed. High-density tape formulations were then manufactured for digital use, and this, combined with new channel codes, improvements in error correction and better head design, led to the use of a relatively low number of tracks per channel, or even single-track recording of a given digital signal, combined with playing speeds of 15 or 30 in. per second. Dedicated rotary-head systems, not based on a VTR, were also developed – the R-DAT format being the most well known.

Digital recording tape is thinner (27.5 μm) than that used for analogue recordings; long-playing times can be accommodated on a reel, but also thin tape contacts the machine's heads more intimately than does standard 50-μm thickness tape which tends to be stiffer. Intimate contact is essential for reliable recording and replay of such a densely packed and high-bandwidth signal.

1.2 Overview of Digital Tape Formats

There have been a number of commercial recording formats over the last 20 years, and only a brief summary will be given here of the most common ones.

Sony's PCM-1610 and PCM-1630 adaptors [11] dominated the CD-mastering market for a number of years, although by today's standards they use a fairly basic recording format and rely on 60 Hz/525 line U-matic cassette VTRs. The system operated at a sampling rate of 44.1 kHz and used 16-bit quantisation, being designed specifically for the making of tapes to be turned into CDs. Recordings made in this format could be electronically

edited using the Sony DAE3000 editing system, and the playing time of tapes ran up to 75 min using a tape specially developed for digital audio use.

The R-DAT or DAT format (Ishida et al. 1985) is a small stereo, rotary-head, cassette-based format offering a range of sampling rates and recording times, including the professional rates of 44.1 and 48 kHz. Originally, consumer machines operated at 48 kHz to avoid the possibility for digital copying of CDs, but professional versions became available which would record at either 44.1 or 48 kHz. Consumer machines will record at 44.1 kHz, but usually only via the digital inputs. DAT is a 16-bit format, but has a non-linearly encoded long-play mode as well, sampled at 32 kHz. Truly professional designs offering editing facilities, external sync and IEC-standard timecode have also been developed. The format became exceptionally popular with professionals owing to its low cost, high performance, portability and convenience. Various non-standard modifications were introduced, including a 96-kHz sampling rate machine and adaptors enabling the storage of 20-bit audio on such a high sampling rate machine (sacrificing the high sampling rate for more bits). The IEC timecode standard for R-DAT was devised in 1990. It allows for SMPTE/EBU timecode of any frame rate to be converted into the internal DAT "running-time" code and then converted back into any SMPTE/EBU frame rate on replay.

The Nagra-D recorder was designed as a digital replacement for the world-famous Nagra analogue recorders and as such was intended for professional use in field recording and studios. The format was designed to have considerable commonality with the audio format used in D1- and D2-format digital VTRs, having rotary heads, although it used open reels for operational convenience (Kudelski et al. 1987). Allowing for 20–24 bits of audio resolution, the Nagra-D format was appropriate for use with high-resolution convertors. The error correction and recording density used in this format were designed to make recordings exceptionally robust, and recording time could be up to 6 h on a 7-in. (18 cm) reel, in two-track mode. The format is also designed for operation in a four-track mode at twice the stereo tape speed, such that in stereo the tape travells at $4.75\,\mathrm{cm\,s^{-1}}$ and in four track at $9.525\,\mathrm{cm\,s^{-1}}$.

The DASH (Digital Audio Stationary Head) format (Kogure et al. 1983) consisted of a whole family of open-reel stationary-head recording formats from two tracks up to 48 tracks. DASH-format machines operated at 44.1 or 48 kHz rates (and sometimes optionally at 44.056 kHz), and they allowed varispeed ±12.5%.

They were designed to allow gapless punch-in and punch-out, splice editing, electronic editing and easy synchronisation. Multitrack DASH machines gained wide acceptance in studios, but the stereo machines did not. Later developments resulted in DASH multitracks capable of storing 24-bit audio instead of the original 16 bits.

In more recent years, budget modular multitrack formats were introduced. Most of these were based on 8-track cassettes using rotary-head transports borrowed from consumer video technology. The most widely used were the DA-88 format (based on Hi-8 cassettes) and the ADAT format (based on VHS cassettes). These offered most of the features of open-reel machines and a number of them could be synchronised to expand the channel capacity.

1.3 Editing Digital Tape Recordings

Razor blade cut-and-splice editing was possible but risky on open-reel digital formats, and the analogue cue tracks could be monitored during these operations. The thin tape could easily be damaged during the cut-and-splice edit procedure making electronic editing far more desirable. A butt joint was used for the editing of digital tape. The discontinuity in the data stream caused by the splice would cause complete momentary dropout of the digital signal if no further action were taken; so circuits were incorporated which sensed the splice and performed an electronic crossfade from one side of the splice to the other, with error concealment to minimise the audibility of the splice.

Electronic editing normally required the use of two machines plus a control unit. A technique was employed whereby a finished master tape could be assembled from source tapes on player machines. This was a relatively slow process, as it involved real-time copying of audio from one machine to another, making modifications to the finished master difficult. The digital editor often stored several seconds of programme in its memory and this could be replayed at normal speeds or under the control of a search knob enabling very slow to-and-fro searches to be performed in the manner of rock-and-roll editing on an analogue machine. Edits could be rehearsed prior to execution. When satisfactory edit points had been determined, the two machines were synchronised using timecode, and the record machine switched to drop in the new section of the recording from the replay machine at the chosen moment. A crossfade was introduced between old and new material to smooth the join. The original source tape was left unaltered.

2 • Digital Recording on Computer Mass Storage Media

The increasing availability, performance and cheapness of computer mass storage media has led to their increasing use for digital audio storage. The majority of digital audio recording and editing is now carried out on computer-based workstations, and the days of dedicated audio formats are waning. It was most common in the early days of hard disk audio systems for manufacturers to develop dedicated systems with fairly high prices. This was mainly because mass-produced desktop computers were insufficiently equipped for the purpose, and because large-capacity mass storage media were less widely available than they are now, having a variety of different interfaces and requiring proprietary file storage strategies. Furthermore, the size of the market was relatively small to begin with, and considerable R&D investment had to be recouped. Nowadays, the average desktop computer is more than capable of handling multiple channels of digital audio, with storage capacity in abundance, even carrying out much of the audio signal processing on the host CPU if required. Alternatively, third-party audio DSP hardware can be employed for greater independence of desktop CPU power.

2.1 Storage Requirements of Digital Audio

Table 1 shows the data rates required to support a single channel of digital audio at various resolutions. Media to be used as primary storage would need to be able to sustain data transfer at a number of times these rates to be useful for audio purposes. The table also shows the number of megabytes of storage required per minute of audio. Storage requirements increase pro rata with the number of audio channels to be handled. Storage systems may use removable media but many have fixed media. It is advantageous to have removable media for audio purposes because it allows different jobs to be kept on different media and exchanged

TABLE 1 Data rates and capacities required for linear PCM

Sampling rate (kHz)	Resolution (bits)	Bit rate (kbit/s)	Capacity/min (Mbytes/min)	Capacity/h (Mbytes/h)
96	16	1,536	11.0	659
48	20	960	6.9	412
48	16	768	5.5	330
44.1	16	706	5.0	303
44.1	8	353	2.5	151

at will, but the use of removable media drives can compromise performance. An alternative is the use of removable drives where the whole mechanism including the storage surfaces can be unplugged and exchanged.

2.2 Disk Drives

Disk drives are probably the most common form of mass storage. They have the advantage of being random access systems – in other words, any data can be accessed at random and with only a short delay. This may be contrasted with tape drives, which only allow linear access, resulting in a considerable delay. Disk drives come in all shapes and sizes from the floppy disk at the bottom end to high-performance hard drives at the top end. The means by which data are stored is usually either magnetic or optical, but some use a combination of the two, as described below.

The general structure of a disk drive is shown in Figure 1. It consists of a motor connected to a drive mechanism, which causes one or more disk surfaces to rotate at anything from a few hundred to many thousands of revolutions per minute. This rotation may either remain constant or may stop and start, and it may either be at a constant rate or a variable rate, depending on the drive. One or more heads are mounted on a positioning mechanism which can move the head across the surface of the disk to access particular points under the control of hardware and software called a disk controller. The heads read data from and write data to the disk surface by whatever means the drive employs. It should be noted that certain disk types are read-only,

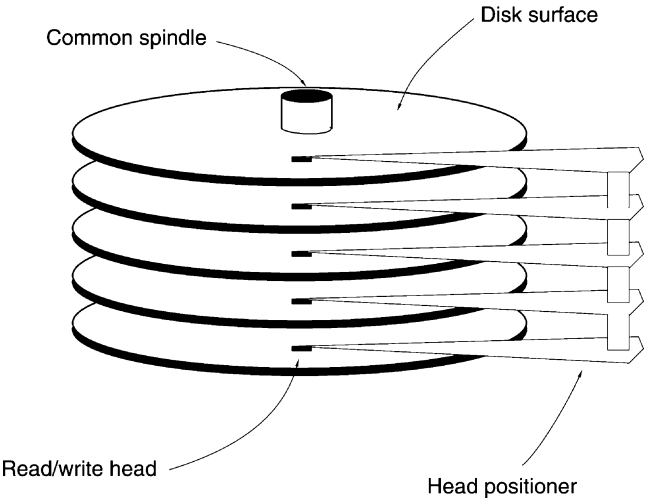


FIGURE 1

some are write-once–read-many (WORM), and some are fully erasable and rewritable.

The disk surface is normally divided up into tracks and sectors, not physically but by means of "soft" formatting (see Figure 2). Formatting places logical markers which indicate block boundaries, among other processes. On most hard disks, the tracks are arranged as concentric rings, but with some optical disks there is a continuous spiral track.

Disk drives look after their own channel coding, error detection and correction so there is no need for system designers to devise dedicated audio processes for disk-based recording systems. The formatted capacity of a disk drive is all available for the storage of "raw" audio data, with no additional overhead required for redundancy and error checking codes. "Bad blocks" are mapped out during the formatting of a disk, and not used for data storage. If a disk drive detects an error when reading a block of data, it will attempt to read it again. If this fails, then an error is normally generated and the file cannot be accessed, requiring the user to resort to one of the many file recovery packages on the market. Disk-based audio systems do not resort to error interpolation or sample hold operations, unlike tape recorders. Replay is normally either correct or not possible.

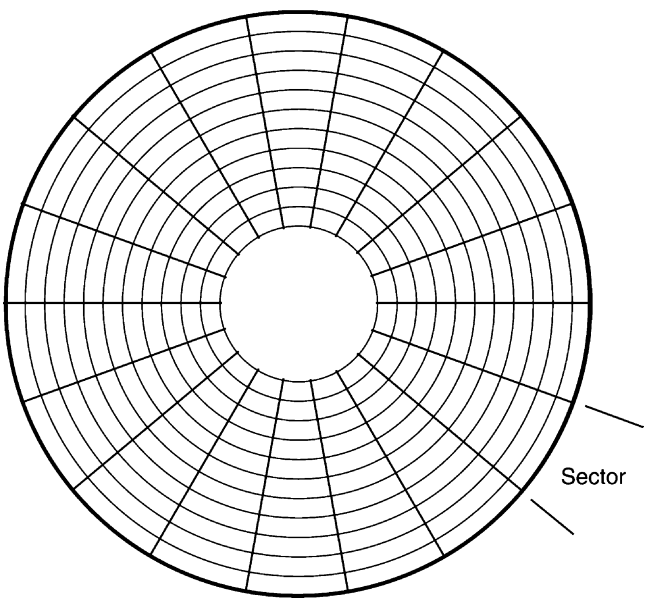


FIGURE 2

2.3 Recording Audio on to Disks

The discontinuous "bursty" nature of recording on to disk drives requires the use of a buffer RAM (random access memory) during replay, which accepts this interrupted data stream and stores it for a short time before releasing it as a continuous stream. It performs the opposite function during recording, as shown in Figure 3. Several things cause a delay in the retrieval of information: the time it takes for the head positioner to move across a disk, the time it takes for the required data in a particular track to come around to the pickup head and the transfer of the data from the disk via the buffer RAM to the outside world, as shown in Figure 4. Total delay, or data access time, is in practice several milliseconds. The instantaneous rate at which the system can accept or give out data is called the transfer rate and it varies with the storage device.

Sound is stored in named data files on the disk, the files consisting of a number of blocks of data stored either separately or together. A directory stored on the disk keeps track of where the blocks of each file are stored so that they can be retrieved in correct sequence. Each file normally corresponds to a single recording of a single channel of audio, although some stereo file formats exist (see below). Multiple channels are handled by accessing multiple files from the disk in a time-shared manner, with synchronisation between the tracks being performed subsequently in RAM. The storage capacity of a disk can be divided between channels in whatever proportion is appropriate, and it is not necessary to pre-allocate storage space to particular audio channels.

The number of audio channels that can be recorded or replayed simultaneously depends on the performance of the storage device and the host computer. Slow systems may only be capable of handling a few channels, whereas faster systems with multiple disk drives may be capable of expansion up to a virtually unlimited number of channels.

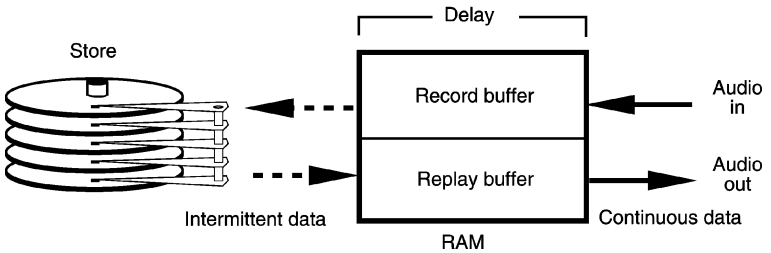


FIGURE 3

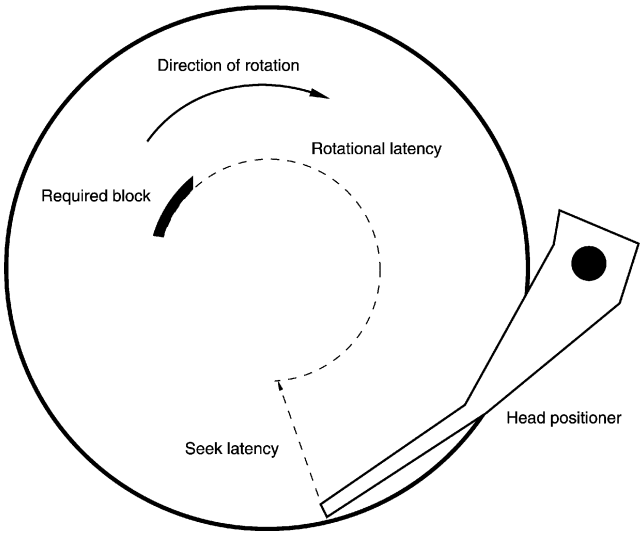


FIGURE 4

2.4 Sound File Formats

As the use of networked workstations grows, the need for files to be transferred between systems also grows, and either by international standardisation or by sheer force of market dominance, certain file formats are becoming the accepted means by which data are exchanged. The recent growth in the importance of metadata (data about data), and the representation of audio, video and metadata as "objects", has led to the development of interchange methods that are based on object-oriented concepts and project "packages" as opposed to using simple text files and separate media files. There is increasing integration between audio and other media in multimedia authoring, and some of the file formats mentioned below are closely related to international efforts in multimedia file exchange. The following is a relatively brief summary of the most commonly encountered file formats. Further details may be found in Rumsey (2004).

2.4.1 Sound Designer Formats

Sound Designer files originate from the Californian company Digidesign. Many systems handle Sound Designer files because they were used widely for such purposes as the distribution of sound effects on CD-ROM and for other short music sample files.

The Sound Designer I format (SD I, file extension ".sd") is for mono sounds and it is recommended principally for use in storing short sounds. It originated from the Apple Macintosh (Mac) computer, so numerical data are stored in big-endian byte order but it has no resource fork. The data fork contains a header of 1,336 bytes, followed by the audio data bytes. The header

contains information about how the sample should be displayed in Sound Designer editing software, including data describing vertical and horizontal scaling. It also contains details of "loop points" for the file (these are principally for use with audio/MIDI sampling packages where portions of the sound are repeatedly cycled through while a key is held down, in order to sustain a note). The header contains information on the sample rate, sample period, number of bits per sample, quantisation method (e.g. "linear", expressed as an ASCII string describing the method) and size of RAM buffer to be used. The audio data are normally either 8 or 16 bits, and always MSbyte (most significant byte) followed by LSbyte (least significant byte) of each sample.

Sound Designer II (file extension ".sd2") has been one of the most commonly used formats for audio workstations and has greater flexibility than SD I. Again it originated from a Mac file and unlike SD I it has a separate resource fork which contains the file's "vital statistics". The data fork contains only the audio data bytes in twos complement form, either 8 or 16 bits per sample. SD II files can contain audio samples for more than one channel, in which case the samples are interleaved, as shown in Figure 5, on a sample-by-sample basis (i.e. all the bytes for one channel sample followed by all the bytes for the next, etc.). It is unusual to find more than stereo data contained in SD II files, and it is recommended that multichannel recordings are made using separate files for each channel.

2.4.2 AIFF and AIFF-C Formats

The AIFF (Audio Interchange File Format) format is widely used as an audio interchange standard because it conforms to the EA IFF 85 (Electronic Arts Interchange File Format) standard for interchange format files used for various other types of information such as graphical images. AIFF (file extension ".aiff" is an Apple standard format for audio data and is encountered widely on Mac-based audio workstations and some Silicon Graphics

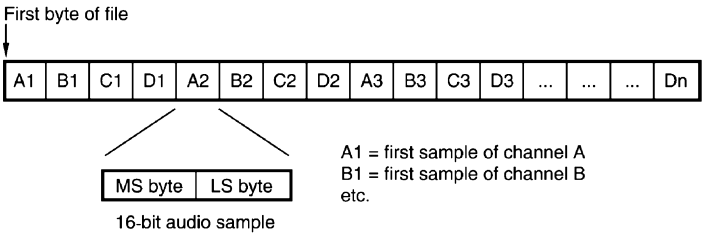


FIGURE 5

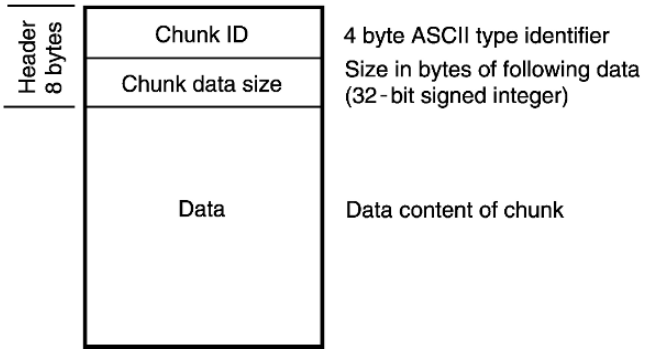


FIGURE 6

systems. Audio information can be stored at a number of resolutions and for any number of channels if required, and the related AIFF-C (file extension ".aifc") format allows also for compressed audio data. A file consists only of a data fork, with no resource fork, making it easier to transport to other platforms.

All IFF-type files are made up of "chunks" of data which are typically made up as shown in Figure 6. A chunk consists of a header and a number of data bytes to follow. The simplest AIFF files contain a "common chunk", which is equivalent to the header data in other audio files, and a "sound data" chunk containing the audio sample data. These are contained overall by a "form" chunk as shown in Figure 7. AIFC files must also contain a "Version Chunk" before the common chunk to allow for future changes to AIFC.

2.4.3 RIFF WAVE Format

The RIFF (Resource Interchange File Format) WAVE (often called WAV) format is the Microsoft equivalent of Apple's AIFF. It has a similar structure, again conforming to the IFF pattern, but with numbers stored in little-endian rather than big-endian form. It is used widely for sound file storage and interchange on PC workstations and for multimedia applications involving sound. Within WAVE files, it is possible to include information about a number of cue points and a playlist to indicate the order in which the cues are to be replayed. WAVE files use the file extension ".wav".

A basic WAV file consists of three principal chunks, as shown in Figure 8, the RIFF chunk, the FORMAT chunk and the DATA chunk. The RIFF chunk contains 12 bytes, the first four of which

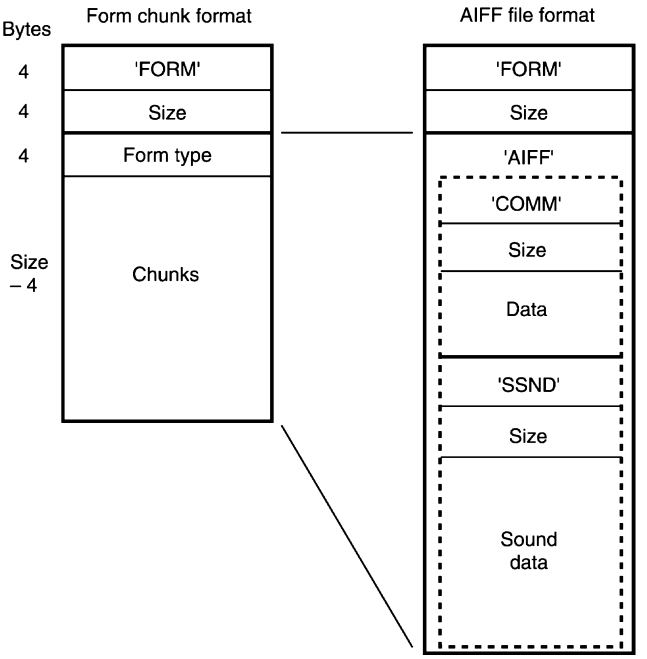


FIGURE 7

are the ASCII characters "RIFF", the next four indicating the number of bytes in the remainder of the file (after the first eight) and the last four of which are the ASCII characters "WAVE". The format chunk contains information about the format of the sound file, including the number of audio channels, sampling rate and bits per sample, as shown in Table 2.

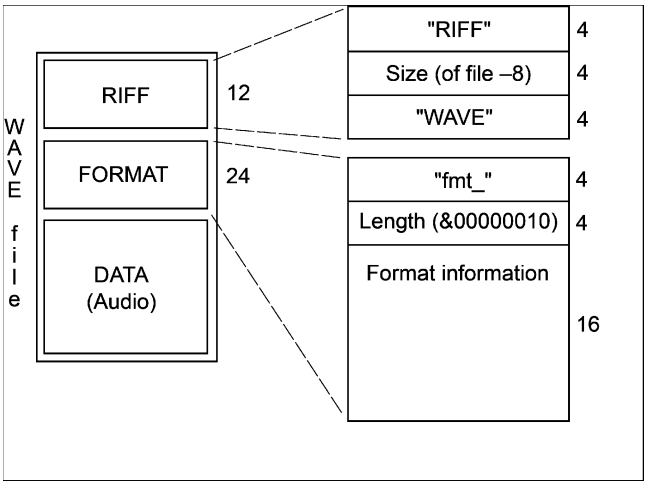


FIGURE 8

TABLE 2 Contents of FORMAT chunk in a basic WAVE PCM file

Byte	ID	Contents
0–3	ckID	"fmt_" (ASCII characters)
4–7	nChunkSize	Length of FORMAT chunk (binary, hex value: &00000010)
8–9	wFormatTag	Audio data format (e.g. &0001 = WAVE format PCM) Other formats are allowed, for example, IEEE floating point and MPEG format (&0050 = MPEG 1)
10–11	nChannels	Number of channels (e.g. &0001=mono, &0002=stereo)
12–15	nSamplesPerSec	Sample rate (binary, in Hz)
16–19	nAvgBytesPerSec	Bytes per second
20–21	nBlockAlign	Bytes per sample: e.g. &0001=8 bit mono; &0002=8 bit stereo or 16 bit mono; &0004=16 bit stereo
22–23	nBitsPerSample	Bits per sample

The audio data chunk contains a sequence of bytes of audio sample data, divided as shown in the FORMAT chunk. Unusually, if there are only 8 bits per sample or fewer, each value is unsigned and ranges between 0 and 255 (decimal), whereas if the resolution is higher than this, the data are signed and range both positively and negatively around zero. Audio samples are interleaved by channel in time order so that if the file contains two channels, a sample for the left channel is followed immediately by the associated sample for the right channel. The same is true of multiple channels (one sample for time-coincident sample periods on each channel is inserted at a time, starting with the lowest numbered channel), although basic WAV files were nearly always just mono or two-channel.

The RIFF WAVE format is extensible and can have additional chunks to define enhanced functionality such as surround sound and other forms of coding. This is known as "WAVE-format extensible" (see http://www.microsoft.com/hwdev/tech/audio/multichaud.asp). Chunks can include data relating to cue points, labels and associated data, for example.

2.4.4 Broadcast WAVE Format

The Broadcast WAVE format, described in EBU Tech. 3285 (EBU 2001), was standardised by the European Broadcasting Union (EBU) because of a need to ensure compatibility of sound files and accompanying information when transferred between workstations. It is based on the RIFF WAVE format described above, but contains an additional chunk that is specific to the format (the "broadcast_audio_extension" chunk, ID = "bext") and

also limits some aspects of the WAVE format. version 0 was published in 1997 and version 1 in 2001, the only difference being the addition of a SMPTE (Society of Motion Picture and Television Engineers) UMID (Unique Material Identifier) in version 1 (this is a form of metadata). Such files currently only contain either PCM or MPEG-format audio data.

A multichannel extension chunk has recently been proposed for Broadcast WAVE files that define the channel ordering, surround format, downmix coefficients for creating a two-channel mix and some descriptive information. There are also chunks defined for metadata describing the audio contained within the file, such as the "quality chunk" (ckID = "qlty"), which together with the coding history contained in the "bext" chunk make up the so-called capturing report. These are described in Supplement 2 to EBU Tech. 3285. Finally, there is a chunk describing the peak audio level within a file, which can aid automatic programme level setting and programme interchange.

2.4.5 MPEG Audio File Formats

It is possible to store MPEG-compressed audio in AIFF-C or WAVE files, with the compression type noted in the appropriate header field. (MPEG is the Motion Pictures Expert Group that defines standards for data-reduced audio.) There are also older MS-DOS file extensions used to denote MPEG audio files, notably ".mpa" (MPEG Audio) or ".abs" (Audio Bit Stream). However, owing to the ubiquity of the so-called MP3 format (MPEG 1, Layer 3) for audio distribution on the Internet, MPEG audio files are increasingly denoted with the extension ".mp3". Such files are relatively simple, being really no more than MPEG audio frame data in sequence, each frame being preceded by a frame header.

2.4.6 DSD-IFF File Format

The DSD-IFF file format is based on a similar structure to other IFF-type files, described above, except that it is modified slightly to allow for the large file sizes that may be encountered with the high-resolution DSD (Direct Stream Digital) format used, for SuperAudio CD. Specifically the container FORM chunk is labelled "FRM8" and this identifies all local chunks that follow as having "length" indications that are 8 bytes long rather than the normal four. In other words, rather than a 4-byte chunk ID followed by a 4-byte length indication, these files have a 4-byte ID followed by an 8-byte length indication. This allows for the definition of chunks with a length greater than 2 Gbytes, which may be needed for mastering SuperAudio CDs. There are also various optional chunks that can be used for exchanging more detailed information and comments such as might be used in

project interchange. Further details of this file format, and an excellent guide to the use of DSD-IFF in project applications, can be found in the DSD-IFF specification (Philips 2002a and b).

2.4.7 Edit Decision List Files

EDL (edit decision list) formats have usually been unique to the workstation on which they are used, but the need for open interchange is increasing the pressure to make EDLs transportable between packages. There is an old and widely used format for EDLs in the video world that is known as the CMX-compatible form. CMX is a well-known manufacturer of video editing equipment, and most editing systems will read CMX EDLs for the sake of compatibility. These can be used for basic audio purposes, and indeed a number of workstations can read CMX EDL files for the purpose of auto-conforming audio edits to video edits performed on a separate system. The CMX list defines the cut points between source material and the various transition effects at joins, and it can be translated reasonably well for the purpose of defining audio cut points and their timecode locations, using SMPTE/EBU form, provided video frame accuracy is adequate.

Software can be obtained for audio and video workstations that translate EDLs between a number of different standards to make interchange easier, although it is clear that this process is not always problem-free and good planning of in-house processes is vital. The Open Media Framework Interchange (OMFI) structure also contains a format for interchanging edit list data. The AES 31 standard (AES 1999, 2001) is now gaining considerable popularity among workstation software manufacturers as a simple means of exchanging audio editing projects between systems. The Advanced Authoring Format (AAF) is becoming increasingly relevant to the exchange of media project data between systems and is likely to take over from OMFI as time progresses.

2.4.8 MXF

MXF (Media Exchange Format) was developed by the Pro-MPEG forum as a means of exchanging audio, video and metadata between devices, primarily in television operations. It is based on the modern concept of media objects that are split into "essence" and "metadata". Essence files are the raw material (i.e. audio and video) and the metadata describes things about the essence (such as where to put it, where it came from and how to process it).

MXF files attempt to present the material in a "streaming" format, that is one that can be played out in real time, but they can also be exchanged in conventional file transfer operations. As such they are normally considered to be finished programme

material, rather than material that is to be processed somewhere downstream, designed for playout in broadcasting environments. The bit stream is also designed to be compatible with recording on digital videotape devices.

2.4.9 AAF

AAF (Advanced Authoring Format) is an authoring format for multimedia data that is supported by numerous vendors, including Avid which has adopted it as a migration path from OMFI (Gilmer 2002). Parts of OMFI 2.0 form the basis for parts of AAF and there are also close similarities between AAF and MXF (described in the previous section). Like the formats to which it has similarities, AAF is an object-oriented format that combines essence and metadata within a container structure. Unlike MXF it is designed for project interchange such that elements within the project can be modified, post-processed and resynchronised. It is not, therefore, directly suitable as a streaming format but can easily be converted to MXF for streaming if necessary.

Rather like OMFI, AAF is designed to enable complex relationships to be described between content elements, to map these elements onto a timeline, to describe the processing of effects, synchronise streams of essence, retain historical metadata and refer to external essence (essence not contained within the AAF package itself). It has three essential parts: the AAF Object Specification (which defines a container for essence and metadata, the logical contents of objects and rules for relationships between them); the AAF Low-Level Container Specification (which defines a disk filing structure for the data, based on Microsoft's Structured Storage); and the AAF SDK Reference Implementation (which is a software development kit that enables applications to deal with AAF files). The Object Specification is extensible in that it allows new object classes to be defined for future development purposes.

3 • Consumer Digital Formats

3.1 Compact Discs and Drives

Compact discs (CDs) are not immediately suitable for real-time audio editing and production, partly because of their relatively slow access time compared with hard disks, but can be seen to have considerable value for the storage and transfer of sound material that does not require real-time editing. Broadcasters use them for sound effects libraries and studios and mastering facilities use them for providing customers and record companies with "acetates" or test pressings of a new recording. They have also become quite popular as a means of transferring finished

masters to a CD pressing plant in the form of the PMCD (pre-master CD). They are ideal as a means of "proofing" CD-ROMs and other CD formats and can be used as low-cost backup storage for computer data. Detailed coverage of CD formats and "books" may be found in Pohlmann (1992).

CDs are familiar to most people as a consumer read-only optical disk for audio (CD-DA) or data (CD-ROM) storage. Standard audio CDs (CD-DA – DA for digital audio) conform to the Red Book standard published by Philips. The CD-ROM standard (Yellow Book) divides the CD into a structure with 2,048 byte sectors, adds an extra layer of error protection and makes it useful for general-purpose data storage including the distribution of sound and video in the form of computer data files. It is possible to find disks with mixed modes, containing sections in CD-ROM format and sections in CD-Audio format. The CD Plus is one such example.

CD-R is the recordable CD and may be used for recording CD-Audio format or other CD formats using a suitable drive and software. The Orange Book, Part 2, contains information on the additional features of CD-R, such as the area in the centre of the disk where data specific to CD-R recordings is stored. Audio CDs recorded according to the Orange Book standard can be "fixed" to give them a standard Red Book table of contents (TOC), allowing them to be replayed on any conventional CD player. Once fixed into this form, the CD-R may not subsequently be added to or changed, but prior to this there is a certain amount of flexibility, as discussed below. CD-RW (RW for read–write) disks are erasable and work on phase-change principles, requiring a drive compatible with this technology, being described in the Orange Book, Part 3.

The degree of reflectivity of CD-RW disks is much lower than that of typical CD-R and CD-ROM. This means that some early drives and players may have difficulties reading them. However, the "multi-read" specification developed by the OSTA (Optical Storage Technology Association) describes a drive that should read all types of CD, so recent drives should have no difficulties here.

3.2 Minidisc

CD has been available for some years now as a 16-bit, 44.1-kHz digital playback medium; it was joined by CD-ROM, CD-R (recordable) and CD-RW (recordable and rewritable). The minidisc (MD) is now an established consumer recording and playback format, and it is of the M-O (magneto-optical) type. Sampling frequency is fixed at 44.1 kHz, and resolution is nominally 16 bits. A coding system similar to those originally developed for digital audio broadcasting (DAB) known as

Adaptive Transform Acoustic Coding (ATRAC) is used (Tsutsui et al. 1992). A data rate of about one-fifth of that required for CD is adequate to encode the resulting signal (CD's data stream is 1.4 Mb/s, MD's is 292 kb/s), and this allows an equivalent playing time to be obtained from a disc which is somewhat smaller than a CD at 64 mm in diameter. Since the format involves considerable data compression (a slight misnomer for data reduction), it is not suitable for professional master recording or archiving, but is used quite widely in applications where the highest sound quality is not required such as broadcast journalism.

3.3 DVD

DVD (Digital Versatile Disk) is the natural successor to CD, being a higher-density optical disk format aimed at the consumer market, having the same diameter as CD and many similar physical features (De Bont et al. 1996). It uses a different laser wavelength to CD (635–650 nm as opposed to 780 nm) so multistandard drives need to be able to accommodate both. Data storage capacity depends on the number of sides and layers on the disk but ranges from 4.7 Gbytes (single-layer, single-sided) up to about 18 Gbytes (double-layer, double-sided). The data transfer rate at "one times" speed is just over 11 Mbit/s.

DVD can be used as a general-purpose data storage medium. Like CD, there are numerous different variants on the recordable DVDs, partly owing to competition between the numerous different "factions" in the DVD consortium. These include DVD-R, DVD-RAM, DVD-RW and DVD+RW, all of which are based on similar principles but have slightly different features, leading to a compatibility minefield that is only gradually being addressed (see Table 3). The "DVD Multi" guidelines produced by the DVD Forum are an attempt to foster greater compatibility between DVD drives and disks, but this does not really solve the problem of the formats that are currently outside the DVD Forum.

Writeable DVDs are a useful option for backup of large projects, particularly DVD-RAM because of its many-times overwriting capacity and its hard disk-like behaviour. It is possible that a format like DVD-RAM could be used as primary storage in a multitrack recording/editing system, as it has sufficient performance for a limited number of channels and it has the great advantage of being removable. However, it is likely that hard disks will retain the performance edge for the foreseeable future.

DVD-Video is the format originally defined for consumer distribution of movies with surround sound, typically incorporating MPEG-2 video encoding and Dolby Digital surround sound encoding. It also allows for up to eight channels of 48 or 96 kHz linear PCM audio, at up to 24-bit resolution. DVD-

TABLE 3 Recordable DVD formats

Recordable DVD type	Description
DVD-R (A and G)	DVD equivalent of CD-R. One-time recordable in sequential manner, replayable on virtually any DVD-ROM drive Supports "incremental writing" or "disk at once" recording. Capacity either 3.95 (early disks) or 4.7 Gbytes per side. "Authoring" (A) version (recording laser wavelength = 635 nm) can be used for pre-mastering DVDs for pressing, including DDP data for disk mastering "General" (G) version (recording laser wavelength = 650 nm) intended for consumer use, having various "content protection" features that prevent encrypted commercial releases from being cloned
DVD-RAM	Sectored format, rather more like a hard disk in data structure when compared with DVD-R. Uses phase-change (PD-type) principles allowing direct over-write. Version 2 disks allow 4.7 Gbyte per side (reduced to about 4.2 Gbytes after formatting). Type 1 cartridges are sealed and Type 2 allow the disk to be removed. Double-sided disks only come in sealed cartridges. Can be rewritten about 100,000 times. The recent Type 3 is a bare disk that can be placed in an open cartridge for recording
DVD-RW	Pioneer development, similar to CD-RW in structure, involving sequential writing. Does not involve a cartridge Can be re-written about 1,000 times. 4.7 Gbytes per side
DVD+RW	Non-DVD-Forum alternative to DVD-RAM (and not compatible), allowing direct over-write. No cartridge. Data can be written in either CLV (for video recording) or CAV (for random access storage) modes. There is also a write-once version known as DVD+R

Audio is intended for very high quality multichannel audio reproduction and allows for linear PCM sampling rates up to 192 kHz, with numerous configurations of audio channels for different surround modes, and optional lossless data reduction (MLP – see below).

DVD-Audio (DVD-A) has a number of options for choosing the sampling frequencies and resolutions of different channel groups, it being possible to use a different resolution on the front channels from that used on the rear, for example. The format is more versatile in respect of sampling frequency than DVD-Video, having also accommodated multiples of the CD sample frequency of 44.1 kHz as options (the DVD-Video format allows only for multiples of 48 kHz). Consequently, the allowed sample frequencies for DVD-Audio are 44.1, 48, 88.2, 96, 176.4 and 192 kHz. The sample frequencies are split into two groups – multiples of 44.1 and multiples of 48 kHz. While it is possible

to split frequencies from one group among the audio channels on a DVD-A, one cannot combine frequencies across the groups for reasons of simple clock rate division. Bit resolution can be 16, 20 or 24 bits per channel, and again this can be divided unequally between the channels, according to the channel group split described below.

Meridian Lossless Packing (MLP) is licensed through Dolby Laboratories and is a lossless coding technique designed to reduce the data rate of audio signals without compromising sound quality (Gerzon et al. 2004). It has both a variable bit rate mode and a fixed bit rate mode. The variable mode delivers the optimum compression for storing audio in computer data files, but the fixed mode is important for DVD applications where one must be able to guarantee a certain reduction in peak bit rate.

3.4 Super Audio CD

Version 1.0 of the Super Audio CD (SACD) specification (Verbakel et al. 1998) is described in the "Scarlet Book", available from Philips licensing department. SACD uses DSD (Direct Stream Digital) as a means of representing audio signals, so it requires audio to be sourced in or converted to this form. SACD aims to provide a playing time of at least 74 min for both two-channel and six-channel balances. The disc is divided into two regions, one for two-channel audio, the other for multichannel, as shown in Figure 9. A lossless data packing method known as Direct Stream Transfer (DST) can be used to achieve roughly 2:1 data reduction of the signal stored on disc so as to enable high-quality multichannel audio on the same disc as the two-channel mix.

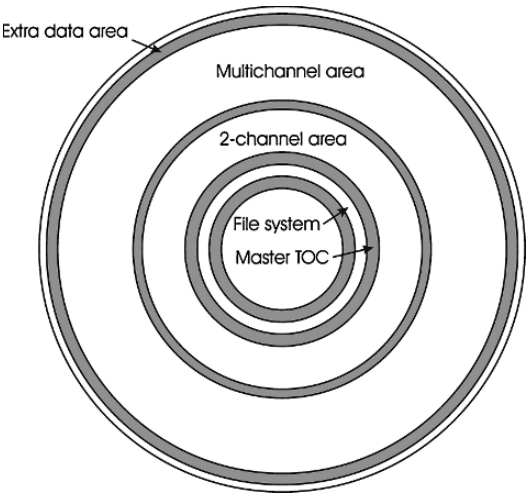


FIGURE 9

SACDs can be manufactured as single- or dual-layer discs, with the option of the second layer being a Red Book CD layer (the so-called hybrid disc). SACDs, not being a formal part of the DVD hierarchy of standards (although using some of the optical disc technology), do not have the same options for DVD-Video objects as DVD-Audio. The disc is designed first and foremost as a super-high-quality audio medium. Nonetheless, there is provision for additional data in a separate area of the disc. The content and capacity of this is not specified but could be video clips, text or graphics, for example.

4 • Digital Audio Editing Using Mass Storage Media

The random access nature of disk-based mass storage media led to the coining of the term non-linear editing for the process of audio editing. With non-linear editing, the editor may preview a number of possible masters in their entirety before deciding which should be the final one. Even after this, it is a simple matter to modify the edit list to update the master. Edits may also be previewed and experimented with in order to determine the most appropriate location and processing. Crossfades may be modified and adjustments made to equalisation and levels, all in the digital domain. Non-linear editing has also come to feature very widely in post-production for video and film.

Non-linear editing is truly non-destructive in that the edited master only exists as a series of instructions to replay certain parts of sound files at certain times, with specified signal processing overlaid, as shown in Figure 10. The original sound files remain intact at all times, and a single sound file can be used as many times as desired in different locations and on different tracks without the need for copying the audio data. Editing may involve the simple joining of sections, or it may involve more complex operations such as long crossfades between one album track and the next, or gain offsets between one section and another. All these things are possible without affecting the original source material.

4.1 Sound Files and Sound Segments

In the case of music editing sound files might be session takes, anything from a few bars to a whole movement, while in picture dubbing they might contain a phrase of dialogue or a sound effect. Specific segments of these sound files can be defined while editing, in order to get rid of unwanted material or to select useful extracts. The terminology varies but such identified parts of sound files are usually termed either “clips” or “segments”. Rather than creating a copy of the segment or clip and storing

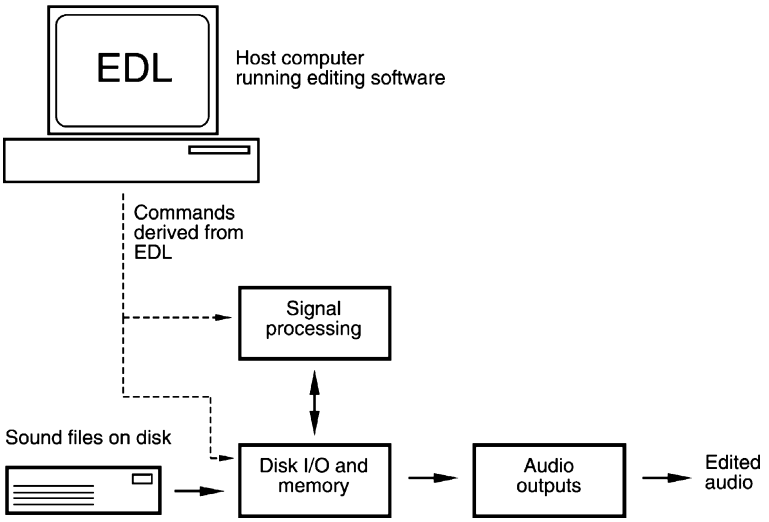


FIGURE 10

it as a separate sound file, it is normal simply to store it as a "soft" entity – in other words as simply commands in an edit list or project file that identify the start and end addresses of the segment concerned and the sound file to which it relates. It may be given a name by the operator and subsequently used as if it were a sound file in its own right. An almost unlimited number of these segments can be created from original sound files, without the need for any additional audio storage space.

4.2 Edit Point Handling

Edit points can be simple butt joins or crossfades. A butt join is very simple because it involves straightforward switching from the replay of one sound segment to another. Since replay involves temporary storage of the sound file blocks in RAM (see above), it is a relatively simple matter to ensure that both outgoing and incoming files in the region of the edit are available in RAM simultaneously (in different address areas). Up until the edit, blocks of the outgoing file are read from the disk into RAM and thence to the audio outputs. As the edit point is reached, a switch occurs between outgoing and incoming material by instituting a jump in the memory read address corresponding to the start of the incoming material. Replay then continues by reading subsequent blocks from the incoming sound file. It is normally possible to position edits to single sample accuracy, making the timing resolution as fine as a number of tens of microseconds if required.

The problem with butt joins is that they are quite unsubtle. Audible clicks and bumps may result because of a discontinuity

in the waveform. It is normal, therefore, to use at least a short crossfade at edit points to hide the effect of the join. This is what happens when analogue tape is spliced because the traditional angled cut has the same effect as a short crossfade (of between 5 and 20 ms depending on the tape speed and angle of cut). Most workstations have considerable flexibility with crossfades and they are not limited to short durations. It is now common to use crossfades of many shapes and durations (e.g. linear, root cosine, equal power) for different creative purposes. This, coupled with the ability to preview edits and fine-tune their locations, has made it possible to put edits in places previously considered impossible.

The locations of edit points are kept in an edit decision list (EDL) which contains information about the segments and files to be replayed at each time, the in and the out points of each section and details of the crossfade time and shape at each edit point. It may also contain additional information such as signal processing operations to be performed (gain changes, EQ, etc.)

4.3 Crossfading

Crossfading is similar to butt joining, except that it requires access to data from both incoming and outgoing files for the duration of the crossfade. The crossfade calculation involves simple signal processing, during which the values of outgoing samples are multiplied by gradually decreasing coefficients while the values of incoming samples are multiplied by gradually increasing coefficients. Time-coincident samples of the two files are then added together to produce output samples, as described in Chapter 35. The duration and shape of the crossfade can be adjusted by altering the coefficients involved and the rate at which the process is executed.

Crossfades are either performed in real time, as the edit point passes, or pre-calculated and written to disk as a file. Real-time crossfades can be varied at any time and are simply stored as commands in the EDL, indicating the nature of the fade to be executed. The process is similar to that for the butt edit, except that as the edit point approaches samples from both incoming and outgoing segments are loaded into RAM in order that there is an overlap in time. During the crossfade, it is necessary to continue to load samples from both incoming and outgoing segments into their respective areas of RAM, and for these to be routed to a crossfade processor. The resulting samples are then available for routing to the output. Alternatively, the crossfade can be calculated in non-real time. This incurs a short delay while the system works out the sums, after which a new sound file is stored which contains only the crossfade. Replay of the edit then involves playing the outgoing segment up to the beginning of the

crossfade, then the crossfade file and then the incoming segment from after the crossfade. Load on the disk drive is no higher than normal in this case.

The shape of the crossfade can usually be changed to suit different operational purposes. Standard linear fades (those where the gain changes uniformly with time) are not always the most suitable for music editing, especially when the crossfade is longer than about 10 ms. The result may be a momentary drop in the resulting level in the centre of the crossfade that is due to the way in which the sound levels from the two files add together. If there is a random phase difference between the signals, as there will often be in music, the rise in level resulting from adding the two signals will normally be around 3 dB, but the linear crossfade is 6 dB down in its centre resulting in an overall level drop of around 3 dB (see Figure 11). Exponential crossfades and other such shapes may be more suitable for these purposes because they have a smaller level drop in the centre. It may even be possible to design customised crossfade laws. It is often possible to alter the offset of the start and end of the fade from the actual edit point and to have a faster fade up than fade down. Many systems also allow automated gain changes to be introduced as well as fades, so that level differences across edit points may be corrected. A lot of the difficulties that editors encounter in making edits work can be solved using a combination of these facilities.

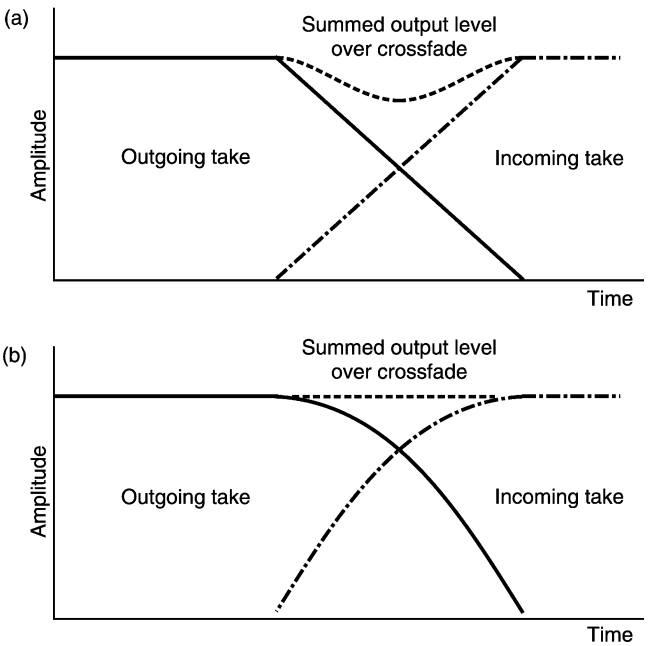


FIGURE 11

4.4 Editing Modes

During the editing process, the operator will load appropriate sound files and audition them, both on their own and in a sequence with other files. The exact method of assembling the edited sequence depends very much on the user interface, but it is common to present the user with a visual analogy of moving tape, allowing files to be "cut-and-spliced" or "copied and pasted" into appropriate locations along the virtual tape. These files, or edited clips of them, are then played out at the timecode locations corresponding to their positions on this "virtual tape". It is also quite common to display a representation of the audio waveform that allows the editor to see as well as hear the signal around the edit point (see Figure 12).

In non-linear systems the tape-based approach is often simulated, allowing the user to roughly locate an edit point while playing the virtual tape followed by a fine trim using simulated reel-rocking or a detailed view of the waveform. Some software presents source and destination streams as well, in further simulation of the tape approach. It is also possible to insert or change sections in the middle of a finished master, provided the EDL and source files are still available. To take an example, assume that an edited opera has been completed and that the producer now wishes to change a take somewhere in the middle. The replacement take is unlikely to be exactly the same length, but it is possible simply to shuffle all of the following material along or back slightly to accommodate it, this being only a matter of changing the EDL rather than modifying the stored music in any way. The files are then simply played out at slightly different times than in the first version of the edit.

It is also normal to allow edited segments to be fixed in time if desired, so that they are not shuffled forwards or backwards

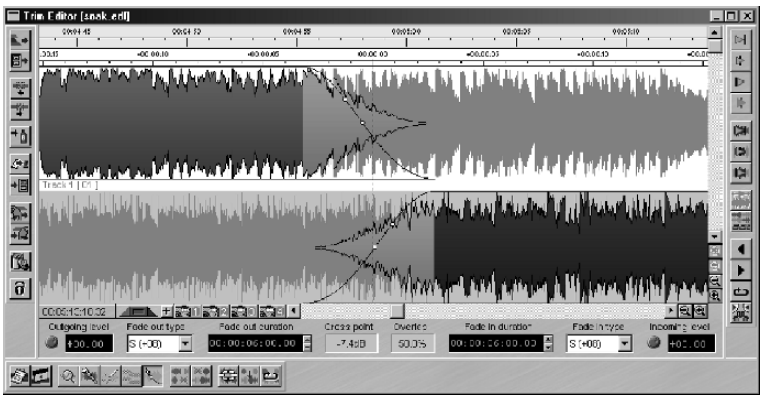


FIGURE 12

when other segments are inserted. This "anchoring" of segments is often used in picture dubbing when certain sound effects and dialogue have to remain locked to the picture.

4.5 Simulation of "Reel-Rocking"

It is common to simulate the effect of analogue tape "reel-rocking" in non-linear editors, providing the user with the sonic impression that reels of analogue tape are being "rocked" back and forth as they are in analogue tape editing when fine-searching edit points. Editors are used to the sound of tape moving in this way and are skilled at locating edit points when listening to such a sound.

The simulation of variable speed replay in both directions (forwards and backwards) is usually controlled by a wheel or sideways movement of a mouse which moves the "tape" in either direction around the current play location. This magnitude and direction of this movement is used to control the rate at which samples are read from the disk file, via the buffer, and this replaces the fixed sampling rate clock as the controller of the replay rate. Systems differ very greatly as to the sound quality achieved in this mode because it is in fact quite a difficult task to provide convincing simulation. So poor have been many attempts that many editors do not use the feature, preferring to judge edit points accurately "on the fly", followed by trimming or nudging them either way if they are not successful the first time. Good simulation requires very fast, responsive action and an ergonomically suitable control. A mouse is very unsuitable for the purpose. It also requires a certain amount of DSP to filter the signal correctly in order to avoid the aliasing that can be caused by varying the sampling rate.

References

AES (1999) *AES 31-3-1999: AES 31-3-1999: AES standard for network and file transfer of audio. Audio file transfer and exchange: Part 3 – Simple project interchange.*

AES (2001) *AES 31-1-2001: AES standard for network and file transfer of audio. Audio file transfer and exchange: Part 1 – Disk format.*

De Bont, F. et al. (1996) The high density DVD. *Proceedings of the AES UK Conference on Audio for New Media*. Audio Engineering Society.

Doi, T., Tsuchiya, Y. and Iga, A. (1978) On several standards for converting PCM signals into video signals. *J. Audio Eng. Soc.*26, 9, pp. 641–649.

EBU (2001) *Tech. 3285. BWF – a format for audio data files in broadcasting. Version 1.* (See also Supplements 1, 2 and 3.). European Broadcasting Union.

Gerzon, M. et al. (2004) The MLP lossless compression system for PCM audio. *J. Audio Eng. Soc.* 52, 3, pp. 243–260.

Gilmer, B. (2002) AAF – the advanced authoring format. *EBU Tech. Rev.*, July.

Ishida, Y. et al. (1985) On the development of a rotary head digital audio tape recorder. *Presented at 79th AES Convention*, Preprint 2271. Audio Engineering Society.

Kogure, T., Doi, T. and Lagadec, R. (1983) The DASH format: an overview. *Presented at 74th AES Convention*, Preprint 2038. Audio Engineering Society.

Kudelski, S. et al. (1987) Digital audio recording format offering extensive editing capabilities. *Presented at 82nd AES Convention*, Preprint 2481. Audio Engineering Society.

Nakajima, H., Doi, T., Tsuchiya, Y. and Iga, A. (1978) A new PCM audio system as an adapter of video tape recorders. *Presented at 60th AES Convention*, Preprint 1352. Audio Engineering Society.

Philips (2002a) *Direct Stream Digital Interchange File Format: DSD-IFF version 1.4, revision 2*. Available from www.superaudiocd. philips.com.

Philips (2002b) *Recommended usage of DSD-IFF, version 1.4*. Available from www.superaudiocd.philips.com

Pohlmann, K. (1992) *The Compact Disc Handbook*. A-R Editions.

Rumsey, F. (2004) *Desktop Audio Technology*. Oxford and Boston: Focal Press.

Tsutsui, K. et al. (1992) ATRAC: Adapative transform acoustic coding for minidisc. *Presented at 93rd AES Convention*, Preprint 3456. Audio Engineering Society.

Verbakel, J. et al. (1998) Super Audio CD format. *Presented at 104th AES Convention*, Preprint 4705. Audio Engineering Society.

37

Audiovisual Interaction

Riikka Möttönen and Mikko Sams

Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, Helsinki, Finland

Information about the objects and events in the external world is received via multiple sense organs, especially via eyes and ears. For example, a singing bird can be heard and seen. Typically, audiovisual objects are detected, localized and identified more rapidly and accurately than objects which are perceived via only one sensory system (see, e.g. Welch and Warren, 1986; Stein and Meredith, 1993; de Gelder and Bertelson, 2003; Calvert et al., 2004). The ability of the central nervous system to utilize sensory inputs mediated by different sense organs is called multisensory processing.

In this chapter we will concentrate on reviewing how perception of acoustic information is affected by concurrent visual information. We first discuss briefly how visual information can affect sound localization. Then we focus on audiovisual speech perception, i.e. how seeing a talker's articulatory gestures affects perception of acoustic speech. We finish by discussing the brain mechanisms of audiovisual perception.

1 • Audiovisual Space

Typically, acoustic and visual inputs coming from the same location at the same time originate from the same audiovisual object. Therefore, spatially coherent acoustic and visual inputs are integrated and perceived as a single audiovisual object. As a result of integration, orientation to new objects in the environment becomes more accurate than orientation to

unimodal objects alone (Stein et al., 1989). In contrast, orientation to spatially disparate, but simultaneous, objects becomes worse than that to unimodal objects alone (Stein et al., 1989).

1.1 Ventriloquism Illusion

When a visual stimulus is presented simultaneously with an acoustic stimulus but in a slightly different location, both inputs are perceived to arrive from the same location. The common location is mainly determined by the location of the visual stimulus. This illusion is known as the ventriloquism illusion (Howard and Templeton, 1966), because an illusory change of sound location is experienced when ventriloquist's speech is perceived to come from the moving mouth of a puppet. In modern everyday life, a similar illusion is experienced when visual events (e.g. actor's articulatory gestures) on a TV or movie screen determine the perceptual locations of sounds (e.g. actor's voice), which actually come from loudspeakers.

The ventriloquism illusion has been extensively studied in laboratories (for a review, see Bertelson and Gelder, 2004). In psychophysical experiments, participants are instructed to point toward locations of sounds when they are presented alone and when they are presented together with simultaneous, but slightly disparate, flashes of light. Perceived sound location is typically affected by flashes of light. The ventriloquism effect is the strongest when the distance between acoustic and visual inputs is small and they are perceived as a unified object (e.g. Wallace et al., 2004). When the distance of acoustic and visual inputs increases, the inputs are more often perceived to be disparate and the ventriloquism effect becomes weaker (e.g. Wallace et al., 2004). Furthermore, temporal asynchrony between acoustic and visual inputs weakens the influence of vision on perceived sound location (e.g. Slutsky and Recanzone, 2001; Wallace et al., 2004).

2 • Audiovisual Speech

Speech perception usually refers to perception of acoustic speech. However, speech is also visual. Viewing a talker's articulatory gestures allows the observer to understand speech to some extent (for reviews, see Dodd and Campbell, 1987; Campbell et al., 1998). This ability to *speechread* varies a lot across observers. Hearing impaired and deaf people are often very skilful speechreaders, but normal-hearing people are able to speechread as well. In everyday conversations, we typically hear talkers' voice and see their articulatory gestures simultaneously, thus we perceive speech audiovisually.

2.1 Audiovisual Perception of Degraded Acoustic Speech

Sumby and Pollack (1954) published a landmark paper on integration of acoustic and visual speech. They presented bisyllabic words at different signal-to-noise ratios (SNR) to subjects who either only heard the stimuli or also saw the talker to articulate them. SNR was varied from zero to −30 dB. In auditory-only condition, the identification of the words deteriorated progressively as SNR decreased. When the same stimuli could also be seen, the overall performance improved greatly. For example, at −18 SNR in the auditory-only condition the subjects correctly identified about 28% of the words, but in the audiovisual condition 78% of them.

In another pioneering study, O'Neill (1954) studied audiovisual perception of vowels, consonants, words and sentences. Seeing the speech improved identification of each stimulus class. MacLeod and Summerfield (1987) used meaningful sentences to study whether the lipreading skill in normal subjects correlates with the ability to utilize visible speech in audiovisual speech perception. Visual-alone performance correlated significantly with the audiovisual benefit. The mean benefit of seeing speech was equivalent to 11-dB increase in SNR, the range being 6–15 dB. In addition, those sentences which were easier to lipread were also recognized at a lower SNR in the audiovisual condition.

Figure 1 shows results from a speech recognition experiment in 20 normal-hearing subjects (Möttönen, 1999). Meaningless acoustic vowel–consonant–vowel stimuli were presented both

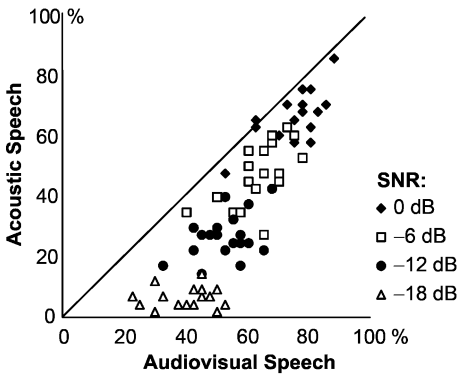


FIGURE 1 Proportions of correctly identified acoustic and audiovisual speech items at four signal-to-noise-ratio (SNR) levels (0, −6, −12, −18 dB) in 20 Finnish subjects. The stimuli were 39 vowel–consonant–vowel words produced by a native Finnish speaker. Each acoustic stimulus was presented alone and together with a matching visual articulation. Modified from Möttönen (1999).

alone and with matching visual speech. Noise was added to acoustic stimuli to acquire four SNR levels: 0, −6, −12 and −18 dB. In each subject, the proportion of correctly recognized audiovisual speech stimuli was greater than that of acoustic speech stimuli at all SNR levels. The lower the SNR of acoustic speech the greater the benefit of visual information.

Voicing and nasality are the articulatory features which are the least affected by noise and the place of articulation is the most difficult one to hear (Binnie et al., 1974; Miller and Nicely, 1955). On the other hand, the place of articulation is relatively easy to identify visually. Therefore, acoustic and visual speech complement each other in an important way. What is hard to hear is easy to see. What is easy to hear needs not to be seen. Binnie et al. (1974) separated five visually distinct consonant clusters, i.e. *visemes*: bilabials, labiodentals, interdentals, rounded labials and linguals. Also other viseme classifications have been suggested (for a review, see Jackson, 1988). In Finnish, Pesonen (1968) demonstrated five consonantal visemes: labials (b, m, p), labiodentals (v, f), incisives (s, t), alveolar dentals (d, l, n, r) and velars (g, k, h). The number of distinct visemes is dependent on the individual talkers (Kricos, 1996). Auer and Bernstein (1997) suggest a measure (phoneme equivalence class – PEC) that reflects both individual speech perceiver skills and the visibility of the visual signal. Twelve PECs suffice to identify most American English words.

2.2 The McGurk Effect

Using videoclips of a single speaker, Harry McGurk and John MacDonald (MacDonald and McGurk, 1978; McGurk and MacDonald, 1976) dubbed the auditory token/ba/onto the visual articulation of/ga/, and something interesting happened. Most perceivers did not recognize the conflict between the stimuli. Moreover, they typically *heard* a new syllable /da/, or even /ga/. Very few perceivers were sensitive to the conflict and reported the acoustic syllable correctly. This change in auditory percept due to conflicting visual stimulus is called the "McGurk effect". An important feature of the McGurk effect is that it occurs when the acoustic stimulus is clear and can be easily recognized, when presented alone.

A strong fusion of acoustic and visual speech stimuli to a new percept does not occur for all stimulus combinations. For example, when Finnish acoustic /o/ (as in the English "dog") is dubbed onto visual /i/ (as in the English "lip"), it is easy to recognize that the two stimuli are discordant and they do not fuse (Klucharev et al., 2003). Strong fusions are obtained with some consonants. When the visual non-labial consonant



FIGURE 2 Examples of the McGurk effect. Photos show a talker who is articulating bilabial/p/and velar/k/sounds. When the articulatory gestures are presented simultaneously with conflicting acoustic stimuli, the auditory perception is modified, i.e. the McGurk effect is produced.

/ka/ is presented with the bilabial acoustic /pa/, people generally perceive either /ta/ or/ ka/(see Figure 2). That is, they *hear* the visual feature "not labial" in this combination. When visual /pa/ is combined with acoustic/ka/, a "combination" such as /pka/or /kpa/ is often *heard* (see Figure 2). The strength and quality of the McGurk effect depends on the vowel context in which the consonant is presented. For example, in American English the effect is stronger in /i/ and /a/ context than in /u/ context (Green, 1996). The nature of the fusion percept also depends on the vowel context. Visual /g/ combined with acoustic /b/ is often perceived as "th" in /a/ vowel context but "d" in /i/ context (Green et al., 1991). A conflicting visual consonant at the beginning of a word may even cause perceptual omission of a consonant. For example, the Finnish visual word /kuola/ combined with acoustic word /puola/ was often perceived as /uola/ (Sams et al., 1998).

2.3 Development of Audiovisual Speech Perception

Babbling is an important stage in the development of speech production and perception. Parents "baby talk" to their infants and also reinforce their utterances by repeating them. Some features of this preliminary speech are hard to hear but they are readily visible, such as bilabials. The syllables /ba/ and /ga/ are acoustically rather similar and difficult to differentiate by hearing, but the corresponding articulations are highly visible

and can be easily discriminated. Both deaf and blind children babble. However, relatively soon important differences can be observed in their babbling. The proportion of easy-to-see bilabials increases in deaf children, but in blind children it decreases. Blind children can face problems in learning such contrasts that are easy to see but hard to hear (Mills, 1987).

Children start to prefer audiovisual speech where the channels are congruent at an early age. Dodd (1979) showed that infants aged 10–20 weeks preferred to watch a face where lip movements were in rather that out of synchrony with the acoustic speech. In this study, the synchrony of the two stimulus sources could have been the critical cue for their common origin. When such a cue was not available and audiovisual vowels were presented in synchrony, 4-month-old infants were able to detect a correct combination of acoustically and visually presented vowels /i/ and /a/ (Kuhl and Meltzoff, 1982). Moreover, infants only 20 weeks old can match the face and voice in their native language, but not in a foreign language (Dodd and Burnham, 1988). Burnham (1998) concludes that the children develop the ability for matching acoustic and visible speech by the age of 3 months, and this ability is the clearest in the native tongue.

Strong evidence that children are really integrating acoustic and visual stimulation would be that they show the McGurk effect. This is the case in 4- to 5-month-old infants (Burnham and Dodd, 1996, 2004). McGurk and MacDonald (1976) compared the audiovisual integration of consonant–vowel syllables in pre-school children (3–5 years old), school children (7–8 years old) and adults (18–40 years old). The McGurk effect was stronger in adults than in the child groups, which did not differ from each other. This suggests that even though infants start to integrate at an early age, the effect of the visible speech on speech perception becomes greater with audiovisual speech experience.

2.4 Effects of Attention

The original study (McGurk and MacDonald, 1976) together with several other subsequent reports suggests that integration of auditory and visual information is strongly automatic since the McGurk effect occurs even when the observer is aware of how the stimuli are constructed. Furthermore, in a study on audio-visual speech perception in children, Massaro (1984) found that audiovisual percepts were very similar both when a child just reported what the talker said and when he also had to report whether the talker's mouth moved. The second task was used to ascertain that the child was paying attention to the visual information. Supposing that visual attention was stronger in the latter case, this result suggests that it does not influence perception of audiovisual speech.

Integration of heard and seen speech is often considered to occur in an obligatory and automatic manner. Soto-Faraco et al. (2004) recently studied a syllabic interference task, where two syllables are presented in rapid consecution. The second syllable only interferes with the classification of the first if it varies; not if it is constant. This happens even when the subjects are instructed to concentrate on the first syllable and ignore the irrelevant second syllable. This is interpreted as auditory attention failing to select the relevant stimulus so that the second syllable is obligatorily processed. The authors made a perceptual change to the second syllable by a conflicting visual syllable, causing the McGurk effect. Under these conditions, the illusory percept rather than the actual acoustic stimulus determined whether interference occurred. Soto-Faraco et al. (2004) concluded that the visual influence on the auditory speech percept must have occurred before attentional selection.

However, the issue of automaticity is not as clear-cut as it may seem. Massaro (1998) has shown that by instructing subjects to respond according to either auditory or visual information only, responses to audiovisual consonants are biased toward the instructed modality. Furthermore, the influence of incongruent visual speech on vowel recognition is greater when subjects are asked to watch and listen to and report what was spoken, than when they are told about the dubbing and asked to report only what they heard (Summerfield and McGrath, 1984).

In order to study the effect of endogenous visual attention on audiovisual speech perception, Tiippana et al. (2004) conducted an experiment in which visual attention was modulated using a distractor stimulus presented together with the talking face, and the McGurk effect was measured in two conditions: with full attention at the talking face and with attention directed towards the distractor. The McGurk effect was found to be weaker in the latter condition for many different combinations of auditory and visual tokens, indicating that visual attention modulated audiovisual speech perception.

Additional evidence of the strong influence of attention was obtained by Andersen et al. (2002). The visual stimulus was a movie of two faces displaced laterally on each side of a central fixation point. Each of the two faces articulated a different utterance. Subjects were instructed to always maintain fixation on the central fixation cross but to covertly attend to the face indicated by the cueing arrow. In audiovisual trials, where a voice was dubbed onto the movie, subjects responded according to what they heard. The voice and the attended face could be matching or conflicting, the latter combination producing a McGurk effect. The results were very clear. The direction of visual

attention strongly determined what the subjects heard. Moreover, Alsius et al. (2005) showed that a difficult concurrent visual or auditory discrimination task impairs audiovisual integration, as indexed by reduced McGurk effect. While audiovisual speech processing may not require extensive attentional resources, nevertheless it is sensitive to attentional contingencies and can be disrupted under specific attentional demand conditions.

2.5 Effects of Spatial and Temporal Asynchronies

A good example of the robustness of the McGurk effect was provided by Green et al. (1991). They dubbed a male voice onto a female talker and vice versa. Although the different genders were easy to recognize, this failed to reduce the strength of the McGurk effect. Interestingly, the McGurk effect is generated even when acoustic and visual speech inputs come from clearly different directions (Jones & Munhall, 1997). This finding is surprising, because a common spatial origin has been thought to be one of the main features of a multisensory event.

The acoustic and visual speech tokens do not need to be in perfect temporal synchrony for the McGurk effect to occur. The time window for the integration can be astonishingly wide. Green (1996) varied the asynchrony of acoustic and visual stimuli from 0 to 225 ms and found no effect on the strength of the McGurk effect. A wide integration window was also found by Munhall et al. (1996). The integration, as indicated by the strength of the McGurk effect, was the strongest with smaller asynchronies but a significant effect was obtained even when the auditory stimulus lagged the visual one by 240 ms. An important finding was that the integration window is strongly asymmetric. Much greater asynchrony is tolerated when the visual stimulus is leading.

3 • Neural Basis of Audiovisual Perception

As the behavioural findings reviewed above show the human brain is able to combine acoustic and visual inputs in order to create unified percepts. Modern non-invasive neuroimaging methods, e.g. electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), provide tools to study how the human brain accomplishes this task (for reviews, see Calvert, 2001; Calvert and Thesen, 2004). These methods provide complementary information about the human brain activity underlying audiovisual processing. EEG and MEG directly measure electric potentials and neuromagnetic fields generated by neural currents, providing information about the brain activity with millisecond accuracy. The haemodynamic responses measured

by fMRI and PET do not provide accurate information about the timing of brain activity. However, fMRI and PET are superior to both EEG and MEG in terms of spatial resolution.

3.1 Multisensory Neurons

Single-cell recording studies in non-human mammals have found multisensory neurons which are activated by inputs mediated by multiple sense organs. These kinds of neurons have been found at various levels of central nervous system in: (1) the subcortical structures (e.g. the superior colliculus (SC), Stein and Meredith, 1993; Stein et al., 2004), (2) the sensory-specific cortices (e.g. auditory belt and parabelt, Schroeder and Foxe, 2002; Schroeder et al., 2003) and (3) the heteromodal cortices (e.g. anterior and posterior regions of superior temporal sulcus (STS), for a review, see Cusick, 1997). These findings support the view that convergence of different sensory inputs to the same neurons enables interaction between sensory modalities.

The most thoroughly studied multisensory neurons are located in the mammalian SC (Stein and Meredith, 1993; Stein et al., 2004), which is a subcortical structure thought to be involved in orientation and attentive behaviours. Some of the multisensory neurons in the SC are able to integrate inputs from different sensory systems: two stimuli presented in the same location at the same time produce *response enhancement* in these neurons. The response to two simultaneous stimuli typically exceeds the sum of responses to the same stimuli presented separately. The weaker the unimodal stimuli the stronger the enhancements; this principle is called *inverse effectiveness*. In contrast, two stimuli presented in different locations (or at different times) produce *response suppression* in these neurons. There are parallels between the responses of the SC multisensory neurons and behaviour of the animal. For example, a cat detects and localizes visual targets more accurately when they are presented together with a spatially coherent sound (Stein et al., 1989). On the other hand, presenting a sound in a different location impairs detection of a visual target (Stein et al., 1989).

3.2 Audiovisual Interaction in the Human Auditory Cortex

A widely used method to study auditory processing in humans is to record the so-called "mismatch responses" to acoustic changes in sound sequences with EEG or MEG (for a review, see Näätänen, 1992). The mismatch responses are considered to be generated by pre-attentive change-detection mechanisms in the auditory cortices. An obvious question which arises is whether these mechanisms are activated when there is no acoustic change in the sound sequence, but a person perceives an illusory auditory change due to McGurk or ventriloquism illusion.

Sams et al. (1991) presented incongruent audiovisual stimuli (acoustic/pa/and visual/ka/), which were perceived as/ta/or/ka/, among frequent congruent syllables (acoustic/pa/and visual/pa/), and measured neuromagnetic responses over the left hemisphere. Perceptually deviant stimuli elicited mismatch responses peaking at 180 ms after stimulus onset in the left auditory cortex. A number of studies have replicated this finding (Möttönen et al., 2002; Colin et al., 2002a, 2004). Furthermore, it has been shown that illusory change in sound location due to ventriloquism illusion is detected in the auditory cortices (Stekelenburg et al., 2004, see also Colin et al., 2002b). These findings show that "illusory" changes in sound sequences are processed in the auditory cortices in a similar manner as "real" acoustic changes.

There is also evidence that the early auditory cortex responses to speech sounds are modulated during presentation of a visual speech stimulus (Klucharev et al., 2003; Besle et al., 2004; Möttönen et al., 2004; van Wassenhove et al., 2005). These findings, together with above-mentioned auditory change-detection studies, suggest that acoustic and visual speech inputs are integrated at an early level of sensory processing, within 200 ms after sound onset.

Which sub-regions of the auditory cortex visual speech has access to can be best studied using fMRI. Calvert et al. (1997) found that auditory regions of the human brain were robustly activated during silent speechreading. Activity was observed in the superior temporal cortex and in medial parts of Heschl's gyrus, where the primary auditory cortex is located. Several later studies have replicated the finding that visual speech has access to the auditory regions in the posterior superior temporal cortex (MacSweeney et al., 2000; Campbell et al., 2001; Calvert and Campbell, 2003; Santi et al., 2003; Pekkola et al., 2005). Some of these studies have found activity also in Heschl's gyrus (MacSweeney et al., 2000; Calvert and Campbell, 2003; Pekkola et al., 2005); however, some others have failed to see such activity during speechreading (Bernstein et al., 2002; Paulesu et al., 2003). Moreover, it has been demonstrated that speechreading can modify processing in the auditory brain stem (Musacchia et al., 2006). A brain stem response component peaking at around 11 ms was of longer latency and smaller amplitude when the acoustic speech stimulus was presented together with visual speech rather than alone.

In sum, the human auditory system is capable of processing visual input, not just acoustic input. Access of visual input into the early levels of the auditory processing stream might explain why vision can influence on auditory perception as demonstrated

by McGurk and ventriloquism illusions. It should however be noted that heteromodal regions in the human brain, such as STS, also participate in the integration of audiovisual speech and other audiovisual objects (Calvert et al., 2000; Raij et al., 2000; Sekiyama et al., 2003; Wright et al., 2003; Callan et al., 2004; Beauchamp et al., 2004; Möttönen et al., 2004; van Atteveldt et al., 2004). Furthermore, there is evidence that motor regions in the frontal lobe are involved especially in audiovisual speech processing (Ojanen et al., 2005; Skipper et al., 2005). Finding out how the different brain regions contribute to integration of inputs from different senses is a challenge for future research.

References

Alsius A, Navarra J, Campbell R, Soto-Faraco S (2005) Audiovisual integration of speech falters under high attention demands. Curr Biol 15:839–843.

Andersen TS, Tiippana K, Sams M (2002) Endogenous visual spatial attention affects audiovisual speech perception. International Multisensory Research Forum, 3rd Annual Meeeting, Geneva, Switzerland, May, 24th–26th.

Auer ET, Jr., Bernstein LE (1997) Speechreading and the structure of the lexicon: computationally modeling the effects of reduced phonetic distinctiveness on lexical uniqueness. J Acoust Soc Am 102:3704–3710.

Beauchamp MS, Lee KE, Argall BD, Martin A (2004) Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41:809–823.

Bernstein LE, Auer ET, Jr., Moore JK, Ponton CW, Don M, Singh M (2002) Visual speech perception without primary auditory cortex activation. Neuroreport 13:311–315.

Bertelson P, de Gelder B (2004) The psychology of multimodal perception. In: Crossmodal Space and Crossmodal Attention (Spence C, Driver J, eds), pp. 141–177. Oxford: Oxford University Press.

Besle J, Fort A, Delpuech C, Giard MH (2004) Bimodal speech: early suppressive visual effects in human auditory cortex. Eur J Neurosci 20:2225–2234.

Binnie CA, Montgomery AA, Jackson PL (1974) Auditory and visual contributions to the perception of consonants. J Speech Hear Res 17:619–630.

Burnham D (1998) Language specificity in the development of auditory-visual speech perception. In: Hearing by Eye 2: Advances in the Psychology Speechreading and Auditory-Visual Speech (Campbell R, Dodd B, Burnham D, eds). Hove, East Sussex, UK: Psychology Press Ltd.

Burnham D, Dodd B (1996) Auditory-visual speech perception as a direct process: the McGurk effect in infants and across languages. In: Speechreading by Humans and Machines: Models, Systems and Applications (Stork DG, Hennecke ME, eds), pp. 103–114. Berlin: Springer-Verlag.

Burnham D, Dodd B (2004) Auditory-visual speech integration by prelinguistic infants: perception of an emergent consonant in the McGurk effect. Dev Psychobiol 45:204–220.

Callan DE, Jones JA, Munhall K, Kroos C, Callan AM, Vatikiotis-Bateson E (2004) Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. J Cogn Neurosci 16:805–816.

Calvert GA (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb Cortex 11:1110–1123.

Calvert G, Campbell R (2003) Reading speech from still and moving faces: the neural substrates of visible speech. J Cogn Neurosci 15: 57–70.

Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. J Physiol Paris 98:191–205.

Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS (1997) Activation of auditory cortex during silent lipreading. Science 276:593–596.

Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Opin Biol 10:649–657.

Calvert G, Spence C, Stein BE (2004) The Handbook of Multisensory Processes. Cambridge, MA: The MIT Press.

Campbell R, Dodd B, Burnham D (1998) Hearing by Eye II: Advances in the Psychology of Speech-Reading and Audio-visual Speech. Hove: Psychology Press.

Campbell R, MacSweeney M, Surguladze S, Calvert G, McGuire P, Suckling J, Brammer MJ, David AS (2001) Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). Brain Res Cogn Brain Res 12:233–243.

Colin C, Radeau M, Soquet A, Dachy B, Deltenre P (2002a) Electrophysiology of spatial scene analysis: the mismatch negativity (MMN) is sensitive to the ventriloquism illusion. Clin Neurophysiol 113: 507–518.

Colin C, Radeau M, Soquet A, Demolin D, Colin F, Deltenre P (2002b) Mismatch negativity evoked by the McGurk–MacDonald effect: a phonetic representation within short-term memory. Clin Neurophysiol 113:495–506.

Colin C, Radeau M, Soquet A, Deltenre P (2004) Generalization of the generation of an MMN by illusory McGurk percepts: voiceless consonants. Clin Neurophysiol 115:1989–2000.

Cusick CG (1997) The superior temporal polysensory region in monkeys. In: Cerebral Cortex (Rockland KS, Kaas JH, Peters A, eds), pp. 435–468. New York: Plenum.

de Gelder B, Bertelson P (2003) Multisensory integration, perception and ecological validity. Trends Cogn Sci 7:460–467.

Dodd B (1979) Lip reading in infants: attention to speech presented in- and out-of- synchrony. Cognit Psychol 11:478–484.

Dodd B, Burnham D (1988) Processing speechread information. Volta Rev 90:45–60.
Dodd B, Campbell R (1987) Hearing by Eye: The Psychology of Lip-Reading. Hove: Lawrence Erlbaum.
Green KP (1996) The use of auditory and visual information in phonetic perception. In: Speechreading by Humans and Machines: Models, Systems, and Applications (Stork DG, Hennecke ME, eds), pp. 55–77. Berlin: Springer.
Green KP, Kuhl PK, Meltzoff AN, Stevens EB (1991) Integrating speech information across talkers, gender, and sensory modality: female faces and male voices in the McGurk effect. Percept Psychophys 50:524–536.
Howard IP, Templeton WB (1966) Human Spatial Orientation. London: Wiley.
Jackson PL (1988) The theoretical minimal unit for visual speech perception: visemes and coarticulation. Volta Rev 90:99–115.
Jones JA, Munhall KG (1997) The effects of separating auditory and visual sources on audiovisual integration of speech. Can Acoust 25: 13–19.
Klucharev V, Möttönen R, Sams M (2003) Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. Brain Res Cogn Brain Res 18:65–75.
Kricos PB (1996) Differences in visual intelligibility across talkers. In: Speechreading by Humans and Machines: Models, Systems, and Applications (Stork DG, Hennecke ME, eds), pp. 43–53. Berlin: Springer.
Kuhl PK, Meltzoff AN (1982) The bimodal perception of speech in infancy. Science 218:1138–1141.
MacDonald J, McGurk H (1978) Visual influences on speech perception processes. Percept Psychophys 24:253–257.
MacLeod A, Summerfield Q (1987) Quantifying the contribution of vision to speech perception in noise. Br J Audiol 21:131–141.
MacSweeney M, Amaro E, Calvert GA, Campbell R, David AS, McGuire P, Williams SC, Woll B, Brammer MJ (2000) Silent speechreading in the absence of scanner noise: an event-related fMRI study. Neuroreport 11:1729–1733.
Massaro DW (1984) Children's perception of visual and auditory speech. Child Dev 55:1777–1788.
Massaro DW (1998) Perceiving talking faces. Cambridge, MA: MIT Press.
McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264:746–748.
Miller GA, Nicely PE (1955) An analysis of perceptual confusions among some English consonants. J Acoust Soc Am 27:338–352.
Mills AE (1987) The development of phonology in the blind child. In: Hearing by Eye: The Psychology of Lipreading (Dodd B, Campbell R, eds), pp. 145–161. London: Lawrence Erlbaum.
Möttönen R (1999) Perception of Natural and Synthetic Audio-visual Finnish Speech. Master's Thesis, Department of Psychology, University of Helsinki.

Möttönen R, Krause CM, Tiippana K, Sams M (2002) Processing of changes in visual speech in the human auditory cortex. Brain Res Cogn Brain Res 13:417–425.

Möttönen R, Schürmann M, Sams M (2004) Time course of multisensory interactions during audiovisual speech perception in humans: a magnetoencephalographic study. Neurosci Lett 363:112–115.

Munhall KG, Gribble P, Sacco L, Ward M (1996) Temporal constraints on the McGurk effect. Percept Psychophys 58:351–362.

Musacchia G, Sams M, Nicol T, Kraus N (2006) Seeing speech affects acoustic information processing in the human brainstem. Exp Brain Res 168:1–10.

Näätänen R (1992) Attention and Brain Function. Hillsdale, NJ: Lawrence Erlbaum.

Ojanen V, Möttönen R, Pekkola J, Jääskeläinen IP, Joensuu R, Autti T, Sams M (2005) Processing of audiovisual speech in Broca's area. Neuroimage 25:333–338.

O'Neill JJ (1954) Contributions of the visual components of oral symbols to speech comprehension. J Speech Hear Disord 19:429–439.

Paulesu E, Perani D, Blasi V, Silani G, Borghese NA, De Giovanni U, Sensolo S, Fazio F (2003) A functional–anatomical model for lipreading. J Neurophysiol 90:2005–2013.

Pekkola J, Ojanen V, Autti T, Jääskeläinen IP, Möttönen R, Tarkiainen A, Sams M (2005) Primary auditory cortex activation by visual speech: an fMRI study at 3 T. Neuroreport 16:125–128.

Pesonen J (1968) Phoneme Communication of the Deaf. Helsinki: Suomalaisen kirjallisuuden kirjapaino Oy.

Raij T, Uutela K, Hari R (2000) Audiovisual integration of letters in the human brain. Neuron 28:617–625.

Sams M, Aulanko R, Hamalainen M, Hari R, Lounasmaa OV, Lu ST, Simola J (1991) Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. Neurosci Lett 127: 141–145.

Sams M, Manninen P, Surakka V, Helin P, Kättö R (1998) McGurk effect in Finnish syllables, isolated words, and words in sentences: effects of word meaning and sentence context. Speech Commun 26:75–87.

Santi A, Servos P, Vatikiotis-Bateson E, Kuratate T, Munhall K (2003) Perceiving biological motion: dissociating visible speech from walking. J Cogn Neurosci 15:800–809.

Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Brain Res Cogn Brain Res 14:187–198.

Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN, Hackett TA (2003) Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. Int J Psychophysiol 50:5–17.

Sekiyama K, Kanno I, Miura S, Sugita Y (2003) Audio-visual speech perception examined by fMRI and PET. Neurosci Res 47: 277–287.

Skipper JI, Nusbaum HC, Small SL (2005) Listening to talking faces: motor cortical activation during speech perception. Neuroimage 25:76–89.

Slutsky DA, Recanzone GH (2001) Temporal and spatial dependency of the ventriloquism effect. Neuroreport 12:7–10.

Soto-Faraco S, Navarra J, Alsius A (2004) Assessing automaticity in audiovisual speech integration: evidence from the speeded classification task. Cognition 92:B13–B23.

Stein BE (1998) Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. Exp Brain Res 123:124–135.

Stein BE, Meredith MA (1993) Merging of the Senses. Cambridge, MA: The MIT Press.

Stein BE, Meredith MA, Huneycutt WS, McDale L (1989) Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. J Cogn Neurosci 1:12–24.

Stein BE, Jiang H, Stanford TE (2004) Multisensory integration in single neurons of the midbrain. In: The Handbook of Multisensory Processes (Calvert G, Spence C, Stein BE, eds), pp. 243–264. Cambridge, MA: The MIT Press.

Stekelenburg JJ, Vroomen J, de Gelder B (2004) Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. Neurosci Lett 357:163–166.

Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26:212–215.

Summerfield Q, McGrath M (1984) Detection and resolution of audio-visual incompatibility in the perception of vowels. Q J Exp Psychol 36A:51–74.

Tiippana K, Andersen TS, Sams M (2004) Visual attention modulates audiovisual speech perception. Eur J Cogn Psychol 16:457–472.

van Atteveldt N, Formisano E, Goebel R, Blomert L (2004) Integration of letters and speech sounds in the human brain. Neuron 43:271–282.

van Wassenhove V, Grant KW, Poeppel D (2005) Visual speech speeds up the neural processing of auditory speech. Proc Natl Acad Sci U S A 102:1181–1186.

Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004) Unifying multisensory signals across time and space. Exp Brain Res 158:252–258.

Welch RB, Warren DH (1986) Intersensory interactions. In: Handbook of Perception and Human Performance: Vol. 1. Sensory Processes and Perception (Boff KR, Kaufman L, Thomas JP, eds), pp. 25.21–25.36. New York: Wiley.

Wright TM, Pelphrey KA, Allison T, McKeown MJ, McCarthy G (2003) Polysensory interactions along lateral temporal regions evoked by audiovisual speech. Cereb Cortex 13:1034–1043.

38

Multichannel Sound Reproduction

Ville Pulkki

Department of Signal Processing and Acoustics, Helsinki University of Technology, Helsinki, Finland, e-mail: Ville.Pulkki@tkk.fi

1 • Introduction

Spatial reproduction of sound is a field in which the spatial attributes of a real recording room or a virtual space are reproduced to the listener. Spatial attributes include for example directions of sound sources, directions of reflections and envelopment by reverberation. Many such systems employ more than two loudspeakers to create virtual sources. This is called multichannel sound or spatial sound reproduction. Binaural techniques also try to recreate sound faithfully but only in listeners' ear canals. These techniques employ in most cases two audio channels through headphone or loudspeaker reproduction. In this chapter, the main methods for spatial reproduction of sound are reviewed, and the quality that can be achieved with them is discussed.

2 • Concepts of Spatial Hearing by Humans

As a sound signal travels from a sound source to the ear canals of a listener, the signals in both ear canals will be different from the original sound signal and from each other. The transfer functions from sound source to the ear canals are called the head-related transfer functions (HRTF) [1]. They are dependent

on the direction of a sound source, and they yield temporal and spectral differences between left and right ear canals. Humans decode these differences and use them to localize sound sources. These differences are called binaural directional cues. Temporal difference is called the interaural time difference (ITD), and spectral difference is called the interaural level difference (ILD). ITD and ILD provide information on where a sound source is in the left vs. right direction dimension. The angle between the sound source direction and the median plane can thus be decoded. The median plane is the vertical plane which divides the space related to a listener to left and right parts. To decode the direction more accurately, monaural spectral cues are applied, and the effect of head rotation on binaural cues is taken into account [1,2]. The precedence effect [1,3] is a suppression of early delayed versions of the direct sound in source direction perception. This helps in reverberant rooms to localize sound sources.

3 • Loudspeaker Layouts

In the history of multichannel audio [4–6], multiple different loudspeaker layouts have been specified. The most frequently used layouts are presented in this chapter. The most common loudspeaker layout is the two-channel setup, called the standard stereophonic setup. It was widely taken into use after the development of single-groove 45°/45° two-channel record in the late 1950s. Two loudspeakers are positioned in front of the listener, separated by 60° from the listener's viewpoint, as presented in Figure 1. In the 1970s, the quadraphonic setup was proposed, in which four loudspeakers are placed evenly around the listener in azimuth angles ±45° and ±135°. This layout was never successful because of the problems related to reproduction techniques of that time and because the layout itself has too few loudspeakers to provide good spatial quality to all directions around the listener [7].

For cinema sound, a system was evolved in which the frontal image stability of standard stereophonic setup was enhanced by one extra center channel, and two surround channels are used to create atmospheric effects and room perception. This kind of a setup was first used in Dolby's surround sound system for cinemas from 1976 [5]. Later, the layout was investigated [8], and ITU gave a recommendation about the layout in 1992 [9]. In the late 1990s, this layout became common also in domestic use. It is widely referred to as the 5.1 system, where 5 stands for the number of loudspeakers, and .1 stands for the low-frequency

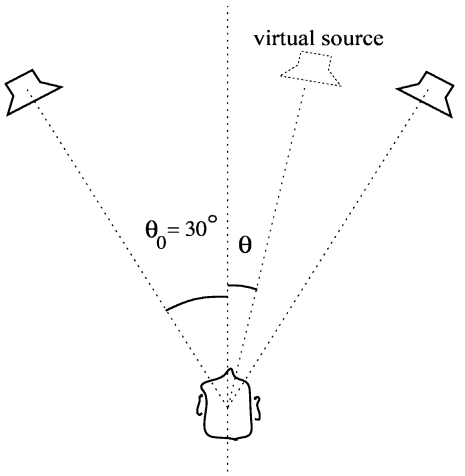


FIGURE 1 Standard stereophonic listening configuration.

channel. In the recommendation, three frontal loudspeakers are in directions 0° and ±30°, and two surround channels in ±110± 10°, as shown in Figure 2. The system has been criticized that it can not deliver good directional quality elsewhere than in the front [7]. To achieve better quality, it can be extended by adding loudspeakers. Layouts with 6–12 loudspeakers have been proposed and are presented in [7].

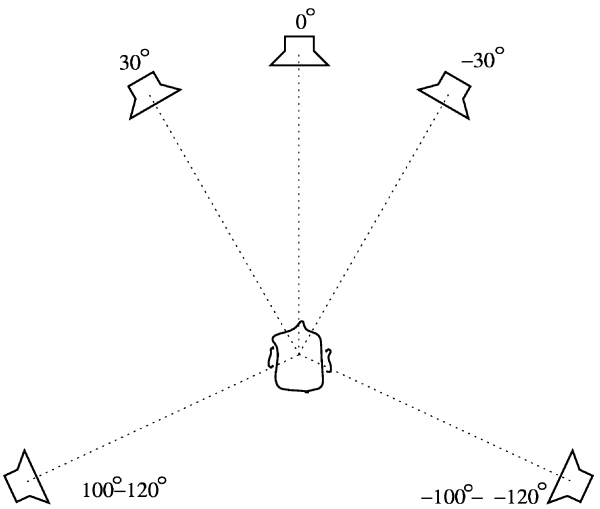


FIGURE 2 Five-channel surround loudspeaker layout based on ITU recommendation BS775.1.

In computer music, media installations and academic projects, loudspeaker setups, in which the loudspeakers are in equal spacing, have been used. In horizontal arrays, the number of loudspeakers can be, for example, six (hexagonal array) or eight (octagonal array). In wave field synthesis, see Sect. 4.3, the number of loudspeakers is typically between 20 and 200. In theaters and in virtual environment systems, there exist systems in which loudspeakers are placed also above and/or below the listener.

4 • Reproduction Methods

In this chapter, reproduction methods which use the loudspeaker layouts presented in the previous section are reviewed. The methods aim to position monophonic sound tracks in arbitrary directions. The directions do not necessarily coincide with any loudspeaker, as shown in Figure 1. Such perceived sound sources are called virtual sources.

4.1 Amplitude Panning

Amplitude panning is the most frequently used virtual source positioning technique. In it, a sound signal is applied to loudspeakers with different amplitudes, which can be formulated as

$$x_i(t) = g_i x(t), \ i = 1, \ldots, N, \tag{1}$$

where $x_i(t)$ is the signal to be applied to loudspeaker i, g_i is the gain factor of the corresponding channel, N is the number of loudspeakers, and t is the time variable. The listener perceives a virtual source, the direction of which is dependent on the gain factors.

4.1.1 Two-Channel Stereophony

In stereophony, two loudspeakers are placed in front of a listener, (Figure 1). Variable θ denotes the perceived azimuth of the virtual source. A panning law estimates θ from the gain factors of loudspeakers. The estimated direction is called the panning direction or panning angle. In [10], it has been found that amplitude panning provides consistent ITD cues up to 1.1 kHz, and roughly consistent ILD cues above 2 kHz for a listener in the best listening position. Outside the best listening position, the virtual source is localized toward the nearmost loudspeaker which emanates sound due to the precedence effect.

In the sine law [11] of perceived direction, the wave propagation path from loudspeakers to ears was modeled with straight lines. The path from the contralateral loudspeaker therefore

penetrates the listener's head, which is highly unnatural. It is presented as

$$\frac{\sin\theta_S}{\sin\theta_0} = \frac{g_1 - g_2}{g_1 + g_2}, \tag{2}$$

where θ_S is an estimate of perceived azimuth angle (panning angle) of a virtual source. θ_0 is the loudspeaker base angle, as in Figure 1. The equation does not set limitations to θ_S, but in most cases, its value is set to satisfy $|\theta_S| \leq \theta_0$. If $|\theta_S| > \theta_0$, the amplitude panning will produce antiphase loudspeaker signals which may distort the virtual source [1].

Bennett et al. [12] derived the tangent law by approximating the propagation path from contralateral loudspeaker to ear with a curved line around the head

$$\frac{\tan\theta_T}{\tan\theta_0} = \frac{g_1 - g_2}{g_1 + g_2}, \tag{3}$$

which has been found to estimate perceived direction best in listening tests in anechoic listening [10]. There are also other panning laws, reviewed in [10].

The panning laws set only the ratio between the gain factors. To prevent undesired change in loudness of virtual source depending on panning direction, the sum of squares of gain factors should be normalized:

$$\sqrt{\sum_{n=1}^{N} g_n^2} = 1. \tag{4}$$

This normalization equation is used in real rooms with some reverberation. Depending on listening room acoustics, different normalization rules may be used [13].

4.1.2 2-D Loudspeaker Setups

In 2-D loudspeaker setups, all loudspeakers are on the horizontal plane. Pair-wise amplitude panning [14] methods can be used in such reproduction systems. The sound signal is applied only to two adjacent loudspeakers of the loudspeaker setup at one time. The pair between which the panning direction lies is selected. Different formulations for pair-wise panning are Chowning's law [14], which is not based on any psychoacoustic criteria, or 2-D vector base amplitude panning (VBAP) [15], which is a generalization of the tangent law (Eq. 3) for stereophonic panning.

In VBAP, a loudspeaker pair is specified with two vectors. The unit-length vectors $\mathbf{l}_m$ and $\mathbf{l}_n$ point from the listening position to the loudspeakers. The intended direction of the virtual source

(panning direction) is presented with a unit-length vector $\mathbf{p}$. Vector $\mathbf{p}$ is expressed as a linear weighted sum of the loudspeaker vectors

$$\mathbf{p} = g_m \mathbf{l}_m + g_n \mathbf{l}_n. \tag{5}$$

Here g_m and g_n are called gain factors of respective loudspeakers. The gain factors can be solved as

$$\mathbf{g} = \mathbf{p}^{\mathrm{T}} \mathbf{L}_{mn}^{-1}, \tag{6}$$

where $\mathbf{g} = [g_m \;\; g_n]^{\mathrm{T}}$ and $\mathbf{L}_{mn} = [\mathbf{l}_m \;\; \mathbf{l}_n]$. The calculated factors are used in amplitude panning as gain factors of the signals applied to respective loudspeakers after suitable normalization, e.g., $||\mathbf{g}|| = 1$.

The directional quality achieved with pair-wise panning was studied in [10]. When the loudspeakers are symmetrically placed on the left and right of the listener, VBAP estimates the perceived angle accurately. When the loudspeaker pair is not symmetrical with the median plane, the perceived direction is biased toward the median plane [10], which can be more or less compensated [16].

When there is a loudspeaker in the panning direction, the virtual source is sharp, but when panned between loudspeakers, the binaural cues are unnatural in some degree. This means that the directional perception of the virtual source varies with panning direction, which can be compensated by applying sound always to more than one loudspeaker [17,18]. Since the perceived direction collapses to the nearmost loudspeaker producing a specific sound, the maximal directional error is of the same order of magnitude with the angular separation of loudspeakers from the listener's viewpoint in pair-wise panning. The curvature of the wavefront generated with pair-wise panning cannot be controlled.

4.1.3 3-D Loudspeaker Setups

A 3-D loudspeaker setup denotes here a setup in which all loudspeakers are not in the same plane with the listener. Typically, this means that there are some elevated and/or lowered loudspeakers added to a horizontal loudspeaker setup. Triplet-wise panning can be used in such setups [15]. In it, a sound signal is applied to a maximum of three loudspeakers at one time that form a triangle from the listener's viewpoint. If more than three loudspeakers are available, the setup is divided into triangles, one of which is used in the panning of a single virtual source at one time, as shown in Figure 3. Three-dimensional vector base amplitude panning (3-D VBAP) is a method to formulate such setups [15]. It is formulated in an equivalent way as with pair-wise panning

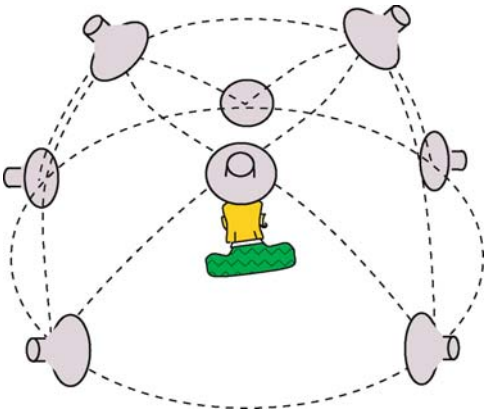


FIGURE 3 A 3-D triangulated loudspeaker system for triplet-wise panning.

in the previous section. However, the number of gain factors and loudspeakers is naturally three in the equations. The angle between the median plane and virtual source is estimated correctly with VBAP in most cases, as in pair-wise panning. However, the elevation of a sound source is individual to each subject [10].

4.1.4 Ambisonics

Ambisonics is basically a microphone technique [19]. However, it can also be simulated to perform synthesis of spatial audio [20] to 2- or 3-D loudspeaker setups. In this case, it is an amplitude panning method in which a sound signal is applied to all loudspeakers placed evenly around the listener with gain factors

$$g_i = \frac{1}{N} \sum_{m=1}^{M} \{1 + 2p_m \cos(m\alpha_i)\}, \tag{7}$$

where g_i is the gain of ith speaker, N is the number of loudspeakers, α is the angle between loudspeaker and panning direction, $\cos(m\alpha_i)$ represents a single spherical harmonic with order of m, M is the order of Ambisonics and p_m are the gains for each spherical harmonic [21,22]. When the order is low, sound signal emanates from all loudspeakers, which causes some spatial artifacts due to unnatural behavior of binaural cues [23]. In such cases also, when listening outside the best listening position, the sound is perceived at the nearmost loudspeaker which produces the sound. This effect is more prominent with first-order Ambisonics than with pair- or triplet-wise panning, since in Ambisonics, virtually all loudspeakers produce the same sound signal. The sound is also colored, for the same reason, i.e., multiple propagation paths of the same signal to the ears produce comb-filter effects. Conventional microphones can be used to realize first-order Ambisonics, as shortly reviewed in Sect. 5.

When the order is increased, the cross talk between loudspeakers can be minimized by optimizing gains of spherical harmonics for each loudspeaker in a listening setup [24]. Using higher order spatial harmonics increases both directional and timbral virtual source quality, since there is less cross talk between loudspeakers. The physical wave field reconstruction is then more accurate, and different curvatures of wavefronts, as well as planar wavefronts can be produced [21], if a large enough number of loudspeakers is in use.

4.2 Time Panning

When a constant delay is applied to one loudspeaker in stereophonic listening, the virtual source is perceived to migrate toward the loudspeaker that radiates the earlier sound signal [1]. Maximal effect is achieved asymptotically when the delay is approximately 1.0 ms or more.

In tests with different signals, it has been found, however, that the perceived direction of virtual sources is dependent on frequency [25,26]. The produced binaural cues vary with frequency, and different cues suggest different directions for virtual sources [27]. It may thus generate a "spread" perception of direction of sound, which is desirable in some cases. The effect is dependent on listening position. For example, if the sound signal is delayed by 1 ms in one loudspeaker, the listener can compensate the delay by moving 30 cm toward the delayed loudspeaker.

4.3 Wave Field Synthesis

Wave field synthesis is a technique that requires a large number of carefully equalized loudspeakers [28,29]. It targets to reconstruct the whole sound field in a listening room. When a virtual source is reproduced, the sound for each loudspeaker is delayed and amplified in a manner that a desired circular or planar sound wave occurs as a superposition of sounds from each loudspeaker. The virtual source can be positioned far behind the loudspeakers, or in some cases, even in the space inside the loudspeaker array, as shown in Figure 4. The loudspeaker signals have to be equalized depending on virtual source position [30].

Theoretically, the wave field synthesis is superior as a technique, the perceived position of the sound source is correct within a very large listening area. Unfortunately, to create a desired wave field in the total area inside the array, it demands that the loudspeakers are at a distance of maximally a half wavelength from each other. The area where a perfect sound field synthesis is achieved shrinks with increasing frequency [21]. In practice, due to the perception mechanisms of humans, more error can be allowed above approximately 1.5 kHz. Arrays for wave field synthesis have been built for room acoustics control

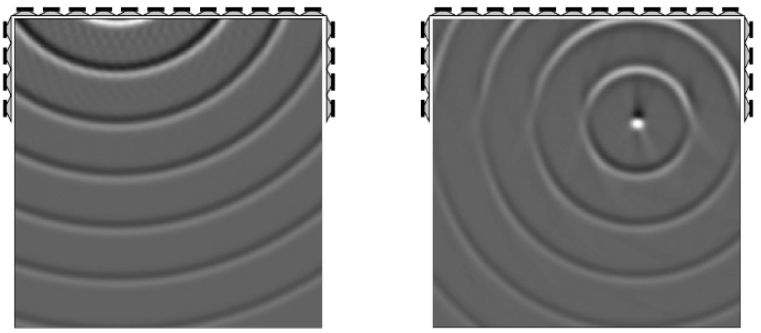


FIGURE 4 Wave field synthesis concept. A desired 2-D sound field can be constructed using a large loudspeaker array. Figure reprinted from [29] with permission.

and enhancement to be used in theatres and multipurpose auditoria [29].

4.4 Binaural Reproduction

Binaural reproduction is mentioned here for completeness, although it is reviewed in more detail in Chapter 11.1, part II. In headphone listening, a monophonic sound signal can be positioned virtually in any direction, if HRTFs for both ears are available for the desired virtual source direction [31,32]. A sound signal is filtered with a digital filter modeling the measured HRTFs. The method simulates the ear canal signals that would have been produced if a sound source existed in a desired direction. If a listener moves his/her head during listening, then the movements of his/her should also be taken into account in processing, otherwise the sound stage will be moving along when the listener rotates his/her head and that may cause inside-head localization.

If the cross talk between a loudspeaker and the contralateral ear in stereophonic setup is canceled by feeding an antiphase signal to the other loudspeaker in a carefully controlled manner, it is possible to control precisely the sound signals in a listener's ear canals [32–35]. However, the best effect is audible only in a small listening area, and the listener should keep his ears in constant distance from the loudspeakers with about 1 cm accuracy. The listening room should also be anechoic, since acoustic reflections deteriorate cross talk canceling.

5 • Capturing Sound for Multichannel Audio

Techniques, in which a microphone is positioned close to a sound source, capture the direct sound with prominently higher amplitude than the reverberation of the room. In mixing

phase, the recorded sound tracks of different sound sources are positioned with a multichannel loudspeaker system using a desired panning technique. However, often it is desired to also introduce reverberation of the room into the mixed signals. It can be captured with distant microphones, and it can be introduced to the sound stage in the mixing. Alternatively, a desired room effect can be introduced using digital reverberators, which either mimic room response with a digital filter structure or use measured room impulse responses in convolution [36,37].

The microphones which have flat frequency response over audible frequency range have first-order directional patterns, such as figure-of-eight, cardioid or hypercardioid, or they have zeroth-order pattern, which means the omnidirectional microphone. Figure-of-eight microphones and hypercardioids have two lobes in their directional patterns. One of the lobes reproduces the sound in correct phase and is called the positive lobe, and the other lobe reproduces the sound phase reversed, which is called the negative lobe. Higher order microphones suffer from coloration in sound spectrum.

When the microphones are relatively far from sound sources, the acoustical influence of the recording room is prominent. There are different approaches to record for a multichannel loudspeaker system in such cases. In the coincident microphone techniques, two or more directive microphones are positioned as close to each other as possible but pointing to different directions [26,38,39]. The microphone signals thus differ generally only in amplitude, in a fashion similar to amplitude panning presented in Sect. 4.1. Only in cases where dipole microphones are used, the negative lobes may introduce 180° phase shifts between microphone channel signals.

In spaced microphone techniques, there are considerable distances between microphones [7,39]. These techniques produce both amplitude and phase differences between microphone channels. They can be treated as a combination of amplitude panning and time panning (Sects. 4.1 and 4.2).

5.1 Microphone Techniques for Stereophonic Setup

The most significant microphone techniques for the two-channel stereophonic setup are presented here briefly. An analysis of the presented techniques computed with a binaural auditory model of the presented techniques has been published in [40]

Coincident techniques are often denoted XY techniques. If cardioids or hypercardioids are used, the angle between microphones is typically 60°–120°, although angles upto 180° may be used [39]. It produces good directional accuracy [26]. However, these techniques have been characterized to lack “sense of space” [39].

Two figure-of-eight microphones in XY positioning was the first XY technique presented by Blumlein [38]. In it, the positive lobes of dipole microphones are directed 45° and −45°, respectively. It reproduces equally the sound energy from all directions. However, the sound arriving from side produces 180° phase reversal between the microphone signals. The sound is thus reproduced in loudspeakers out of phase, which may produce coloration and directional artifacts.

MS stereo was also presented by Blumlein [38]. In it, a directive microphone (M) points to the front and a figure of eight microphones (S) to the side. Later, weighted sums are performed from the signals, which create two virtual microphone signals. The directional patterns of these microphones can be adjusted as needed in mixing phase by changing the weights in summing.

In spaced microphone techniques, the distance between microphones is typically between 60 and 100 cm. The microphones can be either omnidirectional or directional microphones. The spacing between the microphones gives better feeling of "ambience" or "air" [26], however it deteriorates the directional accuracy and may produce some comb-filter effects.

In near coincindent microphone techniques, the distance between microphones is selected in a way that the microphone signals are effectively in same phase at low frequencies, and in different phase at high frequencies. This can be thought to compromise coincident and spaced techniques: they provide moderately good directional quality combined with better "sense of depth" [39]. Probably the most used near coincident technique is the ORTF technique, in which two cardioid microphones are spaced by 17 cm in left–right dimension and face to ±110°.

5.2 Microphone Techniques for 5.1 Loudspeaker Setup

A fundamental problem with microphone techniques for the 5.1 loudspeaker setup is that good-quality microphones are available only with directional patterns of zeroth or first order. When such microphones are used with a 5.1 setup, there will be a considerable amount of cross talk between loudspeaker channels, which results in coloration and directional artifacts. To overcome this, a microphone has been recently proposed which consists of eight omnidirectional microphones and postprocessing which composes virtual microphones with moderately narrow directional patterns [41]. The too high coherence produced by first-order directional patterns can be enhanced using information of directional properties of sound field in frequency bands [46], or by cardioid microphones pointing to opposite direction combined with digital signal processing [47].

In practice, the recording is typically performed by positioning omnidirectional or directional microphones with distances of

10 cm to 1.5 m from each other and by directing the microphones in a desired manner. There is no generally accepted theory for an optimal setup, although some heuristic models to formalize common practices have been presented [42]. A large number of spaced microphone techniques have been tested, which are reviewed by Rumsey [7].

5.3 Microphone Techniques for Loudspeaker Arrays

Both higher order Ambisonics and wave field synthesis utilize large horizontal arrays of loudspeakers. Higher order Ambisonics is based on measuring the sound field with a set of highly directional microphones in a single point. Unfortunately such microphones do not exist as single units. It has been however proven that microphone arrays can be used for recording for high-order Ambisonics [43].

Wave field synthesis requires capturing the sound field with a large number of microphones [28]. Circular microphone arrays have been found superior to linear arrays [44]. The measured impulse responses with such arrays have been used in the analysis of room acoustics, which is called wave field analysis [45].

References

1. J. Blauert. *Spatial Hearing, Revised edition*. MIT Press, Cambridge, MA, 1997.
2. R. H. Gilkey and T. R. Anderson, editors. *Binaural and Spatial Hearing in Real and Virtual Environments*. Lawrence Erlbaum, Mahwah, NJ, 1997.
3. P. M. Zurek. The precedence effect. In W. A. Yost and G. Gourewitch, editors, *Directional Hearing*, p. 3–25. Springer, New York, 1987.
4. G. Steinke. Surround sound — the new phase. an overview. In *Audio Engineering Society 100th Convention Preprint # 4286*, Copenhagen, Denmark, 1996.
5. M. F. Davis. History of spatial coding. *J. Audio Eng. Soc.*, 51(6): 554–569, June 2003.
6. E. Torick. Highlights in the history of multichannel sound. *J. Audio Eng. Soc.*, 46(1/2):27–31, January/February 1998.
7. F. Rumsey. *Spatial Audio*. Focal Press, Oxford, England, 2001.
8. G. Theile. HDTV sound systems: How many channels? In *Proceedings of AES 9th International Conference "Television Sound Today and Tomorrow"*, p. 217–232, Detroit, MI, Feb. 1–2 1991.
9. ITU-R Recommendation BS.775-1. Multichannel stereophonic sound system with and without accompanying picture. Technical report, International Telecommunication Union, Geneva, Switzerland, 1992–1994.
10. V. Pulkki. *Spatial Sound Generation and Perception by Amplitude Panning Techniques*. Doctoral thesis, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland, August 2001.

11. B. B. Bauer. Phasor analysis of some stereophonic phenomena. *J. Acoust. Soc. Am.*, 33(11):1536–1539, November 1961.
12. J. C. Bennett, K. Barker, and F. O. Edeko. A new approach to the assessment of stereophonic sound system performance. *J. Audio Eng. Soc.*, 33(5):314–321, May 1985.
13. F. R. Moore. *Elements of Computer Music*. Prentice Hall, Englewood Cliffs, NJ, 07632, 1990.
14. J. Chowning. The simulation of moving sound sources. *J. Audio Eng. Soc.*, 19(1):2–6, 1971.
15. V. Pulkki. Virtual source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.
16. V. Pulkki. Compensating displacement of amplitude-panned virtual sources. In *the AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, p. 186–195, Espoo, Finland, June 15–17 2002.
17. V. Pulkki. Uniform spreading of amplitude panned virtual sources. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, New Paltz, NY, 1999.
18. R. Sadek and C. Kyriakakis. A novel multichannel panning method for standard and arbitrary loudspeaker configurations. In *AES 117th Convention paper # 6263*, San Francisco, CA, 2004.
19. M. A. Gerzon. Panpot laws for multispeaker stereo. In *The 92nd Convention 1992 March 24–27 Vienna*. Audio Engineering Society, Preprint No. 3309, 1992.
20. D. G. Malham and A. Myatt. 3-d sound spatialization using ambisonic techniques. *Comput. Music J.*, 19(4):58–70, 1995.
21. J. Daniel, R. Nicol, and S. Moreau. Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging. In *Proceedings of 114th Audio Engineering Society Convention*, 2003. Paper # 5788.
22. G. Monro. In-phase corrections for ambisonics. In *Proceedings of Internation Computer Music Conference*, p. 292–295, Berlin, Germany, 2000.
23. V. Pulkki and T. Hirvonen. Localization of virtual sources in multichannel audio reproduction. *IEEE Trans Speech Audio Proc.*, 2005.
24. P. G. Craven. Continuous surround panning for 5-speaker reproduction. In *AES 24th Int. Conf. Multichannel Audio*, 2003.
25. D. H. Cooper. Problems with shadowless stereo theory: Asymptotic spectral status. *J. Audio Eng. Soc.*, 35(9):629–642, September 1987.
26. S. P. Lipshitz. Stereophonic microphone techniques. Are the purists wrong? *J. Audio Eng. Soc.*, 34(9):716–744, 1986.
27. V. Pulkki, M. Karjalainen, and J. Huopaniemi. Analyzing virtual sound source attributes using a binaural auditory model. *J. Audio Eng. Soc.*, 47(4):203–217, April 1999.
28. A. J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.*, 93(5), May 1993.
29. D. Vries and M. Boone. Wave field synthesis and analysis using array technology. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, p. 15–18, Mohonk Mountain House, New Paltz, NY 1999.

30. A. Apel, T. Röder, and S. Brix. Equalization of wave field synthesis systems. In *Audio Eng. Soc. 116th Convention Paper # 6121*, 2004.
31. H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43(5):300–321, May 1995.
32. D. R. Begault. *3-D Sound for Virtual Reality and Multimedia*. AP Professional, Cambridge, MA, 1994.
33. D. H. Cooper and J. L. Bauck. Prospects for transaural recording. *J. Audio Eng. Soc.*, 37(1/2):3–39, January/February 1989.
34. B. Gardner. *3-D Audio Using Loudspeakers*. PhD thesis, Massachusetts Institute of Technology, Massachusetts, USA, 1997.
35. O. Kirkeby, P. A. Nelson, and H. Hamada. Local sound field reproduction using two closely spaced loudspeakers. *J. Acoust. Soc. Am.*, 104:1973–1981, October 1998.
36. B. Gardner. Reverberation algorithms. In M. Kahrs and K. Brandenburg, editors, *Applications of Digital Signal Processing to Audio and Acoustics*. Kluwer Academic Publishers, Norvell, MA, 1998.
37. J. Merimaa and V. Pulkki. Spatial impulse response rendering I: Analysis and synthesis. *J. Audio Eng. Soc.*, 53(12):1115–1127, 2005.
38. A. D. Blumlein. U.K. Patent 394,325, 1931. Reprinted in Stereophonic Techniques, Audio Eng. Soc., NY, 1986.
39. R. Streicher and W. Dooley. Basic stereo microphone perspectives – a review. *J. Audio Eng. Soc.*, 33(7/8), 1985.
40. V. Pulkki. Microphone techniques and directional quality of sound reproduction. In *Proceedings of 112th AES Convention, Preprint # 5500*, Munich, Germany, 2002. Audio Engineering Society.
41. A. Laborie, R. Bruno, and S. Montoya. Designing high spatial resolution microphones. In *Audio Eng. Soc. 117th Convention Paper # 6231*, 2003.
42. M. Williams. Multichannel sound recording using 3, 4 and 5 channel arrays for front sound stage coverage. In *Proceedings of 117th AES Convention, Preprint # 6230*, San Francisco, CA, 2004.
43. R. Nicol and M. Emerit. 3d-sound reproduction over an extensive listening area: a hybrid method derived from holophony and ambisonic. In *Proceedings of AES 16th International Conference*, p. 436–453, 1999.
44. E. Hulsebos, D. de Vries, and E. Bourdillat. Improved microphone array configurations for auralization of sound fields by wave-field synthesis. *J. Audio Eng. Soc.*, 50(10):779–790, October 2002.
45. A. Berkhout, D. de Vries, and J. Sonke. Array technology for acoustic wave field analysis in enclosures. *J. Acoust. Soc. Am.*, 102(5): 2757–2770, 1997.
46. V. Pulkki. Spatial sound reproduction with directional audio coding. *J. Audio Eng. Soc.*, 55(6):503–516, 2007.
47. C. Faller. A highly directive 2-capsule based microphone system. In *Proceedings of 123th AES Convention*, Preprint #7313, New York, 2007.

39

Virtual Acoustics

Tapio Lokki and Lauri Savioja

Department of Media Technology,
Helsinki University of Technology, Helsinki, Finland,
e-mail: Tapio.Lokki@tkk.fi, Lauri.Savioja@tkk.fi

1 • Introduction

The term virtual acoustics is often applied when sound signal is processed to contain features of a simulated acoustical space and sound is spatially reproduced either with binaural or with multichannel techniques. Therefore, virtual acoustics consists of spatial sound reproduction and room acoustics modeling. A synonym for virtual acoustics is auralization–rendering audible–which was first defined by Kleiner et al. [1].

Virtual acoustics can be created with perceptually or physically based modeling techniques. The physically based approach seeks to simulate exactly the propagation of sound from the source to the listener for a given room. Conceptually, this approach is straightforward since by modeling physical principles of sound propagation in air and reflections from boundaries as correctly as possible a high-quality virtual acoustics should be achieved. However, for many applications, e.g., in computer music and professional audio, perceptually based approach produces relevant and accurate enough results. Rendering in perceptually based virtual acoustics is often performed without any knowledge of the room geometry, and perceptually optimized efficient algorithms are used [2]. In other words, perceptually based virtual acoustics endeavors to reproduce only the perceptually salient characteristics of reverberation (Section 75.7) [3].

Room acoustics modeling, required for physically based virtual acoustics, can be realized with several different ways. In principle, the most accurate results are obtained with wave-based techniques such as finite element method (FEM), boundary element method (BEM) [4], and finite-difference time domain (FDTD) methods [5–7]. They all try to numerically solve the wave equation in a modeled space. The more practical techniques are ray based, such as ray-tracing [8,9], beam-tracing [10], and the image source method [11,12]. None of the room acoustics modeling techniques is optimal in all geometries and over the whole audible frequency range, but each of them can be applied in the areas they are best at. A survey of room acoustics modeling has been presented by Svensson and Kristiansen [13].

In this chapter, signal processing methods applied in physically based virtual acoustics are presented. The presentation follows mainly the design of the DIVA auralization system [14,15] which has been developed at the Helsinki University of Technology since 1994. In addition to auralization, one signal processing-oriented room acoustics modeling method, namely the waveguide mesh, is presented in Section 3.

2 • Auralization and Signal Processing

In ray-based methods, the wavelength of sound is assumed to be zero and sound propagation is modeled with sound rays. With them reflection paths inside an enclosure can be efficiently found, but all frequency-dependent acoustical phenomena caused by the wave nature of sound are neglected. An energy response, needed to compute room acoustical parameters of a space, can be obtained with ray-based methods. Such parameters are often computed in octave bands. If path tracing is performed separately on each band, frequency-dependent attenuations can be simulated with gain factors. However, in auralization it is practical to perform room acoustics modeling with sound pressure signals. In addition, the whole audible frequency range is needed to be processed at once. These requirements are fulfilled if room acoustics modeling is performed with the image source method and the frequency-dependent phenomena are added by processing the sound emitted by each image source with digital filters. In other words, with image sources the propagation paths are found and with digital filters each impulse, corresponding to a ray path, is filtered to add wave nature features and correct reproduction direction. Figure 1 presents two reflection paths and the filters applied on both paths.

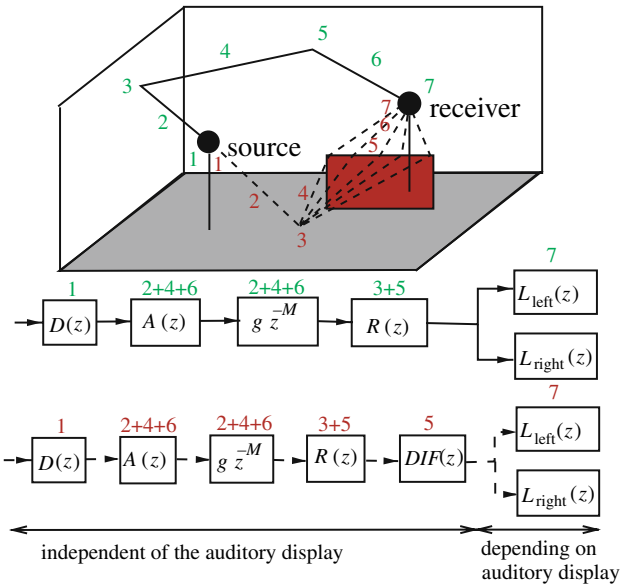


FIGURE 1 Two sound propagation paths with applied filters to perform physically based rendering. The filters simulate sound source directivity ($D(z)$), air absorption ($A(z)$), delay and gain based on distance (gz^{-M}), material absorption ($R(z)$), diffraction (DIF(z)), and binaural filtering ($L_{\text{left,right}}(z)$). Illustration adapted from Strauss [16].

The filtering approach in auralization enables interactive applications in which the listener or the sound source can move in a virtual space. In such dynamic rendering, the filter parameters are updated, e.g., 20 Hz update rate [17] while the sound processing is performed at 48 kHz frequency. The interactivity makes filter design challenging since as low-order filters as possible have to be used. Filter design is thus often optimized from a perception point of view. In addition, filters have to be designed so that interpolation of filter coefficients between updates is possible. In the following, different solutions to design and apply filters, presented in Figure 1, are discussed.

2.1 Sound Source Directivity

Sound sources have radiation characteristics that are dependent on frequency and direction of radiation. In auralization, a sound source is typically modeled as a point source and radiation to different directions is modeled with directional filtering [18–21]. Such simplification overrules the fact that in reality sound sources have a finite volume and they emit sound from several locations of their body. For example, most musical instruments

radiate sound from their whole body, and in addition the radiation characteristics vary according to the played note.

The modeling of sound source directivity is a problem similar to HRTF modeling (Section 11.1). For filter design, the directivity data of each sound source has to be measured or derived with mathematical models. The sparse measurement data is usually completed by interpolating measurements to form a dense grid of directivity data. Then, digital filters are fitted to the frequency response of each direction and these filters are applied in rendering. The filter fitting can be performed in a warped frequency domain [22,23] to optimize low-order filters for human auditory perception [15].

In auralization, one directional filter ($D(z)$ in Figure 1) has to be applied for every single reflection. One way to enhance the performance in real-time rendering is to separate the directivity filter into two parts, namely diffuse field part and directional part. Since the diffuse field filter is common to all directions, it has to be applied only once, while the directional filter is applied to every reflection. The diffuse field power spectrum for diffuse field filter can be derived from all measured responses ($H_{\mathrm{dir},i}(\omega, \theta, \phi)$) by power-averaging:

$$S_{\mathrm{d}}(\omega, \theta, \phi) = \sqrt{\frac{1}{k}\sum_{i=1}^{k} |H_{\mathrm{dir},i}(\omega, \theta, \phi)|^2 \cos\phi}, \tag{1}$$

where k is the number of responses, θ is the azimuth angle, ϕ is the elevation angle, and ω is the angular frequency.

The use of a diffuse field filter has many advantages. For example, it makes the design of the primary directivity filters much easier, by flattening the response at low and high frequencies. This is also called diffuse field equalization and it allows the use of computationally efficient low-order filters for the primary directivity filter.

2.2 Air Absorption

The absorption of sound in air depends mainly on distance, temperature, and humidity. The equations for calculation of air absorption are standardized [24], and they can be applied to calculate the target responses for filter design [14,25]. The filters ($A(z)$ in Figure 1) can be implemented as low-order IIR filters which are fitted, e.g., to minimum-phase complex frequency responses. In dynamic rendering, the coefficients of air absorption filters need not be interpolated, despite the recursive filters. Instead, filter coefficients can be changed without any audible effects if the grid of the precalculated filters is dense enough, e.g., in 1-m steps [15].

2.3 Distance Delay and Attenuation

The distance of an image source to the listener, i.e., the length of the ray path, defines the delay and attenuation of sound. The frequency independent attenuation according to $1/r$-law, caused by the spherical propagation of sound waves (see, e.g., [26]), is trivial to compute. In dynamic rendering, in which sound source or listener can move, the distance delay has to be implemented with fractional delays [27] to obtain smooth and continuous output signal. The first-order FIR fractional delay filter has been found accurate enough for this purpose. Correspondingly, interpolation of the distance attenuation coefficient between updates has to be applied [14].

2.4 Material Absorption

The effect of boundary material to a sound wave should be modeled somehow to obtain realistic auralization. The temporal and spectral behavior of a reflected sound is a function of incident angle. In addition, scattering and diffraction make it impossible to develop numerical models that are accurate in all aspects. For these reasons, in room acoustics modeling usually simplified material absorption models are applied. Typically, only angle-independent material absorption that can be measured in reverberation room or with the impedance tube technique is applied.

In auralization, the material absorption is modeled by utilizing reflection filters ($R(z)$ in Figure 1). The input data for filter design is measured absorption coefficients, available in octave bands. For filter design, absorption coefficients are transformed into reflectance data. The resulting amplitudes are transformed into the frequency domain, and filter fitting is performed in a warped frequency domain as in the case of sound source directivity filters. When sound is reflected from two or more surfaces, the absorption data can be cascaded so that only one filter is needed in rendering [25].

2.5 Diffraction

Modeling of sound propagation paths with rays totally neglects diffraction. However, Svensson et al. [28] have derived a mathematical solution with which edge diffraction can be added to the image source method. In fact, edge diffraction modeling compensates also the low-frequency error of the image source method. Such error is an inherent problem of the image source method due to the finite size of surfaces. In edge diffraction modeling, multiple rays are traced via all points of the edge and the sum of all these rays forms an impulse response for that edge which can be of a finite length. Although sound passes the edge through all points along the edge, most of the energy is concentrated to the shortest path point of the edge. Based on this, it can be assumed that diffraction can be modeled with a

point-like secondary source positioned to the shortest path point. So, diffraction can be approximated with one edge image source, filtered with a diffraction filter DIF(z) in Figure 1. For efficient auralization a low-order filter can be fitted to a computed edge impulse response [29].

2.6 Directional Filtering

Each modeled reflection reaches the listening position from a defined direction, and in auralization, the correct reproduction direction is obtained with directional filtering, filters $L_{\{\text{left,right}\}}(z)$ in Figure 1. The filtering required depends of the applied auditory display. For binaural reproduction HRTF filters [30] (Section 11.1) are applied, and for multichannel reproduction many different techniques are available (Chapter 38).

For dynamic auralization, HRTF filters have to be designed so that interpolation between them is possible. In the DIVA system [14], the HRTF filters were realized by decomposing the measured HRTF functions to a frequency-independent delay (ITD) and a minimum-phase part. Then, FIR filters were fitted to the minimum-phase responses. Such minimum-phase FIR filtering enables direct interpolation of filter coefficients in a dynamic auralization.

2.7 Late Reverberation

In real-time auralization, only early reflections can be modeled with image sources and digital filtering. As reflection order grows, more and more reflections reach the listening position and reflections build a nearly diffuse reverberation field which is close to exponentially decaying random noise [31]. Under these assumptions, the late reverberation does not have to be modeled as individual reflections with certain wave features and directions. Therefore, to optimize computation in late reverberation modeling, recursive digital filter structures have been designed [3,32], whose responses model the characteristics of real room responses, such as the frequency-dependent reverberation time (Section 75.7).

3 • Waveguide Meshes in Room Acoustics

The digital waveguide mesh provides a tool for simulating wave propagation in a multidimensional system. The most practical applications lie in physical modeling of musical instruments and room acoustics.

A digital waveguide mesh [33] is formed of a regular array of discretely spaced 1D digital waveguides arranged along each perpendicular dimension, interconnected at their intersections

(Chapter 25). The resulting mesh of a 3D space is a regular rectangular grid in which each node is connected to its six neighbors by unit delays [33,34].

Two conditions must be satisfied at a lossless junction connecting lines of equal impedance [33]:

1. Sum of inputs equals the sum of outputs (volume velocity flows add to zero).
2. Sound pressures at each intersecting waveguide are equal at the junction (continuity of pressure).

Based on these conditions, a difference equation for the nodes of an N-dimensional rectangular mesh [34] can be derived:

$$p_k(n) = \frac{1}{N}\left[\sum_{l=1}^{2N} p_l(n-1)\right] - p_k(n-2), \tag{2}$$

where p represents the sound pressure at a junction at time step n, k is the position of the junction to be calculated, and l represents all the neighbors of k. This equation is equivalent to a difference equation derived from the Helmholtz equation by discretizing time and space.

3.1 Dispersion Error

An inherent problem with the rectangular digital waveguide mesh method is the direction- and frequency-dependent dispersion of wavefronts [33]. High-frequency signals parallel to the coordinate axes are delayed, whereas diagonally the waves propagate undistorted.

The dispersion error of digital waveguide mesh structures can be investigated with Von Neumann analysis [33,35]. It is based on the two-dimensional discrete-time Fourier transform of the difference scheme. After the Fourier transform, the scheme can be represented by means of the spectral amplification factor, and the stability and the dispersion characteristics of the scheme may be analyzed with it.

3.2 Mesh Topologies

The two-dimensional rectangular mesh is the most simple mesh structure as shown in Figure 2a. Several mesh topologies have been experimented to reduce the direction-dependent dispersion in that structure, but all of them suffer from some kind of dispersion. The two main approaches are triangular [36,37] and interpolated rectangular ones [38] as illustrated in Figure 2b and c, respectively. The 3D counterpart for the triangular mesh is the tetrahedral structure [37], while the interpolated mesh can be applied in both 2D and 3D [39].

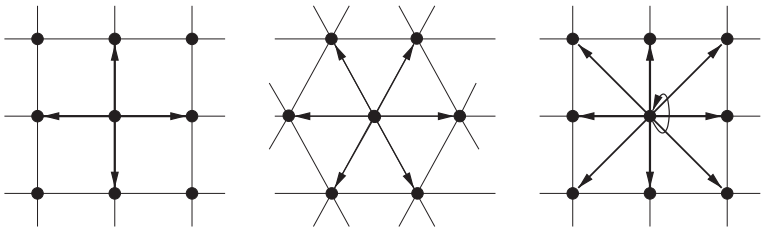


FIGURE 2 Various 2D digital waveguide mesh structures: (**a**) in the rectangular mesh each node is connected to four neighbors with unit delays, (**b**) triangular mesh has six connections per neighbor, and (**c**) in the interpolated mesh the signal propagates to all the eight neighbors and to the node itself.

3.3 Reducing Dispersion with Frequency Warping

The triangular and interpolated rectangular mesh structures render the dispersion practically independent of propagation direction. In addition, the frequency error curves are smooth. Due to these reasons, the dispersion error can be reduced remarkably by applying a frequency warping technique [38,39].

The frequency warping is applied to the input and output signals of the mesh using a warped FIR filter [22,23]. This filter shifts frequencies to compensate for the dispersion error. For this purpose a first-order allpass warping is suitable. The applied amount of warping determines the extent of warping, and there are various criteria to find the optimal value to be used with digital waveguide meshes [38].

3.4 Boundary Conditions

The boundary conditions of a mesh dictate the reflection characteristics of the surfaces surrounding the space under study, and thus are fundamental for acoustic simulation. In general, the boundary conditions can be represented by digital filters [25]. However, in practice setting the boundary conditions accurately are difficult, and the topic is still under research at the time of writing [40]. In general, it is required that users can set arbitrary boundaries for modeling reflection of sound from various materials. Challenging topic, especially, is the modeling of an anechoic boundary that enables simulation of semi-open spaces.

References

1. M. Kleiner, B.-I. Dalenbäck, and U.P. Svensson. Auralization – an overview. *J. Audio Eng. Soc.*, 41(11):861–875, Nov. 1993.
2. J.-M. Jot. Real-time spatial processing of sounds for music, multimedia and interactive human–computer interfaces. *Multimedia Syst., Special Issue Audio Multimedia*, 7(1):55–69, 1999.

3. W.G. Gardner. Reverberation algorithms. In M. Kahrs and K. Brandenburg, editors, *Applications of Digital Signal Processing to Audio and Acoustics*, pp. 85–131. Kluwer Academic Publishers, Norwell, MA, 1997.
4. A. Pietrzyk. Computer modeling of the sound field in small rooms. *Proceedings of the AES 15th International Conference on Audio, Acoustics & Small Spaces*, pp. 24–31, Copenhagen, Denmark, Oct. 31–Nov. 2, 1998.
5. D. Botteldooren. Finite-difference time-domain simulation of low-frequency room acoustic problems. *J. Acoust. Soc. Am.*, 98(6): 3302–3308, 1995.
6. L. Savioja, J. Backman, A. Järvinen, and T. Takala. Waveguide mesh method for low-frequency simulation of room acoustics. *Proceedings of the 15th International Congress Acoustics on (ICA '95)*, volume 2, pp. 637–640, Trondheim, Norway, June 1995.
7. L. Savioja. *Modeling Techniques for Virtual Acoustics*. PhD thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, report TML-A3, 1999.
8. A. Krokstad, S. Strom, and S. Sorsdal. Calculating the acoustical room response by the use of a ray tracing technique. *J. Sound Vib.*, 8(1):118–125, 1968.
9. A. Kulowski. Algorithmic representation of the ray tracing technique. *Appl. Acoust.*, 18(6):449–469, 1985.
10. T.A. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. West, G. Pingali, P. Min, and A. Ngan. A beam tracing method for interactive architectural acoustics. *J. Acoust. Soc. Am.*, 115(2): 739–756, February 2004.
11. J.B. Allen and D.A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.*, 65(4):943–950, 1979.
12. J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. Soc. Am.*, 75(6):1827–1836, 1984.
13. U.P. Svensson and U.R. Kristiansen. Computational modeling and simulation of acoustic spaces. *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, pp. 11–30, Espoo, Finland, June 15–17, 2002.
14. L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen. Creating interactive virtual acoustic environments. *J. Audio Eng. Soc.*, 47(9):675–705, 1999.
15. T. Lokki. *Physically-Based Auralization – Design, Implementation, and Evaluation*. PhD thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, report TML-A5, 2002. Available at http://lib.hut.fi/Diss/2002/isbn9512261588/.
16. H. Strauss. Implementing doppler shifts for virtual auditory environments. *The 104th Audio Engineering Society (AES) Convention*, Amsterdam, the Netherlands, May 16–19, 1998. Preprint no. 4687.
17. T. Lokki, L. Savioja, J. Huopaniemi, R. Väänänen, and T. Takala. Creating interactive virtual auditory environments. *IEEE Comput. Graph. Appl.*, 22(4):49–57, July/Aug. 2002.

18. M. Kompis and N. Dillier. Simulating transfer functions in a reverberant room including source directivity and head-shadow effects. *J. Acoust. Soc. Am.*, 93(5):2779–2787, May 1993.
19. J. Huopaniemi, K. Kettunen, and J. Rahkonen. Measurement and modeling techniques for directional sound radiation from the mouth. *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '99)*, pp. 183–186, Mohonk Mountain House, New Paltz, New York, Oct. 17–20, 1999.
20. J.-M. Jot, V. Larcher, and O. Warusfel. Digital signal processing issues in the context of binaural and transaural stereophony. *The 98th Audio Engineering Society (AES) Convention*, Paris, France, 1995. Preprint no. 3980.
21. M. Karjalainen, J. Huopaniemi, and V. Välimäki. Direction-dependent physical modeling of musical instruments. *Proceedings of the 15th International Congress on Acoustics (ICA '95)*, pp. 451–454, Trondheim, Norway, June 1995.
22. U.K. Laine, M. Karjalainen, and T. Altosaar. Warped linear prediction (WLP) in speech and audio processing. *Proceedings of the International Conference on Acoustics, Speech, Signal Processing (ICASSP'94)*, volume 3, pp. 349–352, Adelaide, Australia, April 19–22, 1994.
23. A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U.K. Laine, and J. Huopaniemi. Frequency-warped signal processing for audio applications. *J. Audio Eng. Soc.*, 48(11):1011–1031, Nov. 2000.
24. ISO Standard 9613-1. *Acoustics – Attenuation of Sound During Propagation Outdoors – Part 1: Calculation of the Absorption of Sound by the Atmosphere*. 1993.
25. J. Huopaniemi, L. Savioja, and M. Karjalainen. Modeling of reflections and air absorption in acoustical spaces – a digital filter design approach. *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'97)*, Mohonk, New Paltz, New York, Oct. 19–22, 1997.
26. A.D. Pierce. *Acoustics – An Introduction to Its Physical Principles and Applications*. The Acoustical Society of America, 2nd ed., 1994.
27. T.I. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine. Splitting the unit delay – tools for fractional delay filter design. *IEEE Signal Process. Mag.*, 13(1):30–60, Jan. 1996.
28. U.P. Svensson, R.I. Fred, and J. Vanderkooy. Analytic secondary source model of edge diffraction impulse responses. *J. Acoust. Soc. Am.*, 106(5):2331–2344, 1999.
29. T. Lokki, U.P. Svensson, and L. Savioja. An efficient auralization of edge diffraction. *AES 21st International Conference on Architectural Acoustics and Sound Reinforcement*, pp. 166–172, St. Petersburg, Russia, June 1–3, 2002.
30. D.R. Begault *3-D Sound for Virtual Reality and Multimedia*. AP Professional, Cambridge, MA, 1994.
31. M.R. Schroeder. Natural-sounding artificial reverberation. *J. Audio Eng. Soc.*, 10(3):219–223, 1962.

32. J.-M. Jot. *Etude et réalisation d'un spatialisateur de sons par modéles physique et perceptifs*. PhD thesis, l'Ecole Nationale Superieure des Telecommunications, Télécom Paris 92 E 019, Sept. 1992.
33. S. Van Duyne and J.O. Smith. Physical modeling with the 2-D digital waveguide mesh. *Proceedings of the International Computer Music Conference*, pp. 40–47, Tokyo, Japan, Sept. 1993.
34. L. Savioja, T. Rinne, and T. Takala. Simulation of room acoustics with a 3-D finite difference mesh. *Proceedings of the International Computer Music Conference*, pp. 463–466, Aarhus, Denmark, Sept. 1994.
35. J. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Chapman & Hall, New York, NY, 1989.
36. F. Fontana and D. Rocchesso. A new formulation of the 2D-waveguide mesh for percussion instruments. *Proceedings of the XI Colloquium on Musical Informatics*, pp. 27–30, Bologna, Italy, Nov. 1995.
37. S. Van Duyne and J.O. Smith. The 3D tetrahedral digital waveguide mesh with musical applications. *Proceedings of the International Computer Music Conference*, pp. 9–16, Hong Kong, Aug. 1996.
38. L. Savioja and V. Välimäki. Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques. *IEEE Trans. Speech Audio Process.*, 8(2):184–194, March 2000.
39. L. Savioja and V. Välimäki. Interpolated rectangular 3-D digital waveguide mesh algorithms with frequency warping. *IEEE Trans. Speech Audio Process.*, 11(6):783–790, Nov. 2003.
40. A. Kelloniemi, L. Savioja, and V. Välimäki. Spatial filter-based absorbing boundary for the 2-d digital waveguide mesh. *IEEE Signal Process. Lett.*, 12(2):126–129, 2005.

40

Audio Restoration

Paulo A.A. Esquef

Nokia Institute of Technology, Rod. Torquato Tapajós, 7200, Tarumã 69048-660 Manaus-AM, Brazil, e-mail: paulo.esquef@indt.org.br

1 • A Brief History of the Recording Technology

The first reproducible recording of human voice was made in 1877 on a tinfoil cylinder phonograph devised by Thomas A. Edison. Since then, much effort has been expended to find better ways to record and reproduce sounds. By the mid-1920s, the first electrical recordings appeared and gradually took over purely acoustic recordings. The development of electronic computers, in conjunction with the ability to record data onto magnetic or optical media, culminated in the standardization of compact disc format in 1980. Nowadays, digital technology is applied to several audio applications, not only to improve the quality of modern and old recording/reproduction techniques, but also to trade off sound quality for less storage space and less taxing transmission capacity requirements. For a comprehensive timeline and description of the most prominent events regarding the recording technology history see [1,2].

2 • Aims and Problems

Audio restoration basically aims at improving the sound of old recordings. The primary goal is to reduce spurious artifacts, which are usually introduced by the recording/playback mechanisms, while preserving, as much as possible, the original recorded sound.

The first step in a typical audio restoration procedure consists of transferring the sound from old matrices to a digital form. This task involves practical matters such as locating the original matrices or the best sounding copies; finding the best way or equipment to play back a given matrix; and dealing with the usual lack of standardization associated with obsolescent recording/playback systems.

Sound transferring from old media to more modern ones, for instance from 78 RPM to LP disks,[1] and attempts to improve the sonic quality of old recordings within the sound transferring process were a common practice even before the digital audio era. Thanks to the progressive increases in the computational power of digital processors, more sophisticated and powerful processing of digitized data has become feasible in practice. Nowadays, audio restoration is accomplished via DSP algorithms devoted to reducing or suppressing from the recorded material spurious noises and distortions introduced by the old recording/playback systems.

The most commonly found types of degradation associated with old recordings can be roughly classified into localized and global disturbances [3]. For example, short impulsive noises (clicks, crackles, and pops) as well as low-frequency long pulses (thumps) belong to the former class. Continuous background disturbances or interferences, such as broadband noise (hiss), buzz, and hum, are usually classified as global disturbances. Other examples of audio degradations include slow and fast frequency modulations, which are known as wow and flutter, respectively; surface noise; and non-linear distortions, e.g., amplitude clipping.

3 • Overview of Digital Audio Restoration

Digital audio restoration takes place after the sound transfer process from analog sources to digital domain. A review of the main techniques available for treating disturbances, such as thumps, clicks, and hiss, is given in the ensuing sections. The presentation follows the order of precedence in which the problems are usually tackled. Methods and algorithms devised for dealing with other types of degradations can be found elsewhere [3].

---

[1]The abbreviations RPM and LP stand for revolutions per minute and long-playing, respectively.

3.1 De-thumping

Thumps are produced by long pulses of low-frequency content that additively corrupt an audio signal. These pulses are related to the mechanical response of the playback mechanisms to an abnormal excitation, e.g., the response of the stylus-arm apparatus to large discontinuities on the groove walls of a disk.

3.1.1 Template Matching

Apart from crude techniques, such as highpass filtering of the signal, a few methods, with various degrees of sophistication, exist for treating long pulses. The template-matching method introduced by [4] figures among one of the first propositions for treating long pulses. Figure 1 shows a block diagram that illustrates the functional stages of the template-matching method for de-thumping. The basic assumption behind this method is that long pulses are identical in shape, varying only in amplitude.

In practical terms, a corrupting pulse is considered to be a scaled version of a representative template of the pulse $\mathbf{d}$ added to the clean signal $\mathbf{x}_n$, leading to $\mathbf{y}_n = \mathbf{x}_n + \mu\mathbf{d}$, with $\mathbf{y}_n = [y_n, y_{n+1}, \ldots, y_{n+N-1}]^{\mathrm{T}}$, $\mathbf{x}_n = [x_n, x_{n+1}, \ldots, x_{n+N-1}]^{\mathrm{T}}$, $\mathbf{d}=[d_0, d_1, \ldots, d_{N-1}]^{\mathrm{T}}$, and μ a scale factor. Moreover, the matched filter, whose coefficients $\mathbf{h}_{\mathrm{mf}}$ consist of the time-reversed samples of $\mathbf{d}$ appropriately scaled, works as an estimator of the cross-correlation coefficient between $\mathbf{d}$ and $\mathbf{y}_n$. Therefore, the presence of pulses is indicated by values of the filter's output that are near unity in magnitude. Thus, after synchronizing the template with a given pulse occurrence, the estimated de-thumped signal is obtained by $\hat{\mathbf{x}}_n = \mathbf{y}_n - \mu\mathbf{d}$, where the scaling factor can be shown to be given by $\mu = (\mathbf{h}_{\mathrm{mf}}^{\mathrm{T}}\mathbf{y}_{\mathbf{n}})/(\mathbf{h}_{\mathrm{mf}}^{\mathrm{T}}\mathbf{h}_{\mathrm{mf}})$.

The high-amplitude click that usually appears at the beginning of a long pulse is to be removed by a de-clicking algorithm (to be discussed in Sect. 3.2). The main limitation of the template-matching method is the lack of robustness in detecting and suppressing pulses whose shape varies over time or pulses that are superimposed [3].

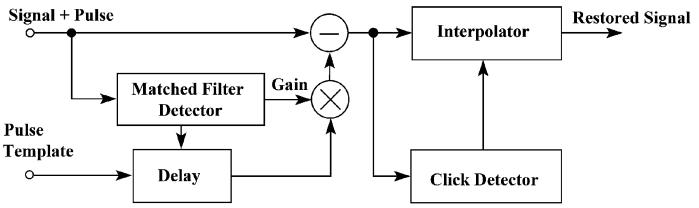


FIGURE 1 Processing stages of the template-matching method for de-thumping. Adapted from [4].

3.1.2 Other Methods

A model-based approach to long pulse removal has been proposed in [5] and further developed in [6]. In this method the corrupted audio signal is modeled as a mixture of two distinct AR processes. Thus, signal restoration is accomplished by separating these two processes. The AR-separation method can produce excellent signal restorations even for superimposed pulses. Moreover, it incorporates the suppression of the initial pulse clicks within the de-thumping procedure. The main drawback of the AR-separation method is its high computational cost.

An alternative audio de-thumping procedure, which is based upon non-linear [7] and polynomial smoothing filters [8], has been proposed in [9]. It is comparable to the AR-separation method in restoration performance, despite being less demanding computationally.

3.2 De-clicking

Impulsive-like disturbances contaminating audio signals are heard as clicks. Such disturbances are usually caused by the intrinsic porosity of the disk material, solid particles adhering to the groove walls, and the presence of superficial scratches on the disk surface [10].

3.2.1 Click Detection

Techniques based on autoregressive (AR) models have been demonstrated to be successful for de-clicking purposes. In these methods the degradation mechanism is assumed to be additive, with the corrupted signal given by $y_n = x_n + i_n d_n$, where x_n is the clean signal and the term $i_n d_n$ corresponds to the noise component. The sequence i_n is binary and indicates the presence of noise at instants n where $i_n = 1$.

The goal of click detection algorithms is to estimate the sequence i_n that minimizes the percentage of both miss detection and false alarm of clicks. Figure 2 depicts a block diagram with the functional stages of a typical AR-based de-clicking scheme.

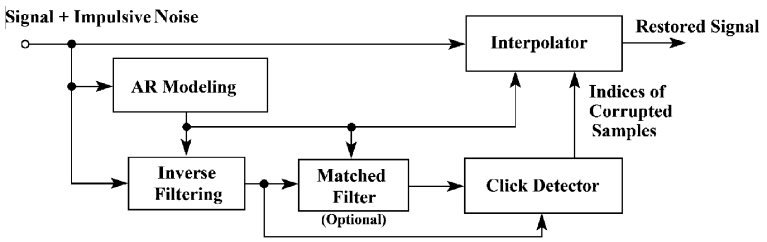


FIGURE 2 Model-based de-clicking method. Adapted from [4].

Given a frame of the corrupted signal $\mathbf{y}_l = [y_l,\ y_{l+1}, \ldots,\ y_{l+N-1}]^{\mathrm{T}}$ and a pth-order estimate $\hat{\mathbf{a}}$ of its AR coefficient vector,[2] the associated sequence i_n can be straightforwardly obtained via

$$i_n = \begin{cases} 1, & |e_n| \geq T \\ 0, & \text{otherwise} \end{cases}, \text{ for } \quad n = l+p, l+p+1, \ldots, l+N-1, \quad (1)$$

where $e_n = \hat{\mathbf{a}}^{\mathrm{T}}\mathbf{y}'$, with $\mathbf{y}' = [y_n,\ y_{n-1}, \ldots,\ y_{n-p}]$, is the output of the inverse filter and $T = g\,\text{median}([|e_{l+p}|,\ |e_{l+p+1}|, \ldots,\ |e_{l+N-1}|])$ is the value of a variable threshold. The parameter g is a scaling factor. Considering a sample rate of 44.1 kHz, the values of the processing parameters are typically set within the following ranges: $500 \leq N \leq 2,000$, $10 \leq p \leq 40$, and $5 \leq g \leq 7$.

Further improvements in click detection performance can be achieved by using matched filtering [4], two-sided linear prediction [11], and more elaborated threshold-based schemes [12].

3.2.2 Signal Reconstruction

Click suppression is attained by replacing the corrupted samples with others that resemble the underlying signal. Since, in general, audio signals exhibit resonant characteristics, AR-based interpolation methods constitute a suitable tool for recovery of missing samples in audio signals. In AR-based interpolators, signal reconstruction is usually attained through the minimization of the variance of the model error w.r.t. the unknown samples [13]. However, due to constraints related to the non-stationarity of audio signals, meaningful data reconstruction is restricted to relatively short-duration (up to about 5 ms) portions of audio material.

Interpolation across longer gaps tends to produce reconstructed signals that are smoother and weaker in energy than the surroundings of the gap. The simplest way to circumvent these shortcomings is to increase the order of the AR model. Other solutions include the use of: extended AR models [11]; interpolators combining AR models and with cosine basis functions [3]; interpolators with constant excitation energy [14,15]; interpolators based on random sampling [16,17]; and interpolators that employ two different AR models, one to model the segment that immediately precedes a gap and another for the fragment that succeeds the gap [18,19].

---

[2] Any standard AR estimation algorithm, such as Yule–Walker and Burg's methods, can be used to obtain this estimate.

3.2.3 Other Methods

Various kinds of approaches have been attempted to tackle the reconstruction of discrete-time signals across gaps of missing samples [20]. In addition to AR-based techniques, audio signal recovery can be attained via: band-limitation criterion [21]; sinusoidal modeling interpolation [22]; waveform substitution [23,24]; multirate signal decomposition and synthesis [25,26]; Bayesian methods [3,5,27]; neural-network algorithms [28]; and model-based adaptive filtering techniques [29,30].

3.3 De-hissing

Signal contamination with additive broadband noise is perceived as hiss. The noise component is usually due to thermic measurement noise, circuitry noise, and intrinsic characteristics of the storage medium, such as in analog tape recordings.

3.3.1 Short-Time Spectral Attenuation Methods

Short-time spectral attenuation (STSA) [31,32] is one of the most common means of audio and speech de-hissing. The basic assumption behind STSA methods is that the samples of the corrupted signal are given by $y_n = x_n + d_n$, where d_n is an additive zero-mean white Gaussian process, uncorrelated to the underlying clean signal x_n.

Due to the non-stationarity of general audio signals, STSA methods employ a block-based overlap-and-add processing scheme [33]. Moreover, the core processing is performed in the frequency domain. In that regard, the transformations from time into frequency and vice versa are carried out through the discrete Fourier transform (DFT). Figure 3 illustrates the typical processing stages of STSA de-hissing methods.

In STSA methods, the frequency components present in a given signal frame are attenuated in magnitude based on a suppression rule, which takes into account the signal-to-noise ratio (SNR) estimated at each frequency bin. For that, estimates of power spectral density of both the disturbance and the corrupted signal

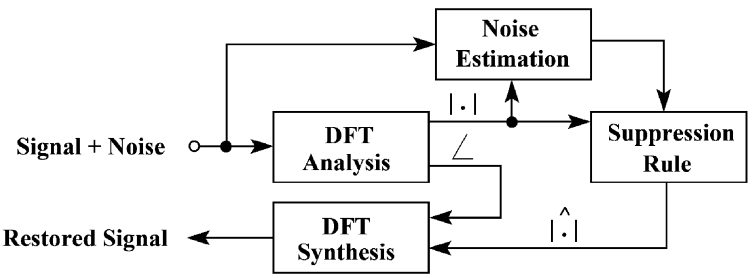


FIGURE 3 STSA-based hiss reduction system. Adapted from [34].

are needed. In practice, those densities can be approximated by the power spectra $|D_k|^2$ and $|Y_k|^2$ of the respective signals, computed via the DFT.

The most common suppression rules can be formulated as

$$|\hat{X}_k| = H_k |Y_k|, \tag{2}$$

where H_k is a frame-varying frequency-domain filter and $|\hat{X}_k|$ is the magnitude spectrum of the estimated clean signal, with k being the DFT bin index. Typically, the filter assumes the form $H_k = [1 - (|D_k|/|Y_k|)^b]^{1/b}$. For example, the so-called magnitude and power spectrum subtraction rules [32] are attained by choosing $b = 1$ and $b = 2$, respectively. The Wiener filter rule is simply given by H_k^2 for $b = 2$. For comparisons among different suppression rules see [33,35].

In general, noise reduction via STSA methods comes at the cost of losses of valuable signal content, especially on the high-frequency range. Another typical side effect affecting the processed signals is the presence of a phenomenon known as musical noise [35]. There exist several means to reduce this artifact. The simplest way consists in overestimating the power of the noise spectrum. Alternatively, musical noise can be reduced by: averaging the spectral information over successive signal frames [32]; applying heuristic rules to the spectral attenuation factor [4,32]; and leaving a minimum noise floor to mask the musical noise components [3,36,37].

The suppression rule proposed by Ephraim and Malah (EMSR) [38,39] is known to be one of the most effective in reducing musical noise. Simpler alternatives to the EMSR have also been proposed in [34]. Recently, some interest has been aroused in incorporating psychoacoustic-related phenomena into the suppression rules used in STSA methods [40–44].

3.3.2 Other Methods

Within the class of non-parametric de-hissing techniques, wavelet-based shrinkage has received substantial attention in the pertinent literature [45–48]. In this method the noisy signal is mapped into a multi-resolution time-frequency representation via a discrete wavelet transform. In the wavelet domain, the underlying signal is supposed to be compactly represented by a few high-amplitude coefficients. Conversely, the noise process is spread over a large number of low-level coefficients. Thus, de-noising can be accomplished, in the simplest way, by discarding these coefficients (hard threshold) before transforming the signal back to the original domain. More involved wavelet shrinkage methods exist [49].

As for parametric audio de-hissing, techniques that make use of AR modeling of degraded speech and audio signals have been proposed in [50]. Statistical-based methods that employ AR and ARMA modeling have also been developed for joint treatment of broadband and impulsive noise [3,29]. Adaptive filtering schemes for audio de-hissing are presented in [4,29,30]. Hybrid methods that combine either different non-parametric schemes [51] or mix non-parametric and parametric strategies [52] can also be found in the relevant literature.

4 • Future Directions

Currently available DSP tools for audio restoration offer a certain degree of automation for some de-noising tasks. However, most restoration methods do not integrate a priori knowledge on the contents of the signal to be restored. Therefore, the sonic quality of a restored sound strongly depends on how an expert listener chooses the processing parameters as well as personally judges the attained results.

It seems reasonable to believe that, in the future, audio restoration techniques can benefit from the incorporation of high-level information on the contents of the signals under processing. However, as pointed out in [53], such strategy implies working in an object-based audio representation [54], which is less general than processing approaches based on conventional signal modeling techniques. Nevertheless, a priori knowledge on the contents of the audio material at hand, for example, genre of the program and types of generative sound sources, could be used to guide the choice of the most suitable restoration technique to employ as well as to better tune its processing parameters. A far more audacious and involved step would consist of first extracting high-level content information from old recordings and then realizing restoration via sound resynthesis within an object-based audio representation framework.

References

1. D. L. Morton, *Sound Recording: The Life Story of a Technology*, Greenwood Technologies, Greenwood Press, Oxford, UK Sept. 2004.
2. S. Schoenherr, "Recording Technology History," `http://history.acusd.edu/gen/recording/notes.html`, Retrieved Dec. 2003.
3. S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration – A Statistical Model Based Approach*, Springer, London, 1998, `http://www-sigproc.eng.cam.ac.uk/~sjg/springer/`.
4. S. V. Vaseghi and R. Frayling-Cork, "Restoration of Old Gramophone Recordings," *J. Audio Eng. Soc.*, vol. 40, no. 10, pp. 791–801, Oct. 1992.

5. S. J. Godsill, *The Restoration of Degraded Audio Signals*, Ph.D. thesis, Engineering Department, Cambridge University, Cambridge, England, Dec. 1993, `http://www-sigproc.eng.cam.ac.uk/~sjg/thesis/`.
6. S. J. Godsill and C. H. Tan, "Removal of Low Frequency Transient Noise from Old Recordings Using Model-Based Signal Separation Techniques," *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, Oct. 1997.
7. W. A. Struzinski and E. D. Lowe, "A Performance Comparison of Four Noise Background Normalization Schemes Proposed for Signal Detection Systems," *J. Acoust. Soc. Am.*, vol. 76, no. 6, pp. 1738–1742, Dec. 1984.
8. W. H. Press and S. A. Teukolsky, "Savitzky-Golay Smoothing Filters," *Comput. Phys.*, vol. 4, no. 6, pp. 669–672, 1990.
9. P. A. A. Esquef, L. W. P. Biscainho, and V. Välimäki, "An Efficient Algorithm for the Restoration of Audio Signals Corrupted with Low-Frequency Pulses," *J. Audio Eng. Soc.*, vol. 51, no. 6, pp. 502–517, June 2003.
10. P. Wilson, "Record Contamination: Causes and Cure," *J. Audio Eng. Soc.*, vol. 13, no. 2, pp. 166–176, Apr. 1965.
11. S. V. Vaseghi and P. J. W. Rayner, "Detection and Suppression of Impulsive Noise in Speech Communication Systems," *IEE Proc.*, vol. 137, no. 1, pp. 38–46, Feb. 1990.
12. P. A. A. Esquef, L. W. P. Biscainho, P. S. R. Diniz, and F. P. Freeland, "A Double-Threshold-Based Approach to Impulsive Noise Detection in Audio Signals," *Proc. X European Signal Processing Conf.*, Tampere, Finland, Sept. 2000, vol. 4, pp. 2041–2044.
13. A. J. E. M. Janssen, R. N. J. Veldhuis, and L. B. Vries, "Adaptive Interpolation of Discrete-Time Signals that Can Be Modeled as Autoregressive Processes," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-34, no. 2, pp. 317–330, Apr. 1986.
14. P. J. W. Rayner and S. J. Godsill, "The Detection and Correction of Artefacts in Degraded Gramophone Recordings," *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct. 1991, pp. 151–152.
15. M. Niedźwiecki, "Statistical Reconstruction of Multivariate Time Series," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 451–457, Jan. 1993.
16. J. J. Ó Ruanaidh and W. J. Fitzgerald, "Interpolation of Missing Samples for Audio Restoration," *Electron. Lett.*, vol. 30, no. 8, pp. 622–623, Apr. 1994.
17. S. J. Godsill and P. J. W. Rayner, "Statistical Reconstruction and Analysis of Autoregressive Signals in Impulsive Noise Using the Gibbs Sampler," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 4, pp. 352–372, July 1998.
18. W. Etter, "Restoration of a Discrete-Time Signal Segment by Interpolation Based on the Left-Sided and Right-Sided Autoregressive Parameters," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1124–1135, May 1996.

19. I. Kauppinen, J. Kauppinen, and P. Saarinen, "A Method for Long Extrapolation of Audio Signals," *J. Audio Eng. Soc*., vol. 49, no. 12, pp. 1167–1180, Dec. 2001.
20. R. Veldhuis, *Restoration of Lost Samples in Digital Signals*, Prentice-Hall, New Jersey, 1990.
21. P. J. S. G. Ferreira, "Iterative and Noniterative Recovery of Missing Samples for 1-D Band-Limited Signals," in *Nonuniform Sampling – Theory and Practice*, F. Marvasti, Ed., Chap. 5, pp. 235–281. Kluwer Academic/Plenum, New York, 2001.
22. R. C. Maher, "A Method for Extrapolation of Missing Digital Audio Data," *J. Audio Eng. Soc*., vol. 42, no. 5, pp. 350–357, May 1994.
23. D. J. Goodman, G. B. Lockhart, O. J. Waen, and W. C. Wong, "Waveform Substitution Techniques for Recovering Missing Speech Segments in Packet Voice Communications," *IEEE Trans. Acoust. Speech Signal Process*., vol. ASSP-34, no. 6, pp. 1440–1448, Dec. 1986.
24. M. Niedźwiecki and K. Cisowski, "Smart Copying – A New Approach to Reconstruction of Audio Signals," *IEEE Trans. Signal Process*., vol. 49, no. 10, pp. 2272–2282, Oct. 2001.
25. S. Montresor, J. C. Valiere, J. F. Allard, and M. Baudry, "Evaluation of Two Interpolation Methods Applied to Old Recordings Restoration," Presented at the 90th Convention of the AES, Paris, France, Feb. 1991, Preprint 3022.
26. B.-S. Chen and Y.-L. Chen, "Multirate Modeling of AR/ARMA Stocastic Signals and Its Application to the Combined Estimation-Interpolation Problem," *IEEE Trans. Signal Process*., vol. 43, no. 10, pp. 2302–2312, Oct. 1995.
27. S. J. Godsill and P. J. W. Rayner, "A Bayesian Approach to the Restoration of Degraded Audio Signals," *IEEE Trans. Speech Audio Process*., vol. 3, no. 4, pp. 276–278, July 1995.
28. A. Czyzewski, "Learning Algorithms for Audio Signal Enhancement, Part I: Neural Network Implementation for the Removal of Impulsive Distortions," *J. Audio Eng. Soc*., vol. 45, no. 10, pp. 815–831, Oct. 1997.
29. M. Niedźwiecki and K. Cisowski, "Adaptive Scheme for Elimination of Broadband Noise and Impulsive Disturbances from AR and ARMA Signals," *IEEE Trans. Signal Process*., vol. 44, no. 3, pp. 528–537, Mar. 1996.
30. A. Bari, S. Canazza, G. De Poli, and G. A. Mian, "Improving the Extended Kalman Filter Method for the Restoration of Electro-acoustic Music," *Proc. Int. Computer Music Conf*., Berlin, Germany, Aug. 2000, pp. 233–236.
31. J. Lim, "Evaluation of a Correlation Subtraction Method for Enhancing Speech Degraded by Additive White Noise," *IEEE Trans. Acoust. Speech Signal Process*., vol. 26, no. 5, pp. 471–472, Oct. 1978.
32. S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," *IEEE Trans. Acoust. Speech Signal Process*., vol. ASSP-27, no. 2, pp. 113–120, Apr. 1979.

33. S. J. Godsill, P. J. W. Rayner, and O. Cappé, "Digital Audio Restoration," in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds., Chap. 4, pp. 133–194. Kluwer Academic, Boston, USA, 1998.
34. P. J. Wolfe and S. J. Godsill, "Efficient Alternatives to the Ephraim and Malah Suppression Rule for Audio Signal Enhancement," *EURASIP J. Appl. Signal Process. – Special Issue on Digital Audio for Multimedia Communications*, vol. 2003, no. 10, pp. 1043–1051, Sept. 2003.
35. O. Cappé and J. Laroche, "Evaluation of Short-Time Spectral Attenuation Techniques for the Restoration of Musical Recordings," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 1, pp. 84–93, Jan. 1995.
36. M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Apr. 1979, vol. 4, pp. 208–211.
37. J. A. Moorer and M. Berger, "Linear-Phase Bandsplitting: Theory and Applications," *J. Audio Eng. Soc.*, vol. 34, no. 3, pp. 143–152, Mar. 1986.
38. Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 32, no. 6, pp. 1109–1121, Dec. 1984.
39. O. Cappé, "Elimination of the Musical Noise Phenomenon with the Ephraim and Malah Noise Suppressor," *IEEE Trans. Speech Audio Process.*, vol. 2, no. 2, pp. 345–349, Apr. 1994.
40. D. Tsoukalas, J. Mourjopoulos, and G. Kokkinakis, "Perceptual Filters for Audio Signal Enhancement," *J. Audio Eng. Soc.*, vol. 45, no. 1/2, pp. 22–36, Jan./Feb. 1997.
41. M. Lorber and R. Hoeldrich, "A Combined Approach for Broadband Noise Reduction," *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct. 1997.
42. P. J. Wolfe and S. J. Godsill, "Perceptually Motivated Approaches to Music Restoration," *J. New Music Res. – Special Issue: Music and Mathematics*, vol. 30, no. 1, pp. 83–92, Mar. 2001.
43. A. Bari, S. Canazza, G. De Poli, and G. A. Mian, "Toward a Methodology for the Restoration of Electroacoustic Music," *J. New Music Res. – Special Issue: Preserving Electroacoustic Music*, vol. 30, no. 4, Dec. 2001.
44. I. Kauppinen and K. Roth, "Adaptive Psychoacoustic Filter for Broadband Noise Reduction in Audio Signals," *Proc. 14th Int. Conf. Digital Signal Processing*, Santorini, Greece, July 2002, vol. 2, pp. 962–966.
45. J. Berger, R. R. Coifman, and M. J. Goldberg, "Removing Noise from Music Using Local Trigonometric Bases and Wavelet Packets," *J. Audio Eng. Soc.*, vol. 42, no. 10, pp. 808–818, Oct. 1994.
46. D. L. Donoho, "De-noising by Soft-Thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
47. H. K. S. Mallat, D. L. Donoho, and A. S. Willsky, "Best Basis Algorithm for Signal Enhancement," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, May 1995, vol. 3, pp. 1561–1564.

48. P. K. Ramarapu and R. C. Maher, "Methods for Reducing Audible Artifacts in a Wavelet-Based Broad-Band Denoising System," *J. Audio Eng. Soc.*, vol. 46, no. 3, pp. 178–190, Mar. 1998.
49. B. Vidakovic, "Nonlinear Wavelet Shrinkage with Bayes Rules and Bayes Factors," *J. Am. Stat. Assoc.*, vol. 93, pp. 173–179, 1998.
50. J. Lim and A. V. Oppenheim, "All-Pole Modeling of Degraded Speech," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 26, no. 3, pp. 197–210, June 1978.
51. S. B. Jebara, A. Benazza-Benyahia, and A. B. Khelifa, "Reduction of Musical Noise Generated by Spectral Subtraction by Combining Wavelet Packet Transform and Wiener Filtering," *Proc. X European Signal Processing Conf.*, Tampere, Finland, Sept. 2000, vol. 2.
52. I. Kauppinen and K. Roth, "Audio Signal Extrapolation – Theory and Applications," *Proc. 5th Int. Conf. on Digital Audio Effects*, Hamburg, Germany, Sept. 2002, pp. 105–110, `http://www.unibw-hamburg.de/EWEB/ANT/dafx2002/papers.html`.
53. P. A. A. Esquef, *Model-Based Analysis of Noisy Musical Recordings with Application to Audio Restoration*, Ph.D. thesis, Helsinki University of Technology, Espoo, Finland, Apr. 2004, `http://lib.hut.fi/Diss/2004/isbn9512269503/`.
54. B. L. Vercoe, W. G. Gardner, and E. D. Scheirer, "Structured Audio: Creation, Transmission, and Rendering of Parametric Sound Representations," *Proc. IEEE*, vol. 86, no. 5, pp. 922–940, May 1998.

41

Audio Effects Generation

Udo Zölzer

Helmut Schmidt University, Hamburg, Germany

An **audio effect** is the modification of an audio signal to make it sound more interesting or to change its sound characteristic. The use of effects for sound modification has changed over the centuries and has been applied with different means. All effects simulate physical sound production and propagation by acoustical, mechanical or electronic means. The human perception of these sound effects leads to a classification according to the psychoacoustic terminology:

- Spectral envelope modification by filtering
- Time-domain amplitude envelope modification by dynamics processing or amplitude modulation
- Pitch shifting (affects the pitch by phase modulation)
- Time stretching (affects the duration without pitch modification)
- Room simulation (affects the room impression)

In the following we will introduce the effect topics:

- **Filters and delays** – like low-/high-/band-pass, shelving, peak and comb filters based on feedforward/feedback delays.
- **Tremolo, vibrato and rotary speaker** – loudness and pitch modification.
- **Chorus, Flanger and Phaser** - doubling, jet airplane effect.
- **Pitch shifting and time stretching**.

- **Dynamics processing** – limiter, compressor, expander, noise gate and distortion.
- **Room simulation** – infuses a room impulse response on the sound.

These effects will be introduced by their musical or physical meaning and then described by their time-domain operations and corresponding frequency-domain representations, where $x(t)$ is the input signal and $y(t) = T\{x(t)\}$ is the output signal with their corresponding Fourier transforms $X(\omega)$ and $Y(\omega)$ with $\omega = 2\pi f$.

1 • Filters and Delays

Filters are applied for simple input spectrum modifications. The standard filter operations for effects generation can be divided into first-order lowpass filters

$$y(t) = x(t) * h_{\mathrm{LP1}}(t) \rightarrow H_{\mathrm{LP1}}(\omega) = \omega_{\mathrm{c}}/[\omega_{\mathrm{c}} + j\omega], \tag{1}$$

first-order highpass filters

$$y(t) = x(t) * h_{\mathrm{HP1}}(t) \rightarrow H_{\mathrm{HP1}}(\omega) = j\omega/[\omega_{\mathrm{c}} + j\omega], \tag{2}$$

and allpass filters (first and second order) given by

$$\begin{aligned} y(t) = x(t) * h_{\mathrm{AP1}}(t) &\rightarrow H_{\mathrm{AP1}}(\omega) = [\omega_{\mathrm{c}} - j\omega]/[\omega_{\mathrm{c}} + j\omega] \\ &= H_{\mathrm{LP1}}(\omega) - H_{\mathrm{HP1}}(\omega) \end{aligned} \tag{3}$$

$$\begin{aligned} y(t) = x(t) * h_{\mathrm{AP2}}(t) \rightarrow H_{\mathrm{AP2}}(\omega) = &[\omega_{\mathrm{c}} - j\omega\omega_{\mathrm{c}}/Q - \omega^2]/ \\ &[\omega_{\mathrm{c}} + j\omega\omega_{\mathrm{c}}/Q - \omega^2]. \end{aligned} \tag{4}$$

These standard filters can easily be controlled by the cutoff frequency ω_{c} and the Q factor, which is the cutoff frequency divided by the $-3\,\mathrm{dB}$ bandwidth. These basic filters can also be implemented by special allpass decompositions which lead to a lowpass filter

$$y(t) = 0.5[x(t) + x(t) * h_{\mathrm{AP1}}(t)] \rightarrow H_{\mathrm{LP}}(\omega) = 0.5[1 + H_{\mathrm{AP1}}(\omega)], \tag{5}$$

highpass filter

$$y(t) = 0.5[x(t) - x(t) * h_{\mathrm{AP1}}(t)] \rightarrow H_{\mathrm{HP}}(\omega) = 0.5[1 - H_{\mathrm{AP1}}(\omega)], \tag{6}$$

and bandpass filter

$$y(t) = 0.5[x(t) - x(t) * h_{\mathrm{AP2}}(t)] \rightarrow H_{\mathrm{BP}}(\omega) = 0.5[1 - H_{\mathrm{AP2}}(\omega)]. \tag{7}$$

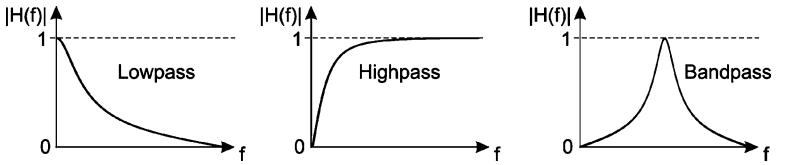


FIGURE 1 Frequency responses of simple filters.

These filters can easily be tuned or adjusted by changing the first- and second-order allpass coefficients to control the cutoff or center frequency of the corresponding filter. The frequency responses are shown in Figure 1.

Based on the allpass decompositions of lowpass, highpass and bandpass filters special weighting filters, such as a shelving filter given by

$$y(t) = x(t) + v\, x(t) * h_{\mathrm{LP/HP}}(t) \rightarrow H_{\mathrm{SF}}(\omega) = 1 + v\, H_{\mathrm{LP/HP}}(\omega) \qquad (8)$$

for low-frequency and high-frequency boost and cut and the peak filter for the mid-frequencies given by

$$y(t) = x(t) + v\, x(t) * h_{\mathrm{BP}}(t) \rightarrow H_{\mathrm{PF}}(\omega) = 1 + v\, H_{\mathrm{BP}}(\omega), \qquad (9)$$

can be implemented [7,13,21,28]. These allpass decompositions allow independent control of boost and cut, cutoff or center frequency and bandwidth for shelving and peak filters (see Figure 2).

Delays occur in sound propagation and can be described for a single delay by

$$y(t) = x(t - \tau) \rightarrow H(\omega) = 1 \exp(-\mathrm{j}\omega\tau). \qquad (10)$$

In a variety of applications we have a sum of several delayed replica of the input signal $x(t)$. If the direct signal $x(t)$ and the

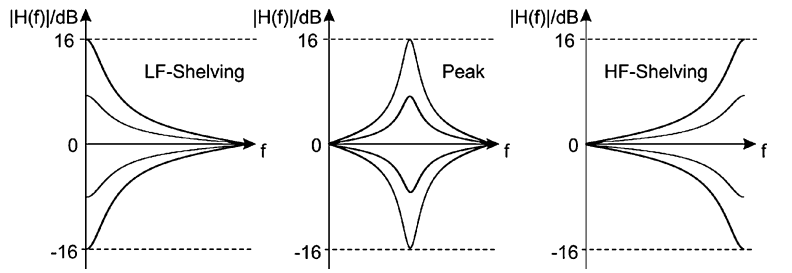


FIGURE 2 Frequency responses of shelving and peak filters.

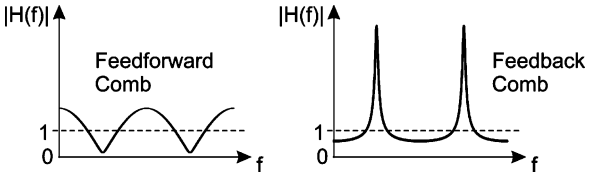


FIGURE 3 Frequency responses of feedforward and feedback comb filters.

delayed signal $x(t-\tau)$ are summed up the so-called feedforward comb filter given by

$$y(t) = x(t) + b\ x(t-\tau) \rightarrow H(\omega) = 1 + b\ \exp(-\mathrm{j}\omega\tau) \tag{11}$$

is performed, which leads to a frequency response with repetitive notches controlled by coefficient b (see Figure 3). If the output is delayed the feedback comb filter given by

$$y(t) = x(t) - a\ y(t-\tau) \rightarrow H(\omega) = 1/(1 + a\ \exp(-\mathrm{j}\omega\tau)) \tag{12}$$

is implemented, which leads to repetitive peaks in the frequency response controlled by coefficient a (see Figure 3).

The combination of weighted and delayed input and output signals leads to an allpass comb filter given by

$$\begin{aligned} y(t) &= ax(t) + x(t-\tau) - a\ y(t-\tau) \rightarrow H(\omega) \\ &= [a + \exp(-\mathrm{j}\omega\tau)]/[1 + a\ \exp(-\mathrm{j}\omega\tau)]. \end{aligned} \tag{13}$$

The allpass comb filter has a unity gain frequency response. All three types of comb filters have been intensively used in room simulation devices [19,24,25].

2 • Tremolo, Vibrato and Rotary Speaker

Tremolo is the rapid repetition of one note in music terminology or a rapid alternation between two or more notes (string instruments, mandolin, flute, piano, harpsichord, spinet, guitar). In technical terminology tremolo is a rapid repetitive increase and decrease in volume. The effect is based on modifying the time-domain envelope (amplitude) of a signal according to

$$y(t) = m(t)\ x(t) \text{ with } \mathrm{FT}\{y(t)\} = Y(\omega) = M(\omega) * X(\omega)/2\pi \tag{14}$$

with the modulating signal $m(t) = 1 + a\sin(\omega_0 t)$. The repetition rate is controlled by $\omega_0 = 2\pi f_0$. The depth of the amplitude modulation is adjusted by a.

Vibrato is the rapid and repetitive variation in pitch for the duration of a note (string instruments, guitar, singing voice, yodeling). In technical terminology the effect is based on modifying the pitch of a signal by delay line modulation (phase modulation) according to

$$y(t) = x(t - m(t)) = x(t) * \delta(t - m(t)) \text{ with}$$
$$\text{FT}\{y(t)\} = Y(t, \omega) = X(\omega) \exp(-\text{j}\omega m(t)) \tag{15}$$

with the time-varying delay $m(t) = t_1 + a \sin(\omega_0 t)$ with the fixed delay t_1 and the sinusoidal term controlled by $\omega_0 = 2\pi f_0$. The depth of the phase modulation is adjusted by a. The pitch modification by delay line modulation has its physical correspondence in the Doppler effect. Methods for varying the length of fractional delay lines are discussed in [4,14].

A **rotary speaker** produces a Doppler effect and an amplitude modulation. The Doppler effect leads to a pitch modulation and the amplitude modulation affects a loudness modulation. This effect can be described by a delay line modulation and an amplitude modulation according to

$$y(t) = a_1(t) x(t - m_1(t)) \tag{16}$$

with parameters $m_1(t) = t_1 + a \sin(\omega_0 t)$ and $a_1(t) = 0.5[1 + \sin(\omega_0 t)]$ [5,6,27].

3 • Chorus, Flanger and Phaser

Chorus simulates the result of an ensemble of musicians (singers, violins) playing together with slight pitch and amplitude differences. The chorus effect performs a loudness increase. These amplitude and pitch differences are simulated by amplitude modulation and delay line modulation (phase modulation) and all signals are summed together to form the output signal

$$y(t) = \Sigma_i x(t - m_i(t)) \tag{17}$$

with $m_i(t) = t_i + r_i(t) * h_{\text{LP}}(t)$, where t_i is a fixed delay, $r_i(t)$ is a random process which is bandlimited by a lowpass filter with impulse response $h_{\text{LP}}(t)$.

A **flanger** simulates a special filter effect which occurs if a slowly time-varying single reflection is added to the direct signal (jet airplane effect). This effect was achieved by playing two copies of a single signal from two tape machines while slowing

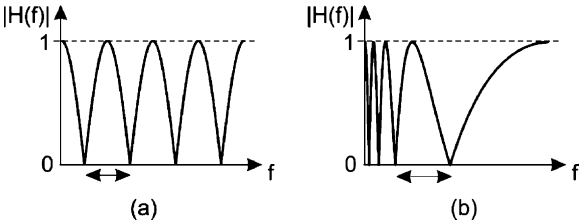


FIGURE 4 Frequency responses of (**a**) flanger and (**b**) phaser.

down the replay speed of one tape machine by placing a finger to the flange of one tape reel. The filter effect is described by

$$y(t) = 0.5x(t) + 0.5x(t - m_1(t)) \text{ and} \tag{18}$$

$$\text{FT}\{y(t)\} = Y(\omega) = X(\omega)H(\omega) = 2X(\omega)|\cos(\omega m_1(t))|\exp(-j\omega m_1(t)), \tag{19}$$

where the slowly time-varying delay is $m_1(t) = t_1 + a\sin(\omega_0 t)$ with $t_1 = 1/f_1$. The frequency f_1 controls the frequency spacing of the zeros in the transfer function. This effect produces equally spaced attenuations to zero across the frequency spectrum, where the zeros move slowly up and down the frequency axis (see Figure 4) by choosing a time-varying delay time $m_1(t)$.

A **phaser** performs a similar effect as a flanger, but instead of the time-varying delay line, which performs a frequency-independent delay, a series of frequency-dependent allpass filters is used which leads to the output signal

$$y(t) = 0.5x(t) + 0.5x(t) * [h_{\text{AP1}}(t) * h_{\text{AP2}}(t) * \cdots * h_{\text{APN}}(t)]. \tag{20}$$

Single allpass sections $h_{\text{AP}i}(t)$ can be first-order allpass filters which have a phase response which ranges from 0 to $-\pi$, in the series connection the allpass phase response ranges from 0 to $-N\pi$. The phase crosses multiples of $-\pi$ $N/2$ times, so notches will be produced at those frequencies. Due to the non-linear phase response the multiples of π are spaced non-uniformly across the frequency spectrum (see Figure 4), where the zeros move slowly up and down the frequency axis when the allpass filter coefficients are varied with a slow speed. Further details on chorus, flanger and phaser can be found in [4,5,10,11,26,29].

4 • Pitch Shifting and Time Stretching

Pitch shifting modifies the pitch of a signal by increasing the playback speed of a recorded signal and thus rising the pitch or decreasing the playback speed of the signal and thus lowering

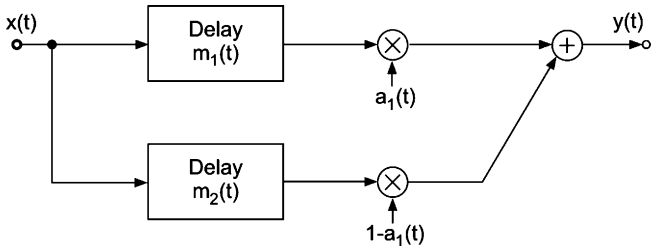


FIGURE 5 Pitch shifting by delay line modulation and overlap and add.

the pitch. The fundamental correspondence is given by the time scaling theorem of the Fourier transform $\mathrm{FT}\{x(at)\} = X(\omega/a)/|a|$, which shows that time-domain scaling leads to a frequency-domain scaling. First pitch shifting devices have been introduced in the late 1930s with the help of tape machines and rotary replay heads for secure audio transmission applications. The basic operations of pitch shifting can also be achieved by two or several parallel modulated delay lines (see Figure 5), where the output signals are cross-faded as given by

$$y(t) = a_1(t)x(t - m_1(t)) + [1 - a_1(t)]x(t - m_2(t)). \tag{21}$$

The time-varying delays are $m_1(t) = t_0 + (t \bmod t_1)$ and $m_2(t) = m_1(t - t_1/2)$, which represent sawtooth signals with a special slope $1/t_1$ for controlling the Doppler effect for pitch shifting each delay line output. The cross-fade functions are given by

$$a_1(t) = 0.5[1 + \sin(\omega_1 t)] \text{ and } \omega_1 = 2\pi/t_1, \tag{22}$$

which allows the cross-fading when the read pointer for delay $m_1(t)$ jumps back to fade over to delay $m_2(t)$ and vice versa [29, Chap. 7]. This time-domain approach for pitch shifting modifies the spectral envelope of the spectrum according to the scaling theorem of the Fourier transform. A method to perform pitch shifting with preservation of the spectral envelope is explained below by using a frequency-domain approach [1,15–17].

Time stretching can be achieved by segmentation of the input signal $x(t)$ into over-lapping blocks with an analysis hop size S_a, taking the unmodified input blocks $x_i(t)$, and performing an overlap and add operation with synthesis hop size $S_s = \alpha\ S_a$ to form the output signal $y(t)$ (see Figure 6). The coefficient α controls the increase ($\alpha > 1$) or decrease ($\alpha < 1$) of the stretching operation. In the overlap region a suitable window function (cross-fading between block overlap) is used [29, Chap. 7].

The same analysis and synthesis approach can be used to perform a frequency-domain processing (see Figure 7) of the individual

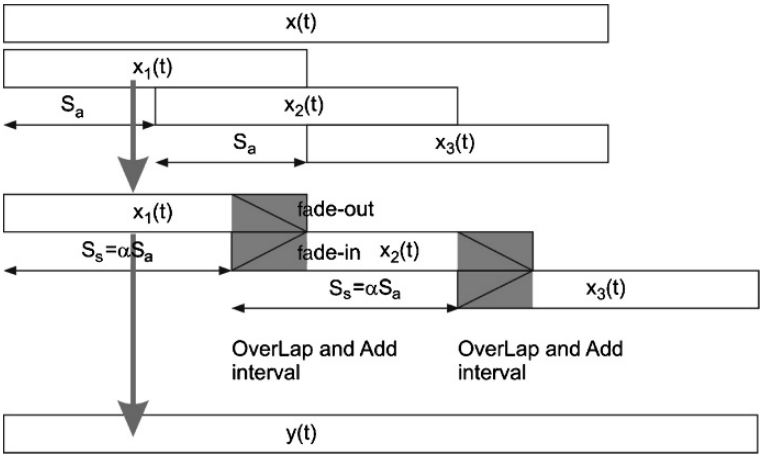


FIGURE 6 Time stretching by input segmentation and modified overlap and add operation.

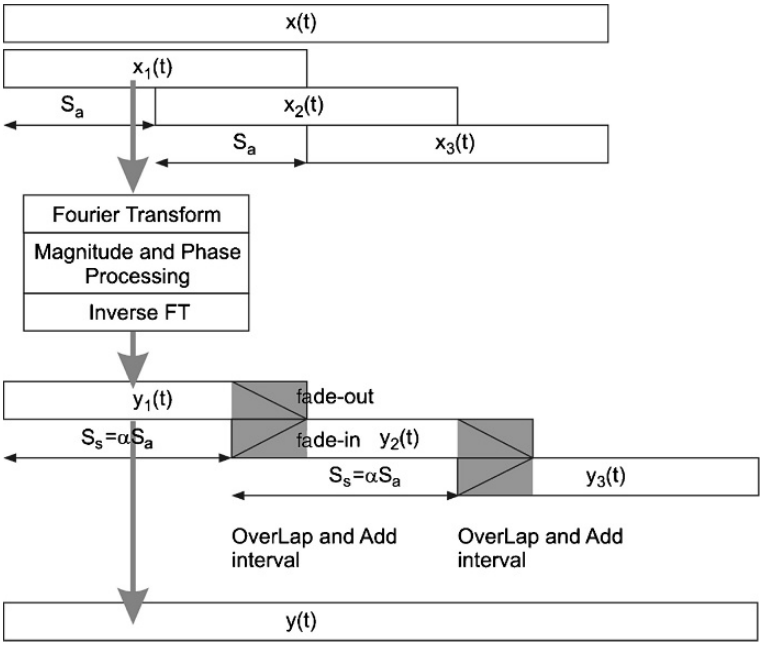


FIGURE 7 Frequency-domain processing for pitch shifting and time stretching.

input blocks $x_i(t)$, modifying the magnitude and the phase of the corresponding Fourier transforms $Y_i(\omega) = f\{X_i(\omega)\}$, and perform an inverse Fourier transform yielding modified blocks $x_i(t)$. The weighted overlap and add operation reconstructs the modified output signal $y(t)$. This frequency-domain approach allows several audio effects (denoising, stable/transients separation, robotization, whisperization, sinusoidal modeling, source separation) besides time stretching and pitch shifting which can preserve the spectral envelope of the input spectrum thus avoiding the so-called Mickey Mouse effect [1,15–17].

5 • Dynamics Processing

Dynamics processing is necessary to reduce or expand the dynamics of the audio signal. It is used in recording or transmission applications and fits the dynamic range to the following signal processing steps. Dynamics processing (see Figure 8) is an operation which performs a modification of the time-domain amplitude envelope according to

$$y(t) = m(t)\, x(t) \text{ with } \mathrm{FT}\{y(t)\} = Y(\omega) = M(\omega) * X(\omega)/2\pi. \quad (23)$$

The time-varying gain function $m(t) = f\{x(t)\}$ is derived from the input signal and represents the instantaneous root mean square or absolute value of the input, which is measured with appropriate time constants. Typical operation modes are limiting, compression, expansion and noise gating [18,28]. The amplitude modification by multiplication shows the convolution of the corresponding spectra. Therefore a bandlimited gain function is necessary, otherwise a spreading of the input spectrum is performed. In several applications a non-linear behavior of the dynamics system is of special interest, for example

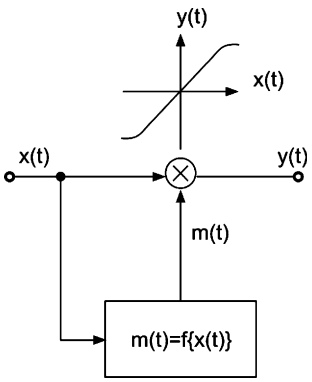


FIGURE 8 Dynamic range processing by non-linear processing.

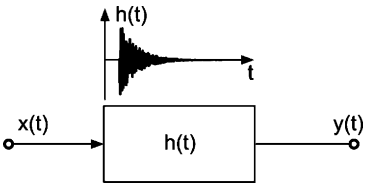


FIGURE 9 Room simulation by fast convolution.

when non-linear transfer characteristics like triode, pentode or tape saturation characteristics are preferred for sound shaping. The spreading of the input spectrum leads to a higher signal bandwidth of the output signal [29].

6 • Room Simulation

Room simulation is the generation of artificial reverberation [2,3,9,12,19,22–25,28] and infuses a room impulse response on the signal. There have been several simplified approximations of room impulse responses over the past 40 years which are based on feedforward and feedback comb and allpass filters in a variety of different configurations. If a room impulse response $h(t)$ (see Figure 9) can be derived by measurements (pseudo-random signals or swept-sine techniques) of a real room or model-based computations (ray-tracing or image model methods), the room simulation applied to the input signal is given by the convolution operation

$$y(t) = x(t) * h(t) \text{ with } \mathrm{FT}\{y(t)\} = Y(\omega) = X(\omega)H(\omega). \tag{24}$$

Digital signal processing techniques allow the segmentation of the input signal and the impulse response thus leading to a parametric convolution operation given by

$$y(t) = x(t) * [a_1 h_1(t) + a_2 h_2(t) + a_3 h_3(t) + a_4 h_4(t)]. \tag{25}$$

Performing the partial convolutions in the frequency domain allows the fast convolution of the entire impulse response [8,20] with the input signal with nearly perfect psychoacoustical properties (see Figure 9).

References

1. D. Arfib, "Time-Frequency Processing" and "Source-Filter Processing", in U. Zölzer (ed.) DAFX – Digital Audio Effects, Wiley, 2002.
2. B. Blesser, "An Interdisciplinary Synthesis of Reverberation Viewpoints", J. Audio Eng. Soc., Vol. 49, No. 10, pp. 867–903, Oct. 2001.

3. J. Dattorro, "Effect Design: Part 1 Reverberator and Other Filters", J. Audio Eng. Soc., Vol. 45, No. 9, pp. 660–684, Sep. 1997.
4. J. Dattorro, "Effect Design: Part 2 Delay-Line Modulation and Chorus", J. Audio Eng. Soc., Vol. 45, No. 10, pp. 764–788, Oct. 1997.
5. J. Dattorro, "Effect Design, Part 3 Oscillators: Sinusoidal and Pseudonoise", J. Audio Eng. Soc., Vol. 50, No. 3, p. 115, 2002.
6. S. Disch, U. Zölzer, "Modulation and Delay Line Based Digital Audio Effects, Proc. DAFx-99 Workshop on Digital Audio Effects, pp. 5–8, Trondheim, 1999.
7. F. Fontana, M. Karjalainen, "A Digital Bandpass/Bandstop Complementary Equalization Filter with Independent Tuning Characteristics", IEEE Signal Process. Lett., Vol. 10, No. 4, pp. 119–122, Apr. 2003.
8. W. G. Gardner, "Efficient Convolution Without Input–output Delay," J. Audio Eng. Soc., Vol. 43, No. 3, pp. 127–136,1995.
9. W. G. Gardner, "Reverberation Algorithms," in M. Kahrs and K. Brandenburg (eds), Applications of Digital Signal Processing to Audio and Acoustics, Kluwer, pp. 85–131, 1998.
10. W. M. Hartmann, Signals, Sound, and Sensation, AIP Press, 1997.
11. W. M. Hartmann, "Flanging and Phasers", J. Audio Eng. Soc., Vol. 26, No. 6, pp. 439–443, June 1978.
12. J. M. Jot, A. Chaigne, "Digital Delay Networks for Designing Artificial Reverberators", Proc. 94th AES Convention, Preprint No. 3030, 1991.
13. F. Keiler, U. Zölzer, "Parametric Second- and Fourth-Order Shelving Filters for Audio Applications", Proc. of IEEE 6th Workshop on Multimedia Signal Processing, Siena, Italy, September 29–October 1 , 2004.
14. T. I. Laakso, V. Välimäki, M. Karjalainen, U. K. Laine, "Splitting the Unit Delay—Tools for Fractional Delay Filter Design", IEEE Signal Process. Mag., Vol. 13, No. 1, pp. 30–60, Jan. 1996
15. J. Laroche, "Time and Pitch Scale Modifications", in M. Kahrs and K.-H. Brandenburg (eds), Applications of Digital Signal Processing to Audio and Acoustics, Kluwer 1998.
16. J. Laroche and M. Dolson, "Improved Phase Vocoder Time-Scale Modification of Audio", IEEE Trans. Speech Audio Process., Vol. 7, No. 3, pp. 323–332, 1999.
17. J. Laroche and M. Dolson, "New Phase-Vocoder Techniques for Pitch Shifting, Chorusing, Harmonizing, and Other Exotic Audio Modifications", J. Audio Eng. Soc., Vol. 47, No. 11, pp. 928–936, 1999.
18. G. McNally, "Dynamic Range Control of Digital Audio Signals", J. Audio Eng. Soc., Vol. 32, No. 5, pp. 316–327, 1984.
19. J. A. Moorer, "About this Reverberation Business", Comput. Music J., Vol. 3, No. 2, pp. 13–28, 1978.
20. A. V. Oppenheim, R. W. Schafer, Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1975.
21. S. J. Orfanidis, Introduction to Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1996.

22. D. Rocchesso, “Spatial Effects”, in U. Zölzer (ed.) DAFX – Digital Audio Effects, Wiley, 2002.
23. D. Rocchesso and J. O. Smith, “Circulant and Elliptic Feedback Delay Networks for Artificial Reverberation”, IEEE Trans. Speech Audio Process., Vol. 5, No. 1, pp. 51–63, Jan. 1997.
24. M. R. Schroeder, “Natural Sounding Artificial Reverberation”, J. Audio Eng. Soc., Vol. 10, No. 3, pp. 219–223, 1962.
25. M. R. Schroeder, B. F. Logan, “Colorless Artificial Reverberation,” J. Audio Eng. Soc., Vol. 9, No. 3, pp. 192–197, 1961.
26. J. O. Smith, “An Allpass Approach to Digital Phasing and Flanging”, Proceedings of the 1984 International Computer Music Conference, pp. 103–108, 1984.
27. J. O. Smith, S. Serafin, J. Abel, D. Berners, “Doppler Simulation and the Leslie”, Proc. of the 5th Int. Conference on Digital Audio Effects (DAFx-02), Hamburg, Germany, September 26–28, 2002.
28. Udo Zölzer, Digital Audio Signal Processing, 2e, Wiley, 2008.
29. U. Zölzer (ed), DAFX - Digital Audio Effects, Wiley, 2002.

PART VII
TELECOMMUNICATIONS

H.W. Gierlich

Telecom Division, HEAD acoustics GmbH, Herzogenrath, Germany

43 Speech Communication and Telephone Networks
H.W. Gierlich 821
General Aspects of Speech Communication • Modern Telephone Networks

44 Methods of Determining the Communicational Quality of Speech Transmission Systems
H.W. Gierlich 831
Assessment Methods • Subjective Assessment • Objective Methodologies • Network Planning

45 Efficient Speech Coding and Transmission Over Noisy Channels
Ulrich Heute and Norbert Goertz 853
Introduction and Fundamentals • Speech Coding • Speech Transmission over Noisy Channels • Further Work, Results and Literature

46 Echo Cancellation
Walter Kellermann 883
Echo Cancellation and Suppression

47 Noise Reduction and Interference Cancellation
Walter Kellermann 897
Single-Channel Techniques • Multichannel Techniques • Blind Signal Separation of Convolutive Mixtures

48 Terminals and Their Influence on Communication Quality
Walter Kellermann 909
Acoustic Interfaces

49 Networks and Their Influence on Communication Quality

H.W. Gierlich 915

PSTN (Public Switched Telephone Networks) • Packet-Based Transmission • Mobile Networks

50 Interaction of Terminals, Networks and Network Configurations

H.W. Gierlich 921

Tandeming of Speech Coders • Delay • Tandeming of Speech Echo Cancellers • Tandeming VAD and NLP • Other Signal Processing

List of Important Abbreviations 927

43

Speech Communication and Telephone Networks

H.W. Gierlich

HEAD accoustics GmbH, Herzogenrath, Germany

Speech communication over telephone networks has one major constraint: The communication has to be "real time". The basic principle since the beginning of all telephone networks has been to provide a communication system capable of substituting the air path between two persons having a conversation at 1-m distance. This is the so-called orthotelephonic reference position [7]. Although many technical compromises must be made to enable worldwide communication over telephone networks, it is still the goal to achieve speech quality performance which is close to this reference.

This chapter gives an overview of terminals, networks, their effects on speech quality, and the digital signal processing involved in keeping the quality as high as possible.

1 • General Aspects of Speech Communication

Speech communication is influenced by a variety of parameters. In contrast to broadcast services, the parameters contributing to the quality of communication service perceived by the user – for the service as a whole – are more complex than the factors contributing to speech quality only. As described by Möller [13], various quality entities are relevant. The main quality entities and their relationship are shown in Figure 1.

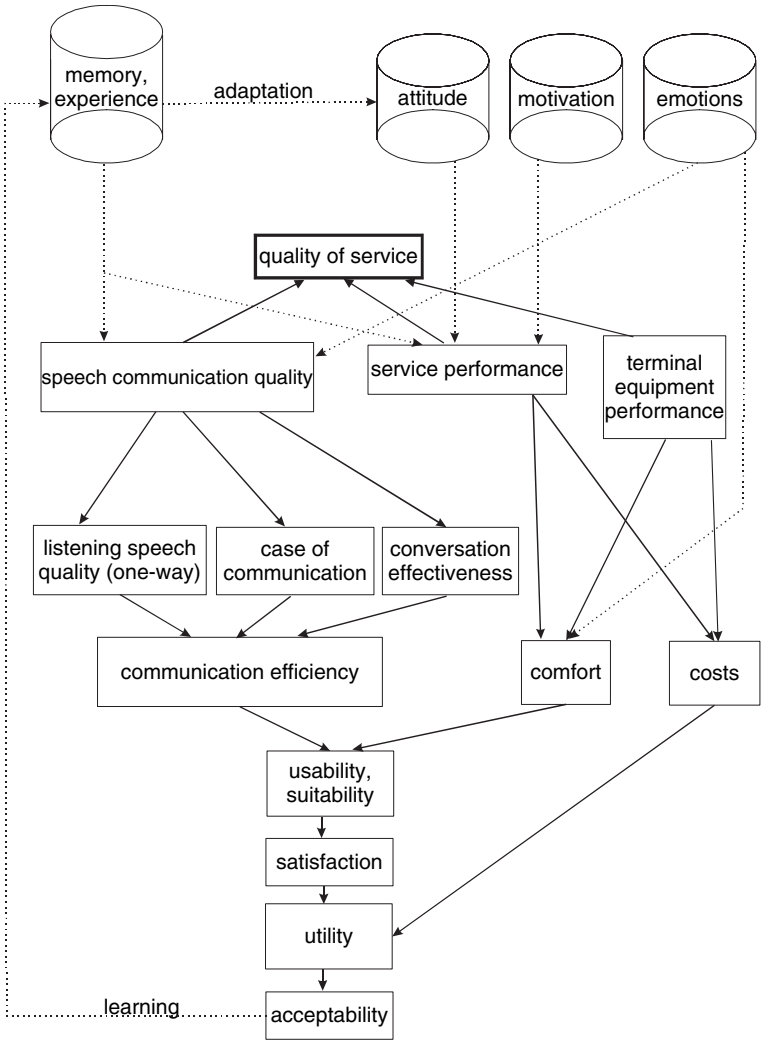


FIGURE 1 Parameters influencing the quality of service (QoS) for a telecommunication service (from [13]).

When focusing on the speech communication aspect, the listening situation, the talking situation, the conversational situation (including double talk), and the background noise situation have to be taken into account and need to be regarded separately. An overview of the parameters influencing speech quality is shown in Figure 2.

The major criterion used in early times for the optimization of traditional telephone communication systems is "speech

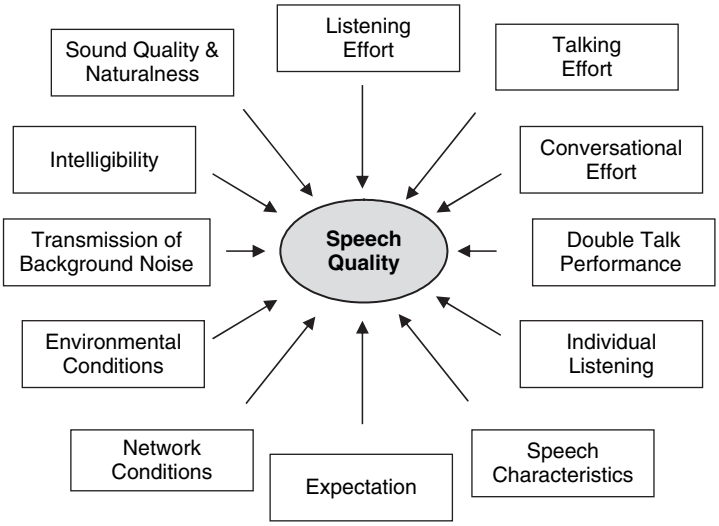


FIGURE 2 Factors influencing the speech quality perceived by the user.

intelligibility" [6,14]. This parameter purely focuses on the transport of information. Based on different kinds of intelligibility tests [6], the requirements for bandwidth, signal-to-noise ratio, distortion, and loudness requirements of the traditional telephone system (defined in ITU-T as 3.1 kHz telephony) were introduced from the beginning of telecommunication system development. Although the technical restrictions are no longer in place for many of the modern communication systems, narrowband telephony is still used in the majority of the connections today. Most speech quality parameters were derived for narrowband (3.4 kHz) telecommunication systems. Wideband communication systems are available, but not widely used yet. This may change with the use of IP-based transmission where a much more flexible bandwidth allocation for the individual call is possible.

When judging the listening situation, the quality of the transmitted speech signal is of major importance besides the intelligibility.

In the *listening situation* the speech signal may be degraded by a variety of parameters.

Loudness. The loudness perceived in the listening situation mainly contributes to the perceived speech quality. The speech signal loudness should be comparable to the loudness perceived when having a conversation between two persons at 1-m distance, the "orthotelephonic reference position" [7]. The

loudness in today's telecommunication networks is considered the main parameter used for the planning of telephone networks and the main parameter which has to be determined for all devices involved in a communication chain. The loudness calculation in telephonometry is different from the loudness calculation used, for example, in psychoacoustics and is described in detail in Chapter 44.3.

Frequency response, linearity, distortion, and noise. In modern telephone systems, the frequency response characteristics are mainly determined by the telephone terminal. Except for older analog transmission systems where the telephone line may strongly influence the overall connection quality due to frequency-dependent attenuation and group delay distortion, it is assumed that the network components typically do not influence the frequency response characteristics significantly. Requirements for distortion, linearity, and noise can be found in many relevant standards.

Speech sound quality. The speech sound quality in classical telephones can be described most comprehensively by the loudness, frequency response, distortion, and the noise introduced by the transmission system. In modern telecommunication, however, speech coding and other types of non-linear and time-variant signal processing are widely used. The speech quality of these systems must be determined subjectively. For known transmission systems, the speech quality can be described using objective speech quality measures like PESQ or TOSQA [4,12]. However, whenever new non-linear or time-variant transmission systems are to be investigated, subjective evaluation is required.

In the *talking situation* the speech quality is influenced by impairments typically related to the user's own voice, which is referred to as *sidetone*, coupled either acoustically or electrically (by the handset receiver) to the user's ear and echo. The echo signal is produced either by acoustical coupling between the receiver and the microphone of the far-end terminal, or may be due to a hybrid echo, e.g. resulting from a hybrid in an analog telephone or a hybrid somewhere else in the connection. The perception of echo depends mainly on two parameters: the echo delay and the echo attenuation. The longer the delay the more echo attenuation is required (see [9]). For typical telephones, an echo loss of 45 dB is required as a minimum.

In the *conversational situation*, speech quality is influenced by many parameters.

Delay. Delay is introduced in every communication system. Besides delay introduced by the acoustical paths between

mouth and microphone of a telephone and between the telephone receiver and the listener's ear, there are other sources of delay. Algorithmic delay may be introduced by the signal processing in the terminal such as background noise reduction (noise cancellation), echo cancellation, speech coding, voice-controlled attenuation, and others. Packetization, for example, as used in VoIP (voice over Internet protocol) systems may further contribute to the delay introduced in a connection. The main effect of delay is the impairment of the conversational dynamic. For highly interactive conversational tasks, a delay of not more than 150 ms is recommended (see [8]). Another effect of delay is the increased audibility of echo components for increasing delay times (see [9]).

Double talk capability. The double talk capability of a telephone system is mainly determined by its echo and switching performance. The longer the delay in a connection the more echo attenuation is required. Since echo cancellation is never perfect, additional means have to be introduced in order to avoid echo components. Typically, additional gain switching is introduced in the terminals as well as in speech echo cancellers in the network. The switching may be introduced in sending and receiving directions. Requirements for switching and echo during double talk and their relationship to the perceived quality of a telephone system can be found in [11].

A parameter not directly related to speech but highly influencing the overall quality perceived by the user is the *background noise transmission performance*. Background noise is present during many telephone calls and may influence the perceived overall quality of a connection quite significantly. Background noise may influence the transmission quality during periods where no speech signal is present, during periods with near-end speech and with far-end speech. Furthermore, the speech signal may be degraded by noise reduction systems in the presence of background noise. The subjective (auditory) relevance of the background noise situation in comparison to the other conversational situation is still under investigation.

2 • Modern Telephone Networks

2.1 PSTN (Public-Switched Telephone Networks)

In general, all parameters determining the speech quality in circuit-switched networks are well controlled. PSTN networks are typically analog-switched networks or TDM (time division multiplex)-type networks. Proper transmission planning generally can ensure high speech quality in these networks. Digital PSTN-type networks work isochronously. The loss of data

is fairly unlikely. The delays in a connection are constant and within the control of the network operator. The network interfaces in such networks are well defined. In analog networks, the transmission line characteristics (length and type of lines) have to be taken into account since it may have impact on signal level, frequency response characteristics, and noise. The delay inserted by traditional national networks is well controlled; proper transmission planning guarantees low delay (< 15 ms) and echo-free connection for national connections. Network speech echo cancellation which is under the control of the network operator ensures mostly echo-free connections in case of interconnection to international networks or satellite communication. Echoes which might result from insufficient hybrid loss of analog terminals connected to the PSTN network are cancelled.

2.2 Packet-Based Transmission

Packet-based networks can be divided into ATM types (asynchronous transfer mode) of networks and IP networks. IP networks differ considerably from circuit-switched networks but in ATM-type networks, the same rules can be applied as for TDM-type networks, except for the packetization process itself. Control of the parameters relevant for achieving a sufficient or good communication quality is impossible, or achievable only within limits. Therefore, transmission planning only can be made based on assumptions, statistical knowledge, and best-worst-case estimations within the constraints given by a specific network and its interconnection to other types of networks. Any sort of guaranteed speech quality can be given on a statistical basis only.

In packet-based networks, the input signal (speech signal) is typically segmented into packets of fixed length. In IP networks, all users connected to the specific IP-network share the bandwidth given by the IP network; no guarantee for a required bandwidth is given. The packets of different sources (e.g. terminals or subscribers) have to be sequenced for transmission. Neither delay nor jitter is predictable in IP networks. Both can be described only by statistical parameters such as average delay (including delay distribution) and packet loss distribution.

IP networks do not provide any means to control the speech signals, so the sender and receiver have to provide the appropriate signal processing mechanisms. Speech coding, packetization, echo cancellation, noise reduction, jitter buffer, packet loss concealment, and other types of signal processing have to be integrated either into the terminals or the gateways interconnecting the IP network to other types of networks.

2.3 Mobile Networks

In mobile networks, the speech quality is highly dependent on the transmission network. A wide variety of network conditions can be expected due to the time-variant characteristics of the radio channel. In all mobile networks, speech coding is used in order to reduce the bandwidth required for the individual speech channel and to increase the number of channels which may be transmitted. The upper limit for the listening speech quality, therefore, is set by the codecs (see [1–3]). While the speech codec quality is predictable as long as clean speech is encoded, the overall speech quality may be degraded by different impairments introduced by the network or the mobile terminals.

The mobile terminal itself should provide all means to ensure a good conversational quality within the constraints given by the mobile transmission itself. Background noise cancellation, speech coding, acoustical echo cancellation, frame loss concealment, and other types of signal processing are integrated in the mobile terminals. The mobile network itself is responsible for routing and switching of the speech signal. The general configuration of a mobile network is shown in Figure 3 for the example of a GSM network.

Disturbances introduced by the radio channel between the mobile station and the base transceiver station result in bit errors. A high bit error rate (BER) contributes to degraded speech. In addition, transmission errors may occur between BTS and MSC. While the PCM-coded speech used between TRC and MSC is less sensitive to the effects of high BER, the GSM-

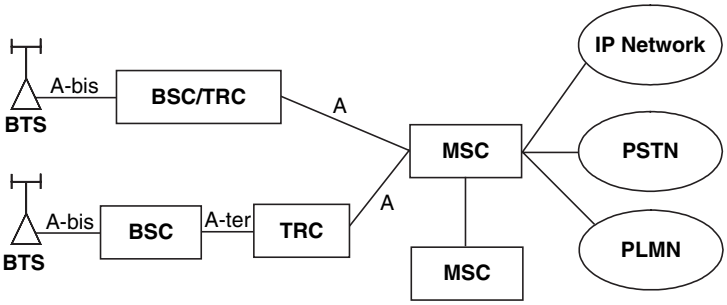


FIGURE 3 Network configuration (see [5]). BTS, base transceiver station; BSC, base station controller; TRC, transcoder controller; MSC, mobile switching centre; PLMN, public land mobile network; PSTN, public-switched telephone network; A, interface between the BSS and the MSC, manages the allocation of suitable radio resources to the MSs and mobility management; A-bis, BSS internal interface linking the BSC and a BTS; A-ter, BSS internal interface linking the BSC and a TRC.

coded speech transmitted between the BTS and the TRC is more sensitive to bit errors. The audible effects on speech quality due to increasing or decreasing BER, on the A-bis and A-ter interfaces, are unnoticeable, provided the transmission fulfills the performance objectives stated in [10].

Varying speech levels in the network affect the loudspeaker level, distortion, and echo canceller performance. Poor network echo cancelling results in disturbing echo for calls to the PSTN and PLMN. The network echo cancellation is provided by the PSTN. If a proper echo control cannot be guaranteed by the PSTN, echo cancellation has to be provided by the MSC (see Figure 3). In the mobile transmission network, a high bit error rate causes serious degradation of the listening speech quality. The most sensitive part of the connection is the connection between the A-bis and A-ter interfaces. The radio network environment is affected by the following factors (see [5]):

- Co-channel interference (C/I),
- Noise limitations (C/N),
- Mobile speed (fading frequency),
- Time dispersion
- Adjacent channel interference (C/A)

During a call, these factors together contribute to the bit error rate (BER – the average amount of bit errors in a speech frame) and frame erasure rate (FER – the percentage of erased frames). In a mobile network, good cell planning is of critical importance to minimize these types of interferences. The quality of the radio environment can be further improved by using the radio network features such as DTX, power control, and frequency hopping.

Discontinuous transmission (DTX) reduces the total interference in the network since signal is only transmitted during speech activity. DTX and comfort noise generation are parts of the speech codec and can therefore differ for different speech codecs.

Frequency hopping can reduce the effect on speech quality caused by multipath fading and interference. The signal strength variations, or interference, are broken up into pieces of duration short enough for the de-interleaving and speech-decoding process to correct the errors.

Dynamic power control may result in a further increase in the protection against fading dips and interference. The output power is controlled with respect to signal strength of the received radio signal as well as with respect to the BER.

The speech quality is always impaired during *handover* (when the call is passed from one BSC to another). During handover, the speech is interrupted shortly; speech frames are used to transmit signalling information.

References

1. 3GPP: TS 46.010: Full Rate Speech Encoding, Third Generation Partnership Project (3GPP), 2002.
2. 3GPP: TS 46.051: GSM Enhanced Full Rate Speech Processing Functions: General Description, Third Generation Partnership Project (3GPP), 2002.
3. 3GPP: TS 46.090: AMR Speech Codec; Transcoding Functions, Third Generation Partnership Project (3GPP), 2002.
4. Berger, J.: Instrumentelle Verfahren zur Sprachqualitätsschätzung – Modelle auditiver Tests, PHD thesis, Kiel, 1998.
5. ETSI EG 201 377-2: Specification and Measurement of Speech Transmission Quality; Part 2: Mouth-to-Ear Speech Transmission Quality Including Terminals, 2004.
6. Fletcher, H. and Galt, R.H.: The Perception of Speech and Its Relation to Telephony. J. Acoust. Soc. Am., 1950, Vol. 22, No. 2, pp. 90–119.
7. ITU-T Handbook on Telephonometry, ITU-T, Geneva, 1992.
8. ITU-T Recommendation G. 114: One-Way Transmission Time, International Telecommunication Union, Geneva, 2000.
9. ITU-T Recommendation G. 131: Control of Talker Echo, International Telecommunication Union, Geneva, 2000.
10. ITU-T Recommendation G. 821: Error Performance of an International Digital Connection Operating at a Bit Rate Below the Primary Rate and Forming Part of an Integrated Services Digital Network, International Telecommunication Union, Geneva, 2002.
11. ITU-T Recommendation P. 340: Transmission Characteristics and Speech Quality Parameters of Hands-Free Telephones, International Telecommunication Union, Geneva, 2000.
12. ITU-T Recommendation P.862: Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs, International Telecommunication Union, Geneva, 2001.
13. Möller, S.: Assessment and Prediction of Speech Quality in Telecommunications, Kluwer Academic Press, 2000, ISBN 0-7923-7894-6.
14. Richards, D.L.: Telecommunication by Speech, Wiley New York, 1973, ISBN 0-470-71949 4.

44

Methods of Determining the Communicational Quality of Speech Transmission Systems

H.W. Gierlich

HEAD accoustics GmbH, Herzogenrath, Germany

Telecommunication systems generally are designed for bi-directional exchange of information. Therefore, quality assessment of telecommunication systems has to reflect all aspects of realistic conversational situations. Speech quality assessment is part of the more global assessment of the service quality [34]. Since all quality parameters are based on sensations perceived by the subjects using the communication system, the basis of all objective assessment methods are subjective tests. They have to be defined carefully in order to reflect the real use situation as closely as possible and, on the other hand, provide reproducible and reliable results under laboratory conditions.

1 • Assessment Methods

Modern telecommunication networks provide a wide range of voice services. The rapid deployment of different types of new technologies leads to an increased need for evaluating the

transmission characteristics of new transmission equipment. In many circumstances, it is necessary to determine the subjective effects of new transmission equipment or modification to the transmission characteristics of a telephone network. Subjective assessment techniques are required to (a) investigate the performance of new and unknown services and systems where no objective evaluation procedure is known yet and (b) form the basis of objective evaluation and measurement procedures. Every objective evaluation method has to be validated based on results achieved from subjective assessment of the impairment the objective method is developed for. Therefore new services and systems typically have to be evaluated subjectively first before an objective testing procedure can be developed and applied.

2 • Subjective Assessment

The basic description of the subjective assessment of speech quality in telecommunication is ITU-T Recommendation P.800 [24]. The methods described here are intended to be generally applicable whatever forms of degradation factors are present. Further recommendations [25–28] describe more specialized tests for specific processing techniques and devices used in modern networks such as speech codecs, speech echo cancellers, hands-free equipment and noise reduction techniques.

2.1 Opinion Scales Used for Subjective Performance Evaluation

Subjective testing requires exactly defined scaling of the scores derived from the subjects participating in subjective experiments. The scale needs to be easily understandable by the subjects, unambiguous and widely accepted. The layout and wording of opinion scales, as seen by subjects in experiments, is very important. Different ways of scaling and different types of scales may be used [32,34]; their individual benefits and drawbacks are not discussed here, only those most often used in telecommunications are described.

The following opinion scale is the most frequently used one for telecommunication applications, and equivalent wording is to be used in different languages. This is a category rating obtained from each subject at the end of each subjective experiment [24,25]:

Opinion of (the overall quality, the listening speech quality, etc.) of the connection you have just been using:

1 – Excellent
2 – Good
3 – Fair
4 – Poor
5 – Bad

If the subjects' opinion is asked for in conversational tests, the result is termed *mean opinion score, conversational* (MOSc). If the scale is used for speech quality rating in listening tests then the result is called *mean listening quality-opinion score* (MOS). Such tests are named *absolute category rating* (ACR).

Often, especially in conversational tests, a binary response is obtained from each subject at the end of each conversation, asking for difficulty in talking or hearing over the connection used. The experimenter allocates the score "Yes" or "No". If a subject experienced difficulties, the experimenter may carefully ask for further information about the type of difficulties experienced by the subject.

If a degradation introduced to a transmission system is to be evaluated, degradation category rating (DCR) tests are required. The scale used is [24]

5 – Degradation is inaudible.
4 – Degradation is audible but not annoying.
3 – Degradation is slightly annoying.
2 – Degradation is annoying
1 – Degradation is very annoying.

The quantity evaluated from the scores is termed *degradation mean opinion score* (DMOS).

Sometimes, when evaluating higher quality systems, comparison rating is used. The scale used for *comparison category rating* (CCR) tests is given as follows [24]:

"The quality of the second system compared to the quality of the first one is –"

3 – Much better
2 – Better
1 – Slightly better
0 – About the same
−1 – Slightly worse
−2 – Worse
−3 – Much worse

Other scales are also used; additional information can be found in [24–27].

2.2 Conversation Tests

Conversation tests best represent typical communication: both conversational partners exchange their information bi-directionally, acting as both talker and listener. The conversation test is the most natural test used in telecommunication and allows the subjects to behave very similarly to a real conversation. The naturalness which can be achieved in a conversation test depends highly on the task chosen within the test. The task should

- Be easy to perform for each subject
- Avoid emotional engagement of the test subjects
- Be mostly symmetrical (the contribution of each test subject to the conversation should be similar)
- Be independent of the individual personal temperament (the task must stimulate people with low interest in talking and must reduce the engagement of people who always like to talk)
- Reflect a typical conversation

Typical tasks used in telecommunication are the so-called "*Kandinsky tests*" [26,27] or "*short conversational tests*" [27,33].

In the "Kandinsky test", the subjects are asked to describe to their partner the position of a set of numbers on a picture. Both subjects have similar pictures, but with some of the numbers in different positions. The pictures used should have simple figures and objects which are easy to describe. This can be achieved by using pictures consisting of colored, geometrical figures (e.g., paintings by Kandinsky or others, see Figure 1).

In the "short conversational tests", the test subjects are given a task to be conducted at the telephone similar to a daily-life

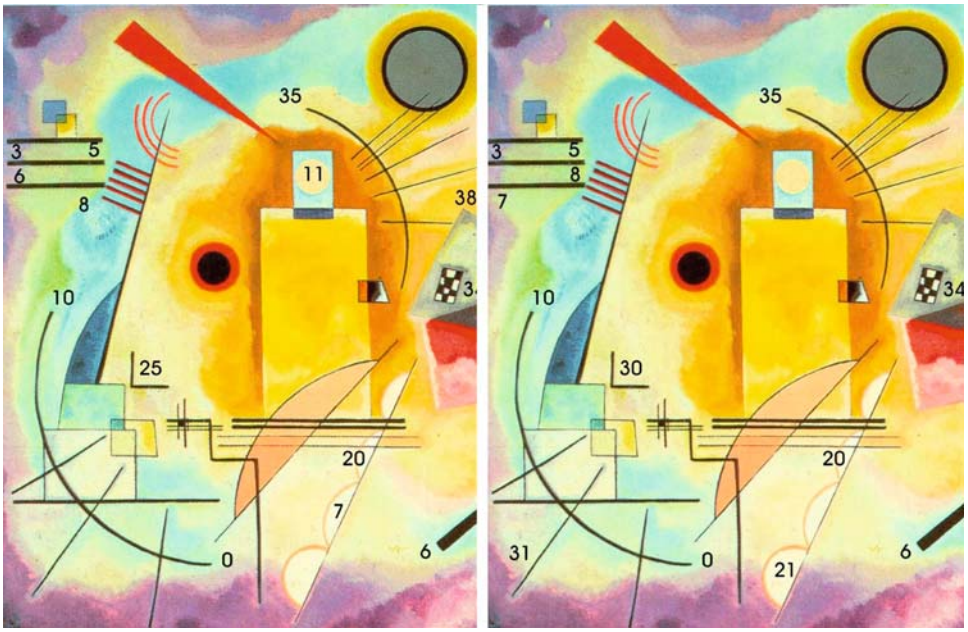


FIGURE 1 Two pictures used by two subjects for conversational tests, "Kandinsky" pictures with different numbering of the different geometrical figures.

situation. Ordering a specific pizza at a pizza service and finding a specific railway connection are examples of typical tasks.

Many conversational tests have been carried out successfully with observers (operators) present in the test rooms together with the subjects. It is their task to register and list all comments which the subjects mention during or after test. This can be useful for further analysis.

In a conversation test, a sufficient number of speakers (voices) should be used in order to minimize talker/speaker-dependent effects.

Subjects mostly rate their opinion about the overall quality of a connection based on the ACR scale. Furthermore, they are asked about difficulties in talking or listening during the conversation. Careful investigation of the nature of these difficulties may require more specialized tests than described below.

2.3 Double Talk Tests

Due to the interactive nature of a conversation, the ability to interact in all conversational situations and especially to interact during double talk with no or only minor impairments is of vital importance. Since many technical implementations may impair the speech quality during double talk, specific double talk tests may help to evaluate the system performance under such conditions. The benefit of this double talk testing method [27] is that it is designed especially for quality assessment during double talk periods and that the test duration is very short. Therefore, double talk tests are very efficient in evaluating this very important quality aspect. Because of the short duration for single tests, double talk tests efficiently determine the influence of the relevant parameters perceived subjectively in detail.

Comparable to the conversational test, the double talk test involves two parties. The purpose of double talk testing tends to be different depending on whether experienced or untrained subjects are used. Untrained subjects are used when it is important to get an indication of how the general telephone-using population would rate the double talk performance of a hands-free telephone, for example. The test procedure is sensitive enough that untrained subjects can assess the relevant parameters even during sophisticated double talk situations. Experienced subjects are used in situations where it is necessary to obtain information about the subjective effects of individual degradations.

During double talk tests, two subjects have a written text in front of them. The texts differ slightly. Subject 1 (talking continuously) starts reading the text. It consists of simple, short and meaningful sentences. Subject 2 (double talk) follows the text

(without reading aloud) and starts reading his text simultaneously at a clearly defined point.

Typical questions within a double talk test are about the dialog capability, the completeness of the speech transmission during double talk, echo and clipping during double talk.

The scales used are typically ACR or DCR scales, but other scales can be used as well. More information can be found in [47].

A drawback of double talk tests is that the testing method is more artificial compared to conversational tests. Even if the text is very simple, the subjects have to concentrate in a different way compared to a free conversation.

2.4 Talking and Listening Tests

In all connections where delay is inserted (e.g., network delay, delay due to signal processing or delay due to packetization) talking-related disturbances like echo or background noise modulation are of importance. In order to investigate such types of impairments in more detail, talking and listening tests can be used. Such tests are mainly used to investigate the performance of speech echo cancellers (EC). All aspects of EC function that influence the transmission quality for subscribers while they are either talking and listening or only listening (without having a conversational partner on the other end of the connection) are covered by this procedure.

A typical test setup is shown in Figure 2.

This measurement setup shows an echo canceller on one side of the connection which is the one under test. From the subject's point of view, this is the far-end echo canceller. The far-end subscriber is simulated by an artificial head. If double talk sequences are required, then the artificial mouth is used to produce exactly defined double talk sequences. Environmental conditions will depend on the characteristics of the test room used, the background noise and other factors. Generally the test procedure is separated into two parts. The first part is designed to examine initial convergence of an echo canceller and the second part has to measure the performance during steady-state conditions. Which of these procedures is used will depend on the purpose of the test. Subjects are first instructed on how to conduct the test.

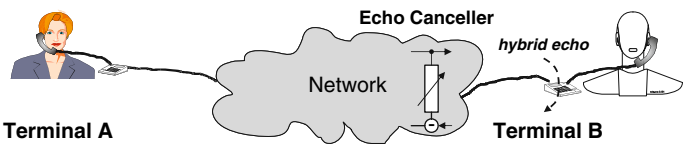


FIGURE 2 Test setup for a talking and listening test.

When testing the initial convergence of a speech echo canceller, subjects answer an incoming telephone call with the same greeting, for example, "[company], [name], [greeting]". After the greeting, the call is terminated, and subjects give their rating.

When testing "steady-state conditions", the echo canceller should first be fully converged. Subjects are asked to perform a task, such as to describe the position of given numbers in a picture similar to the "Kandinsky" test procedure described for the conversational tests. The artificial head can be used to generate double talk at defined points in time in order to introduce disturbing signal components for the speech echo canceller and test the canceller's ability to handle double talk. After the termination of the call, the subjects are asked to give a rating.

The scales used are typically ACR or DCR scales, but other scales can be used as well. More information can be found in [25].

2.5 Listening-Only Tests and Third Party Listening Tests

Listening-only tests are most commonly used in telecommunication. The main purpose of listening-only tests (LOT) is the evaluation of impairments under well-defined and reproducible conditions. When used to evaluate parameters influencing the speech quality in telecommunications, it should be noted that listening tests are very artificial. Listening-only tests are typically used to evaluate the speech sound quality or other listening-related impairments of speech. They are strongly influenced by the selection of the speech material used in the tests; the influence of the test stimuli is much stronger than in conversation tests. Therefore, a very careful test design is required. The test sequences used must represent the typical phoneme distribution of the language used [24]. Furthermore, at least two male and two female speakers should be used for each condition in order to avoid speaker-dependent effects. A sufficient number of presentations must be integrated into a test, ranging from the best to the worst-case condition of the impairment investigated in the test. Reference conditions (simulated, defined impairments) can be integrated in order to check the validity of the test. Additional rules on how to conduct listening-only tests for various purposes are found in [25–28].

In general, listening-only tests are used by presenting pre-recorded, processed speech to the subjects, using reference handset terminals which are operated by the subjects. In third party listening tests [26,27], the speech material is recorded by an artificial head which is used to record the complete acoustical situation including the background. By this procedure, all types of handset, headset and hands-free configurations can be

included into a listening-only test including the environmental conditions at the terminal location. For playback, equalized headphones are used. The equalization must guarantee that during playback the same ear signals are reproduced which were measured during recording. By this procedure, binaural reproduction (for details see [4]) is possible which leads to a close-to-original presentation of the acoustical situation during recording. The third party listening setup allows use of this type of test for investigating conversational situations by third parties. Therefore, a complete conversation is recorded and presented to the listeners. Although the listeners are not talking themselves but listening to other persons' voices, these tests have proven their usefulness in investigating conversational impairments.

The ACR or DCR scales are typically used but other scales, like loudness preference, can be used as well. More information can be found in [26,27].

Instead of ACR or DCR tests, the *comparison category rating* (CCR) can be used. CCR tests are based on paired comparisons of samples. They offer a higher sensitivity and may be used for the quality evaluation of high-quality systems, typically to compare processed and unprocessed speech material.

3 • Objective Methodologies

Besides the typical evaluation procedures used for electroacoustical systems, such as frequency responses, distortions, dynamic ranges and others, a variety of analyses exist for telecommunications. While the well-known procedures for electroacoustical systems are not described here, an overview is given of the objective evaluation procedures specific to telecommunication systems. Table 1 gives an overview of the relationship between impairments perceived subjectively and the correlating objective parameters.

3.1 Loudness

Speech is a time-varying process. Speech sounds are short-term stationary within a time interval of approximately 20 ms. The speech spectrum is continuous. In contrast to loudness calculations (such as are described in ISO R 532 B [33], which is based on the work of Zwicker [38] and others), in telephonometry the concept of *loudness ratings* is used. Although the speech signal is not continuous, useful conclusions can be drawn from experiments with speech and referring them to a continuous spectrum-based model, the loudness ratings concept. This concept is based on the speech signal energy calculated in fixed (frequency) bands and a subjective comparison of these values to the signal energy

TABLE 1 Impairments perceived subjectively and their correlating objective parameters

Impairments perceived subjectively	Description	Correlating objective parameters
Loudness of the speech signal	In send and receive direction, perceived in the listening situation	Loudness ratings in send and receive
Speech sound quality	In send and receive direction, perceived in the listening situation	Frequency response Distortions Objective speech quality measures based on hearing models
Delay	Impairment discovered in conversational situations, especially with high interaction between the conversational partners leading to reduced conversational dynamics and unintended double talk situations	Delay (Time-variant) echo loss TCL (terminal coupling loss) Switching Occurrence of double talk impairments
Echo disturbances (under single talk conditions)	Talking-related disturbance correlated to the user's voice, perceived in the presence of transmission delay	Echo attenuation Echo level fluctuation over time Spectral echo attenuation
Double talk capability	In send and receive direction: Loudness variation between single and double talk periods Loudness variation during double talk Echo disturbances Occurrence of speech gaps Occurrence of artifacts uncorrelated to the transmitted speech signal	Attenuation range Attenuation in send and receive direction during double talk Switching characteristics Minimum activation level to switch over from receive to send direction and from send to receive direction Echo attenuation during double talk Spectral and time-dependent echo characteristics
Quality of background noise transmission	Impairments introduced to the background noise signal in the send direction during Idle mode Far-end speech active in the presence of background noise Near end speech active in the presence of background noise	Minimum activation level in send direction Comfort noise spectral and level adaptation Level variation due to noise reduction systems Attenuation in send direction Sensitivity of background noise detection (activation level, absolute level, level fluctuations)

(*Continued*)

TABLE 1 (Continued)

Impairments perceived subjectively	Description	Correlating objective parameters
		Attenuation introducing noise modulation Switching characteristics D-value (sensitivity for background noise compared to the speech signal sensitivity)
(System) noise	In send and receive direction, perceived during silent intervals	Noise level Level fluctuations Spectral characteristics

of a reference system, the IRS [11,21,35], when adjusting the perceived loudness of the unknown system to the reference system. The different frequency-dependent contributions to the overall loudness impression are reflected by a weighting factor B':

$$B' = \frac{(\Delta l)_c}{(\Delta f)_c} \cdot \frac{1}{l_{max}} = \frac{0.0156}{10^{K_c/10}},$$

where K_c is the critical bandwidth expressed in dB, $(\Delta l)_c$ is the loudness contribution in the frequency band $(\Delta f)_c$.

The total loudness impression is proportional to the integral over $Q(Z) \times B'$ where $Q(Z)$ is the loudness growth function depending on the speech level Z measured in dB [35]. $Q(Z)$ can be approximated by

$$Q(Z) = \text{const.} \cdot (10^{Z/10})^m,$$

where a single mean value for m can be determined by auditory experiments depending on the interest on the range of Z. Z is defined as

$$Z = \beta'_S - L_{ME} - (\beta_0 - K_C),$$

β'_S	long-term power density spectrum of speech referred to the mouth reference point MRP [20]
β_0	pure-tone threshold of hearing
K_χ	width of the critical band frequencies
$(\beta_0 - K_\chi)$	can be treated as the perception threshold of sounds with a continuous spectrum
L_{ME}	mouth-to-ear transmission loss (between the reference points β'_S and β_0)

From these terms, with

$$Z_{R0} = \beta'_S - (\beta_0 - K_C)$$

the loudness for an unknown object (λ_U) can be determined as

$$\lambda_U = \text{const.} \int_0^\infty 10^{(-m/10)\cdot L_{UME}} \cdot 10^{(m/10)\cdot Z_{R0}} B' \mathrm{d}f$$

and can be compared to the loudness of the reference system (λ_R):

$$\lambda_R = \text{const.} \int_0^\infty 10^{(-m/10)\cdot L_{RME}} \cdot 10^{(m/10)\cdot Z_{R0}} B' \mathrm{d}f.$$

The *loudness rating* is defined as the loudness loss Δx, removed from the speech path of the unknown system to achieve $\lambda_U = \lambda_R$, independent of frequency:

$$\textit{Loudness Rating} = \Delta x = \overline{L_{UME}} - \overline{L_{RME}}.$$

For calculation, the frequency band of interest is divided into N bands (20 bands approximately spaced in one-third octaves from 100 Hz to 8 kHz).

Using the substitution

$$G = [10^{(m/10)Z_{R0}} \mathrm{B}']$$

and using the values G for each frequency band based on the subjective loudness experiment conducted in the CCITT laboratory, the tabulated data for the frequency weights (w-weights) are derived. Based on this, the weighted average mouth-to-ear loss of the speech path for an unknown object can be written as

$$\overline{L_{UME}} = -\frac{10}{m} \log_{10} \sum_{i=1}^{N} 10^{(-m/10)(L_{UME,i} + w_i)}$$

with the values for w_i found in ITU-T Recommendation P.79 [21]. The loudness rating is then defined as

$$\textit{Loudness Rating} = -\frac{10}{m} \log_{10} \sum_{i=1}^{N} 10^{(-m/10)(L_{UME,i} - \overline{L_{RME}} + w_i)}.$$

For a given telephone or transmission system, the values of L_{UME} can be derived from the measurement of different sensitivities $S_{M\vartheta}$ (mouth-to-junction) for calculation of the SLR (sending loundness rating), from $S_{\vartheta E}$ (junction-to-ear) for the calculation of the RLR (receiving loundness rating) or from S_{ME} (mouth-to-ear) for the overall loudness rating, OLR.

In a similar manner, the sidetone paths can be described: STMR is the sidetone masking rating describing the perceived loudness of the user's own voice and LSTR is the listener sidetone rating describing the perceived loudness of room noise coupled to the user's ear.

3.2 Speech Sound Quality – Listening Speech Quality

Speech sound quality is influenced by many parameters: frequency response, loudness rating, non-linear distortions, switching, noise and others. Since any degradation of the speech sound quality is perceived mostly in the listening situation, the main focus of speech sound quality is during one-way transmission (single talk condition).

Due to the use of different types of speech coders and the impairments introduced by the transmission network (packet loss, jitter, interference), the listening speech quality may be degraded. The type of distortion added to the signal is typically uncorrelated with the speech signal. This effect cannot be assessed by loudness ratings, frequency response or traditional distortion measurements. The assessment methods used under these conditions are hearing model-based comparison techniques. The principle of these assessment methods is shown in Figure 3.

All methods [2,3,36] compare an original signal $x(t)$ with a degraded signal $y(t)$ that is the result of passing $x(t)$ through a communication system. The output is a prediction of the perceived quality that would be given to $y(t)$ by subjects in a

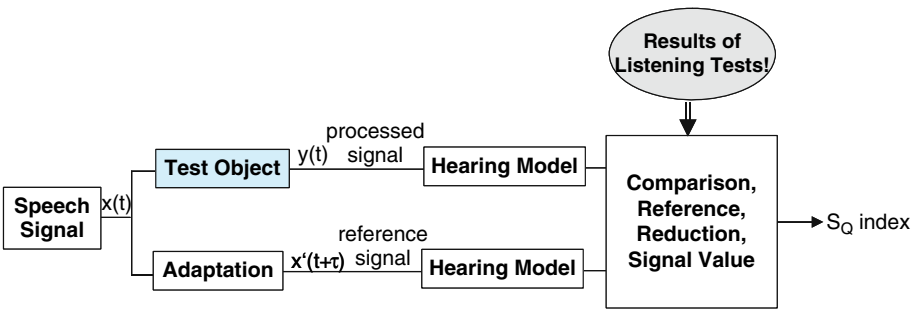


FIGURE 3 Principle of reference signal-based objective speech quality assessment procedures.

subjective listening test. In a first step the signal $x(t)$ is adapted to the delay and the bandwidth limitation introduced by the transmission system. The time alignment between original signal and transmitted signal is made in blocks given by the objective method in order to take care of time-varying delays in transmission systems. It is assumed that the delay within a block is constant.

Based on the time-aligned signals $x'(t + \tau)$ and $y(t)$, the comparison is made. The key to this process is transformation of both the original and the degraded signals to an internal representation that is comparable to the psychophysical representation of audio signals in the human auditory system, taking into account the frequency resolution and the loudness sensation of the human ear.

Almost all of the procedures provide level alignment to a calibrated listening level, time–frequency mapping, frequency warping and loudness scaling. Since not all differences found in the representations of $x'(t)$ and $y(t)$ contribute to the perceived degradation of the processed signals minor steady-state differences between original and degraded signals are compensated and not taken into account for the comparison. More severe effects, like rapid (level) variations, are typically only partially compensated so that a residual effect remains and contributes to the overall perceptual disturbance. Segments with missing signal energy in the processed signal are weighted less than any components added to the processed signal $y(t)$ (asymmetry weighting). By this procedure a small number of quality indicators are used to model the most important subjective effects.

The differences at the hearing model outputs are computed in a cognitive model and combined to an internal objective listening quality score. By appropriate mapping, this internal score is adjusted to the MOS results of the subjective listening experiments (MOS-LQS, mean opinion score-listening quality subjective [23]) which are conducted using the same types of impairments which the objective speech quality assessment method is intended to be used for. In order to distinguish this objective result from the subjective MOS-LQS results, this score is named MOS-LQO (mean opinion score-listening quality objective [23]).

The ITU-T recommended objective speech quality evaluation procedure is PESQ, as described in P.862 [29]. As with many other methods, this procedure is suitable for the electrical access of speech signals, such as using interfaces provided at the network termination points. The influence of terminals, such as degradations of the speech signals introduced by non-ideal

frequency responses, transducer distortions, etc., is not modeled. TOSQA2001 [3,6] allows the acoustical access and takes into account the impairments introduced by the acoustical components. In combination with a proper acoustical coupling, taking into account the behavior of a terminal when used by a human operator, a realistic speech quality assessment can be made. In order to achieve realistic conditions, an artificial head (HATS – head and torso simulator, [17,18]) equipped with an artificial mouth and a type 3.4 artificial ear is used for assessing the acoustical input and output signals of a terminal.

A general drawback of all hearing model-based approaches is the assumption of the same listening level in a subjective listening test under all conditions. The standard listening level assumed is 79 dB_{SPL} (dB re 20 μPa). For higher- or lower-level acoustical output signals, deviating from the standard listening level, the predicted MOS-LQO value may be incorrect.

3.3 Delay and Echo Loss

Delay is introduced by any communication system. Algorithmic delay may be introduced by signal processing in the terminal, such as background noise reduction (noise cancellation), echo cancellation, speech coding, voice-controlled attenuation and others. Packetization, as used in VoIP systems, may further contribute to the delay introduced in a connection. The main effect of delay is the impairment of the conversational dynamics. For highly interactive conversational tasks a delay of not more than 150 ms is recommended [14], although more recent investigations [31] indicate that in typical conversations higher delays may be acceptable. Another effect of delay is the increased sensitivity of the human ear for echo components with increasing delay [15]. The more delay in a connection, the higher the importance of the echo loss between the acoustical receiver and the terminal microphone. Any signal component received by the terminal and transmitted back to the far-end user may be perceived as echo by the far-end user. A further source of echo may be insufficient trans-hybrid loss in analog networks, which is not properly compensated by the network speech echo cancellers. From the perception point of view the actual echo source is of minor importance. The requirement for the echo attenuation depending on the delay inserted in a connection is described in [15]. The echo loss has to be provided during the entire call without any temporal change. Testing techniques useful for evaluating the convergence behavior of echo cancellers in the networks can be found in [16], those for hands-free terminals in [22].

3.4 Double Talk Performance

The double talk performance of a hands-free terminal and the resulting requirements have been investigated in great detail (see, e.g., [33]). These investigations led to a classification system which allows to assess the performance of a system during double talk objectively by a set of parameters [15,22]. Although the subjective tests were mainly based on the evaluation of hands-free systems, the results can be used for other terminals and network equipment as well, since the impairments are mostly perceived by the far-end user. From the perception point, the source of the impairments (terminal or network component) is not relevant. Table 2 gives an overview of the required performance parameters and their impact on the performance perceived subjectively.

The double talk performance is described by a set of parameters:

- *Attenuation range in send and receive directions*: The attenuation range is the difference between minimum and maximum amplification in one transmission direction. For the measurement, typically a speech-like, time-multiplexed test signal is inserted simultaneously in send and receive direction. The test signal consists of a series of partially overlapping composite source signals [8,19].
- *Talker echo loudness rating/echo attenuation*: The talker echo contributes significantly to the perceived speech quality in the double talk situation. The talker echo assessment during double talk can be made by inserting a speech-like frequency-multiplexed test signal consisting of two voiced sounds with non-overlapping spectra simultaneously in send and receive directions [19]. By applying an appropriate comb filter technique, the echo signal can be extracted from the double talk signal and the echo loss during double talk can be determined.

TABLE 2 Definition of double talk performance according to [22]

	Type 1 (full duplex)	Type 2 (partial duplex)			Type 3 (non-duplex)
		Type 2a	Type 2b	Type 2c	
$TELR_{DT}$ (dB)	≥ 37	≥ 33	≥ 27	≥ 21	< 21
a_{Hsdt} (dB)	≤ 3	≤ 6	≤ 9	≤ 12	> 12
a_{Hrdt} (dB)	≤ 3	≤ 5	≤ 8	≤ 10	> 10

$TELR_{DT}$, talker echo loudness rating during double talk; a_{Hsdt}, attenuation range in send direction during double talk; a_{Hrdt}, attenuation range in receive direction during double talk

- *Switching characteristics, Switching times*: The measurement and optimization of system's gain switching and switching times are required for all systems which are not fully duplex-operable. Such systems typically show a switching behavior, sometimes a certain (speech) level is required to activate the system. The switching characteristics can be evaluated basically in a similar manner as the evaluation of the attenuation range. The definitions of switching times, etc., can be found in, e.g., [22].

Further characterization may be useful depending on the type of duplex capability of a communication system:

- Minimum activation level to switch over from receive to send direction and from send to receive direction
- Frequency responses in send and receive directions during double talk
- Loudness ratings in send and receive directions during double talk

Details can be found in [9,22].

3.5 Background Noise Performance

In background noise situations, the perceived speech quality is influenced by a variety of parameters. The human listener focuses on two signals: the speech signal and the background noise signal. The quality judgment is made based on the perception of the transmitted background noise both during periods where no speech signal is present and during periods with a speech signal. Furthermore, the perception is based on the quality of the transmitted speech signal. Therefore, an objective method describing the perceived quality must take into account these different quality dimensions. Initial subjective experiments conducted with expert listeners indicate that at least two groups of "typical listeners" seem to exist: The first group is very sensitive to any impairments introduced to the speech signal but is not sensitive against the background noise level, the other group is less sensitive against any impairments added to the speech signal but prefers a maximum of noise reduction. A complete background noise evaluation method does not exist yet, but for the different situations instrumental procedures can be used.

When focusing purely on the background noise transmission without speech signal present simultaneously, a level variation analysis can be used to analyze the transmitted signal. Typically the measured output signal level is compared to the input signal level and the level difference is analyzed. This method can only be applied if the reference (input signal) can be assessed. Preliminary auditory investigations indicate that a level variation of

±3 dB should not be exceeded. More information about the transmission characteristics of background noise, however, can be found when comparing output and input signals based on a spectral analysis. The output/input signals are analyzed spectrographically and displayed as a spectral difference between input and output. Again, a reference signal is needed to perform the analysis. Once more, there is some indication that variations in time and/or frequency should not exceed ±3 dB. The disadvantage of such simple spectral difference methods is the non-existent relationship to the human ear signal processing. There is a high probability of getting misleading results due to poorly adapted frequency resolution, not taking into account the masking effects in time and frequency, disregarding the non-linearity of the human ear signal processing and others.

Another possibility is the use of the psychoacoustically motivated "relative approach" [7] which seems to be more promising. The basis for the analysis is a hearing model according to [37]. In contrast to all other methods, the "relative approach" does not use any reference signal in its present form. The non-linear relationship between sound pressure level and auditory perceived loudness is taken into account by time/frequency warping in a Bark filter bank and proper integration of the individual outputs. The filter bank is realized in the time domain. The output signals of the filter bank are rectified and integrated, thus generating the envelope. The three-dimensional output of the hearing model is the basis for the relative approach. In each critical band the long-term level (integration time 2–4 s) is compared to the short-term level (2 ms).

An overall value can be derived for example by applying the following equation [7]:

$$Q = f(N,S) + f\left(\sum_{i=1}^{24}\left[\,|F_G(i-1) - F_G(i)\,|\cdot w_1(i,F_G(i)) + \sum_{n=1}^{T}|F_G(i,n) - F_G(i,n+1)\,|\cdot w_2(i,F_G(i))\right]\right),$$

where $F_G\ (i)$ is a mean value of the critical band level over a period T of 2–4 s, $F_G\ (0) = F_G\ (1)$, $F_G\ (i,\ n)$ is a mean value of the critical band level over a much shorter period (approx. 2 ms), n is the current (time-dependent) value. The weighting factors $w_1(i,\ F_G\ (i))$, $w_2(i,\ F_G\ (i))$ depend on the critical band level $F_G\ (i)$. In addition, the overall value is influenced by the function *f (N,S)* which describes an auditory factor, dependent on loudness N and sharpness S. Figure 4 shows the block diagram including the

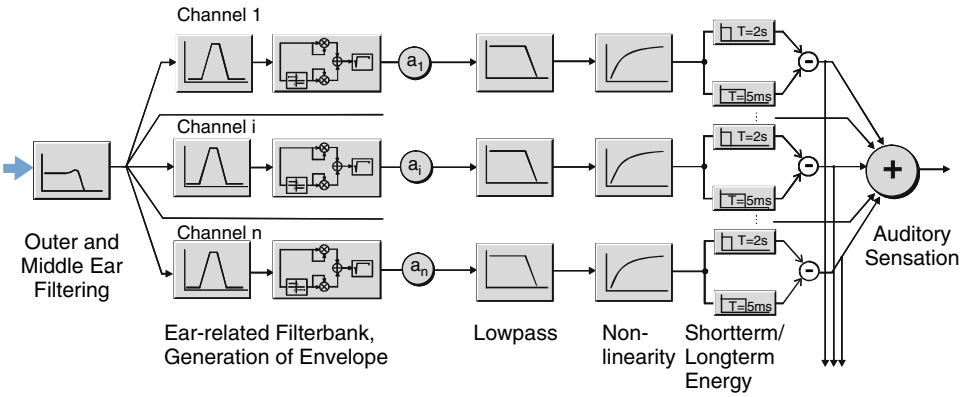


FIGURE 4 Model of the "relative approach": hearing model and calculation of the energy differences in critical bands.

hearing model [37] and the calculation of the relative approach output signals in the Bark domain.

The method can be applied for short-term stationary background noise signals, such as car noise during constant driving conditions, or office noise with low dynamic variations.

When focusing on the quality of the transmitted speech signal, all objective methods proposed so far have some disadvantages. *S/N*-based methods, as described in [1], do not take into account the perceptual effects of the human ear and are limited in their application to specific types of noise reduction algorithms. More advanced methods like PESQ [29] or TOSQA [3] take into account the human ear signal processing; however, it is known [3,5] that such technologies do not yield sufficient results in conditions with *S/N* ratios of less than about 20 dB.

4 • Network Planning

The state-of-the-art planning model in today's communication is the "*E-model*" [12]. The E-model is intended for assessing the combined effects of variations in several transmission parameters that determine conversational quality (except double talk and background noise transmission) of 3.1 kHz handset telephony. The E-model is based on the OPINE model [30] which first introduced the concept of impairment factors. The typical application of the E-model is to help planners to ensure that users will be satisfied with end-to-end transmission performance.

The primary output of the model is the "*rating factor*" *R*; but this can be transformed to give estimates of customer opinion. They can be used for transmission planning purposes, but are

not suitable for a customer opinion prediction, since the E-model has not been fully verified. Nevertheless, for many combinations important to transmission planners, the E-model can be used with confidence. Time-variant behavior of transmission systems, however, is not modeled.

The E-model splits a connection into a send side and a receive side. It estimates the conversational quality from mouth to ear as perceived by the user at the receive side, both as listener and talker. The basic assumption of the E-model is the additivity theorem: all impairments are additive on a psychological scale. Based on this assumption a rating factor (R) is calculated:

$$R = R_{\mathrm{o}} - I_{\mathrm{s}} - I_{\mathrm{d}} - I_{\mathrm{e,eff}} + A.$$

R_{o} is the basic signal-to-noise ratio which all impairments are subtracted from. I_{s} (simultaneous impairment factor) describes all impairments simultaneous to the speech signal. I_{d} (delayed impairment factor) summarizes all impairments caused by delays in a connection. $I_{\mathrm{e,eff}}$ describes the impairments introduced by low-bit-rate codecs. For the different codecs used in telecommunication, the I_{e} values can be derived from tables found in [13]. The perceived listening speech quality in the case of packet loss for the different types of speech coders can be modeled by the packet-loss robustness factor *Bpl* in combination with the packet-loss probability *Ppl* [12,13]. In this situation the following formula is applicable:

$$I_{\mathrm{e,eff}} = I_{\mathrm{e}} + (95 - I_{\mathrm{e}}) \cdot \frac{\mathrm{Ppl}}{\mathrm{Ppl} + \mathrm{Bpl}}.$$

The advantage factor A allows for compensation of impairment factors when there are other advantages of access to the user which are not covered by the E-model.

The calculation of R_{o} can be found in [12].

I_{s} is defined as

$$I_{\mathrm{s}} = I_{\mathrm{olr}} + I_{\mathrm{st}} + I_{\mathrm{q}}.$$

I_{olr} represents the decrease in quality caused by too-low values of OLR (overall loudness rating), I_{st} represents the impairment caused by non-optimum sidetone and I_{q} represents impairment caused by quantizing distortion. The detailed description is found in [12].

I_{d} is defined as

$$I_{\mathrm{d}} = I_{\mathrm{dte}} + I_{\mathrm{dle}} + I_{\mathrm{dd}}.$$

I_{dte} represents the impairments due to talker echo, I_{dle} represents impairments due to listener echo and I_{dd} represents the impairment caused by too-long absolute transmission delay. The detailed description is found in [12].

The advantage factor A may be used to model the advantage of a transmission system in special conditions or situations (for example, a system used for communication in emergency situations where the reliable function is more important than excellent speech quality). There is no relation to all other transmission parameters and typically the factor is not used in transmission planning.

The transmission rating for a narrowband transmission system with very high quality is 93.2 (close to a maximum of 100).

If required, an MOS value can be calculated from the R value in order to better interpret the output of the E-model as an estimated average subjective rating. The following equations are used for mapping:

$$\begin{aligned} &\text{For } R < 0: && \text{MOS} = 1 \\ &\text{For } 0 < R < 100: && \text{MOS} = 1 + 0.035R + R(R-60)(100-R)7 \cdot 10^{-6} \\ &\text{For } R > 100: && \text{MOS} = 4.5 \end{aligned}$$

References

1. 3GPP TS 26.077: 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Minimum Performance Requirements for Noise Suppresser; Application to the Adaptive Multi-Rate (AMR) Speech Encoder, 3GPP, 2003
2. Beerends, J.G., Stemerdink J.A.: A Perceptual Speech-Quality Measure Based on a Psychoacoustic Sound Representation, J. Audio Eng. Soc., Vol. 42, No. 3, pp. 115–123, March 1994
3. Berger, J.: Instrumentelle Verfahren zur Sprachqualitätsschätzung – Modelle auditiver Tests, PhD thesis, Kiel, 1998
4. Blauert, J.: Spatial Hearing: The Psychophysics of Human Sound Localization, MIT Press, Cambridge, MA, 1997
5. Dreiseitel, P.: Quality Measures for Single Channel Speech Enhancement Algorithms. 7th International Workshop on Acoustic Echo and Noise Control, 10–13 Sept. 2001, Darmstadt
6. ETSI EG 201 377-2: Specification and Measurement of Speech Transmission Quality; Part 2: Mouth-to-Ear Speech Transmission Quality including Terminals, 2004
7. Genuit, K.: Objective Evaluation of Acoustic Quality Based on a Relative Approach. Internoise '96, 1996, Liverpool, UK
8. Gierlich, H.W.: A Measurement Technique to Determine the Transfer Characteristics of Hands-Free Telephones, Signal Process., Vol. 27, No. 3, pp. 281–300, 1992
9. Gierlich, H.W.: The Auditory Perceived Quality of Hands-Free Telephones: Auditory Judgements, Instrumental Measurements

and Their Relationship, Speech Commun., Vol. 20, p. 241–254, October 1996
10. ISO R 532 B: Acoustics – Method for Calculating Loudness Level, International Organization for Standardization, Geneva, 1975
11. ITU-T: ITU-T Handbook on Telephonometry, ITU-T, Geneva, 1992
12. ITU-T Recommendation G.107: The E-Model, a Computational Model for Use in Transmission Planning, International Telecommunication Union, Geneva, 2003
13. ITU-T Recommendation G.113: Transmission Impairments due to Speech Processing, International Telecommunication Union, Geneva, 2001
14. ITU-T Recommendation G.114: One-Way Transmission Time, International Telecommunication Union, Geneva, 2000
15. ITU-T Recommendation G.131: Control of Talker Echo, International Telecommunication Union, Geneva, 2000
16. ITU-T Recommendation G.168: Digital Network Echo Cancellers, International Telecommunication Union, Geneva, 2004
17. ITU-T Recommendation P.57: Artificial ears, International Telecommunication Union, Geneva, 2002
18. ITU-T Recommendation P.58: Head and Torso Simulator for Telephonometry, International Telecommunication Union, Geneva, 1996
19. ITU-T Recommendation P.501: Test Signals for Use in Telephonometry, International Telecommunication Union, Geneva, 2000
20. ITU-T Recommendation P.64: Determination of the Sensitivity/Frequency Responses of a Local Telephone Systems, International Telecommunication Union, Geneva, 1999
21. ITU-T Recommendation P.79: Calculation of Loudness Ratings for Telephone Sets, International Telecommunication Union, Geneva, 2000
22. ITU-T Recommendation P.340: Transmission Characteristics and Speech Quality Parameters of Hands-Free Telephones, International Telecommunication Union, Geneva, 2000
23. ITU-T Recommendation P.800.1: Mean Opinion Score (MOS Terminology), International Telecommunication Union, Geneva, 2003
24. ITU-T Recommendation P.800: Methods for Subjective Determination of Speech Quality, International Telecommunication Union, Geneva, 2003
25. ITU-T Recommendation P.830: Subjective Performance Assessment of Telephone-Band and Wideband Digital Codes, International Telecommunication Union, Geneva, 1996
26. ITU-T Recommendation P.831: Subjective Performance Evaluation of Network Echo Cancellers, International Telecommunication Union, Geneva, 1998
27. ITU-T Recommendation P.832: Subjective Performance Evaluation of Hands-Free Terminals, International Telecommunication Union, Geneva, 2000

28. ITU-T Recommendation P.835: Subjective Performance of Noise Suppression Algorithms, International Telecommunication Union, Geneva, 2003
29. ITU-T Recommendation P.862: Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs, International Telecommunication Union, Geneva, 2001
30. ITU-T Supplement 3 to ITU-T Series P Recommendations: Models for Predicting Transmission Quality from Objective Measurements, International Telecommunication Union, Geneva, 1988
31. ITU-T SG12: Delayed Contribution D.45: Report on a New Subjective Test on the Relationships Between Listening, Talking and Conversational Qualities when Facing Delay and Echo, France Telecom, Geneva, 18–27 January 2005
32. Jekosch, U.: Voice and Speech Quality Perception, Springer, Berlin, 2005, ISBN3-540-24095-0
33. Kettler, F., Gierlich, H.W., Diedrich, E.: Echo and Speech Level Variations During Double Talk Influencing Hands-Free Telephones Transmission Quality. IWAENC '99, 27–30 Sept. 1999, Pocono Manor, USA
34. Möller, S.: Assessment and Prediction of Speech Quality in Telecommunications, Kluwer Academic, New York, 2000, ISBN 0-7923-7894-6
35. Richards, D.L.: Telecommunication by Speech, Wiley, New York, 1973, ISBN 0-470-71949-4
36. Rix, A.W., Reynolds R., Hollier M.P.: Perceptual Measurement of End-to-End Speech Quality over Audio and Packet-Based Networks. 106th AES Convention, pre-print no. 4873, May 1999
37. Sottek, R.: Modelle zur Signalverarbeitung im menschlichen Gehör, PhD thesis, RWTH Aachen, 1993
38. VDA: VDA-Specification for Car Hands-Free Terminals, Version 1.5, VDA, 2005

45

Efficient Speech Coding and Transmission Over Noisy Channels

Ulrich Heute[1] *and Norbert Goertz*[2]

[1] *Inst. for Circuit and System Theory, Faculty of Engineering, Christian-Albrecht University, Kaiserstr. 2, D-24143 Kiel, Germany*
[2] *Inst. for Digital Communications, School of Engineering and Electronics, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JL, UK*

1 • Introduction and Fundamentals

1.1 Goals

"Efficient speech coding" addresses the digital representation of discrete-time acoustic signals by few bits. "Transmission over noisy channels" may corrupt these bits. Error correction and, if imperfect, concealment are needed.

1.2 Signal Digitization and Efficient Coding

Bandwidth and sampling: The continuous speech signal $u_0(t)$ is sampled at instances $t_k = kT_s, k \in \mathbb{Z}$. The sampling frequency $f_s = 1/T_s$ is chosen according to the sampling theorem:

$$f_s \geq 2f_c. \tag{1}$$

"Narrow-band telephone speech" is usually band-limited to

$$f_c < 4\,\text{kHz} \qquad (\text{e.g. } 3.4\,\text{kHz}); \tag{2}$$

"wide-band speech" has an improved naturalness due to a larger value

$$f_c < 8\,\text{kHz} \qquad \text{(e.g. } 7.0\,\text{kHz)}. \tag{3}$$

Quantization and coding: A sample $u_0(kT_s) \doteq u(k)$ is then sorted into the ith of 2^w possible quantization intervals of size Q_i and described by its w-bit binary index i. Globally, the sizes Q_i are proportional to the signal-amplitude range D; usually, D is chosen symmetrically as $\pm D/2$ around $u = 0$ and proportionally to the signals' root-mean-square value σ_u:

$$D = \lambda \cdot \sigma_u \text{ (with, e.g. } \lambda \approx 4). \tag{4}$$

So, halving D allows to have the same interval sizes with one bit less.

In the simplest case of a linear quantizer (or "linear ADC", or "linear *p*ulse-*c*ode *m*odulation, PCM"), the size of all intervals is constant:

$$Q_i \equiv Q = D \cdot 2^{-w}.$$

At a DAC output as well as for further digital processing, all signal values $u(k)$ within an interval indexed by $i = i(k)$ are then represented by the same "representative" $[u(k)]_Q$. The quantization error $[u(k)]_Q - u(k)$ is usually modelled as an additive random signal. For linear PCM, with the representatives chosen as the interval centres, signal-independent, white, uniformly distributed quantization noise is a good model, with zero mean and variance

$$\sigma_Q^2 = Q^2/12. \tag{5}$$

The description of $[u(k)]_Q$ by an index $i(k)$ is a digital coding step. "Coding" is also often used, however, to address error protection of $i(k)$. This will be addressed separately (see Section 1.4). At this point, "efficient coding" denotes the replacement of the w-bit indices $i(k)$ by other, linearly or non-linearly "transformed" values $x(k)$ such that the bit rate f_B needed to reconstruct $\tilde{u}(k) \approx u(k)$ at a prescribed quality is much smaller than that needed for a linear PCM. So, "coding" here refers to *data compression*.

A common (though not generally "valid") quality measure is the *s*ignal-to-*n*oise (power) *r*atio, defined logarithmically as

$$\text{SNR/dB} = 10 \cdot \lg[\sigma_u^2/\sigma_n^2] = 20 \cdot \lg[\sigma_u/\sigma_n], \tag{6}$$

with σ_n^2 denoting the general distortion variance. We got used to

$$\mathrm{SNR} \approx 35, \ldots, 40\,\mathrm{dB} \tag{7}$$

in the "classical" telephone network. Equations (1), (2), and (4)–(7), with $\sigma_n \doteq \sigma_Q$, tell us, after some considerations about the speaker-dependent, varying dynamic range D, that $w = 11$ bits are necessary for linear PCM; thereby,

$$f_{\mathrm{B_{PCM}}} \doteq w \cdot f_\mathrm{s} = 11\ \mathrm{bit} \cdot 8000/\mathrm{s} = 88\,\mathrm{kbit/s}. \tag{8}$$

Speech production and perception modelling: Understanding how some information in a speaker's brain is transferred into a speech sound and how this sound is processed by a listener is more than background knowledge: models for both operations are practical tools in efficient speech coding.

Figure 1 shows a simplified production model [62,77]: the "vocal tract" or "acoustic tube" between larynx and lips can be described, under certain assumptions, by an all-pole-filter transfer function of order $n(\approx 10)$:

$$H^{-1}(z) = z^n / \sum_{\nu=0}^{n} \beta_\nu \cdot z^\nu = 1 / \sum_{\nu=0}^{n} \beta_{n-\nu} \cdot z^{-\nu}. \tag{9}$$

The frequency response $H^{-1}(\mathrm{e}^{\mathrm{j}\Omega})$, with the normalized frequency $\Omega \doteq 2\pi \cdot f/f_\mathrm{s}$, is time constant for some $5, \ldots, 10, \ldots, 20, \ldots$ ms. It defines the uttered sound for this time. Especially, the maxima of $|H^{-1}(\mathrm{e}^{\mathrm{j}\Omega})|$, closely related to the poles of $H^{-1}(z)$, are sound-characteristic and therefore termed "formants".

Some information about the sound stems also from the excitation signal $d(k)$ in Figure 1. It can be white noise with

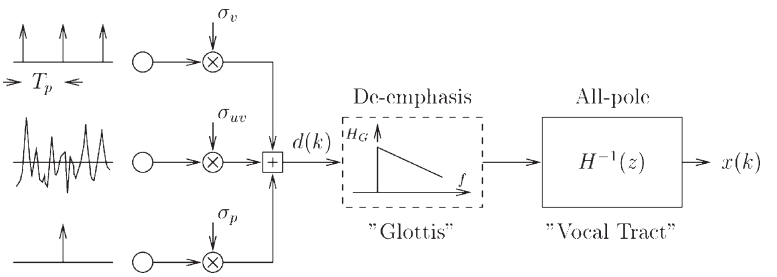


FIGURE 1 Simplified speech-production model with three generators exciting the vocal-tract filter (including a possible "de-emphasis" glottal pre-filter).

a variance σ_{uv}^2 for *u*n*v*oiced sounds, a pulse of size σ_{pl} for *p*losives, a periodic pulse sequence with period T_p and size σ_v for *v*oiced sounds which are periodic with a fundamental (or "*p*itch") frequency $f_p = 1/T_p$, or a combination of these. So, $d(k)$ has, ideally, a constant spectral envelope in all cases. The actually slightly decaying spectra can be taken into account by inserting a fixed de-emphasis-type "glottis pre-filter" $H_G(e^{j\Omega})$. Refinements of this model are discussed in [53].

When a sound reaches a listener, it is first filtered in the outer and middle ear. This causes a spectral shaping and a band-limitation to an upper value $f_c \leq 20\,\text{kHz}$ and a lower boundary $f_L \geq 20\,\text{Hz}$. However, not all signals within these limits are perceived: a frequency-dependent minimum level is needed.

In the inner ear, the signal's frequency components are separated by a bank of (...24...) band-pass filters; the bandwidths grow with a warped frequency axis termed "Bark scale" [77,79]. Non-linear processing follows: depending on both their absolute and relative amplitudes, one of two neighbouring components may be "masked" by the other, i.e. become inaudible. So, beyond the absolute hearing threshold, a signal-dependent masking threshold can be computed. Besides, wide-band but time-dependent maskings of weaker succeeding or preceding sounds are possible. Finally, the resulting excitation of the inner-ear sensors is non-linearly transformed into a perceived loudness.

1.3 Speech Coding and the "Coding Gap"

In the "classical" speech-coding literature, "high-rate" and "low-rate" codings were discriminated: while digital speech transmission or storage was possible with "telephone quality" as in Eqs. (2) and (7) via various schemes between log. PCM at $f_B = 64\,\text{kbit/s}$ (ITU G.711, [43]) and ADPCM at $f_B = 32\,\text{kbit/s}$ (ITU G.726, [45]), really low rates like $f_B = 2.4\,\text{kbit/s}$ were possible, for the same narrow-band speech, only with "synthetic" quality via VOCODER techniques like LPC-10 [75]. Here, no waveform reproduction is aimed at and, therefore, no SNR description can be used appropriately. Early attempts to code wide-band speech according to Eq. (3) somehow "naturally" worked at "high rates", like a split-band ADPCM at $f_B = 64\,\text{kbit/s}$ (ITU G.722, [44], see also [61]). So, there is a "gap" – topic of "lab research" mainly in the 1970s and beginning 1980s, with results invading realization in the 1990s [38].

1.4 Robust Transmission

Bit errors: Up to now, we assumed that the source decoder would receive the bitstream as sent out by the optimized coder. In practice, however, the transmission can be corrupted by noise and interference, perhaps strongly, like in mobile radio. Since

the quality of the speech output is severely deteriorated by bit errors (the lower the data rate, the higher the sensitivity), error protection is needed.

Real channels may vary widely. Errors may happen randomly or can become "bursty". We assume error probabilities which may be very low sometimes, but may go up to, say, $p_e = 0.2$. Only if required for analysis, we will assume errors and noise to be uncorrelated and independent of other signals.

Channel coding: A channel coder computes "channel codewords" from the source-coder output, enlarging the bitstream by redundant bits which are to be used to correct errors. The ratio $R = f_B/f_{total}$ of the source coder and the total bit rates is termed "code rate". As detailed later, error-correcting codes are often concatenated by error-detecting codes, enabling at least the awareness of (perhaps "bursty") disturbances which are too strong to be corrected. This means that, in fact, no ideal error protection is assumed. This is due to realizability restrictions violating the conditions of ideal channel coding, especially that of infinite-length data.

Source and channel coder combination: Realizability also prevents the source coder from being ideal – its output bits are not at all free of redundancy or of equal importance. So, the conditions for applying Shannon's separability theorem are violated altogether, and optimizing a source coder to some best possible average quality at a small bit rate which is then fed into a separately optimized channel coder does not lead to the actual aim: the system should yield the highest possible similarity between input and output signals at a possible *total* bit rate after source *and* channel coding.

2 • Speech Coding

2.1 Origin of the Gap

The rate reduction from that of a simple linear PCM with $f_{B_{PCM}} = 88\,\text{kbit/s}$ as in Eq. (8) over 64 kbit/s down to 32 kbit/s at the same quality uses three essential steps, namely, non-linear quantization, adaptivity and, last but not least, the concept of linear prediction (LP): a non-recursive filter with a transfer function $H_0(z) = \sum_{i=1}^{n} a_i \cdot z^{-i}$ computes a linear combination of past signal values as an estimation $\hat{u}(k) = \sum_{i=1}^{n} a_i \cdot u(k-i)$ of the next sample. The subtraction of the actual sample $u(k)$ yields the difference or residual signal $d(k) = \hat{u}(k) - u(k)$ as the output of a prediction-error filter with the transfer function

$$H(z) = 1 - H_0(z) = 1 - \sum_{i=1}^{n} a_i \cdot z^{-i}. \tag{10}$$

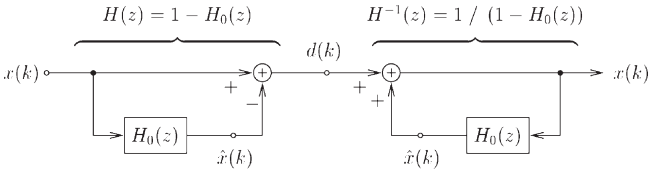


FIGURE 2 Typical structure of a predictive system.

Excitation of $1/H(z)$ by $d(k)$ in the receiver regenerates $u(k)$ (see Figure 2).

The minimization of the residual variance σ_d^2 leads to the optimal predictor-coefficient vector $\mathbf{a} = \mathbf{R}^{-1} \cdot \mathbf{r}_0$, found from the correlation matrix $\mathbf{R}$ and correlation vector $\mathbf{r}_0$. Due to this minimization, $d(k)$ is "smaller" than $u(k)$, and as concluded after Eq. (4), it can be quantized with fewer bits, even considerably less if both prediction and quantizer are adapted with time. This is no heavy algorithmic load, since the above matrix operations can be computed very efficiently, e.g. by the "Levinson–Durbin" recursion [62,77]. But, if the SNR quality is to be kept, $w \geq 4$ is found to be necessary. Together with $f_s = 8\,\text{kHz}$, this limits the rate reduction to $f_B = 32\,\text{kbit/s}$ as mentioned above. On the other hand, large parts of the signal information *do now* reside in the predictor and quantizer parameters, either adapted backwards (ADPCM) in both transmitter and receiver or transmitted as side-information with an additional rate $f_B = \ldots 2 \ldots \text{kbit/s}$ (APC). This, however, is the data rate of the (LPC-) VOCODER [75], synthesizing speech from parameters – with synthetic quality: following the speech-production model of Figure 1, the filter with the transfer function $1/H(z)$ is used to model the vocal tract. This is possible, since it has the same all-pole form as found in Eq. (9); we just have to identify the coefficients in Eqs. (9) and (10) as

$$\beta_{n-i} = -a_i, i \in \{1, \ldots, n\}, \text{ and } \beta_n = 1.$$

Beyond, little information is needed to replace the large residual data volume and control the possible excitations – their type, size, and periodicity.

2.2 Approaches to Close the Gap

RELP coding: If "parameters-only" representations reduce quality, and residual rates cannot be reduced by smaller values of w, f_s should be diminished. This is not possible for speech, because of Eqs. (1) and (2), but for the residual, which is whitened by the (necessarily) decorrelating filter $H(z)$: since $H(z)$ *exploits* the speech correlation to reduce σ_d^2, the correlation remaining in $d(k)$ must be smaller, ideally, a pulse-like autocorrelation; thus a

constant power spectrum should remain – the residuum should become white. This can be interpreted as an inversion of the vocal-tract filtering – $H(z)$ denotes the inverse of the vocal-tract filter, which for this reason was already termed $H^{-1}(z)$ before.

After band-limiting $d(k)$ to $f_c/r, r \in < 001 >$, the remaining (constant!) base band needs only a sampling rate f_s/r; in the receiver, a spectrally widened signal can be applied to $1/H(z)$ in a "baseband residual"-excited LP system (BB-RELP). However, some unnatural, metallic distortions are unavoidable. Two reasons cause these artefacts: an *unflexible down-sampling*, first, which especially does not take care of the *spectral fine structure* second, in $d(k)$, namely its lines at harmonics of the fundamental frequency f_p.

Further redundancy reduction – LTP: Equispaced spectral lines are due to periodicity. They carry redundancy like the spectral shape, now (more or less) removed by $H(z)$. So, they can be removed in a similar way by a second, "pitch predictor", which uses correlation not at distances of $n \cdot T_s (n \approx 10)$, but of $m_p \cdot T_s \approx T_p$ ($m_p \in \{32, 160\}$ for normal voices). This long-time range gives rise to the notation "long-term predictor" (LTP; see Figure 3). The parameter β_p may be found by the same principle used before (see Section 2.1) for determining the coefficients a_i in Eq. (10): If m_p is known, e.g. from some pitch-detection algorithm [37], the variance σ_e^2 of the second-difference signal $e(k)$ is minimized with respect to β_p. However, the search for m_p can also be included into this minimization. Then, m_p may be not a "good" pitch estimate, but it is most suitable in terms of the criterion σ_e^2.

Flexibilities in down-sampling – RPE and MPE: Down- and upsampling of the baseband signal results, for voiced sounds, in spectral lines which are harmonic, i.e. lie at multiples of f_p within the baseband, but are then shifted according to f_s/r, irrespective of f_p.

Instead of taking every rth sample of the residual $d(k)$ in a naive manner, one may try to dissolve the misfit between sampling and pitch frequencies by, e.g. a decimation switching between various factors r with a reasonable mean value $\bar{r}$. However, this is less

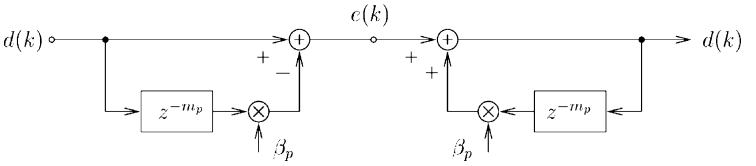


FIGURE 3 Long-term predictor ("1-tap LTP") and its inversion.

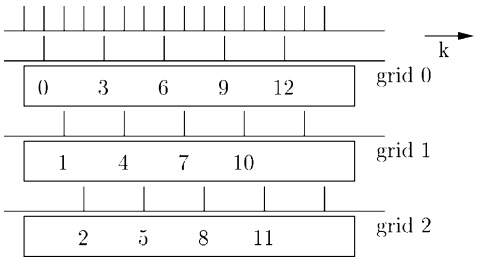


FIGURE 4 The RPE principle with $r = 3$ possible sub-sampling phases.

successful than a simple switch between the possible r phases of the down-sampling grid, according to some optimality criterion (see Figure 4). What remains is a decimated residual, applied in the receiver as a "regular-pulse excitation" (RPE) of $1/H(z)$. As always, an inverse LTP has to precede $1/H(z)$, if an LTP is included in the transmitter.

In a generalized RPE, optimized amplitudes may be applied instead of just samples. Furthermore, the N/r remaining values after down-sampling a block of N data could as well be placed on any optimal positions rather than a grid, leading to the concept of multi-pulse excitation (MPE). Both RPE and MPE can, on the other hand, be simplified by allowing only simple amplitudes like $\{-1, 0, 1\}$, with a common gain factor per block (or for suitable sub-blocks), thus separating a simplified shape and a gain of the excitation.

Vector quantization (VQ): After quantization, samples with a limited number of allowed values remain. So, only a certain set of possible excitation blocks is available at the receiver. Another, "direct" way to such excitations applies a table ("codebook") with, say, $L = 2^w$ length N (decimated or non-decimated) vectors from which the best one is chosen and coded by w bits per N samples; for, e.g. $L = 1024 = 2^{10}$ and $N = 40$, this means that $w = 10/40 = 1/4$ bits/sample are used in such a "vector-quantization" (VQ) scheme. The same principle can then also be applied to parameters like the predictor-coefficient vector **a** or any equivalent, transformed coefficient set, like the so-called reflection-coefficient vector **k**, the log-area-ratio vector **lar** [67], or the line-spectral-frequencies' vector **lsf** [42,77].

2.3 Closing the Gap

A first important practical realization helping to bridge the coding gap was the GSM full-rate (FR) codec (see Figure 5), standardized at the end of the 1980s by ETSI, applied world-wide since 1992 [12], based on those of the above ideas which were

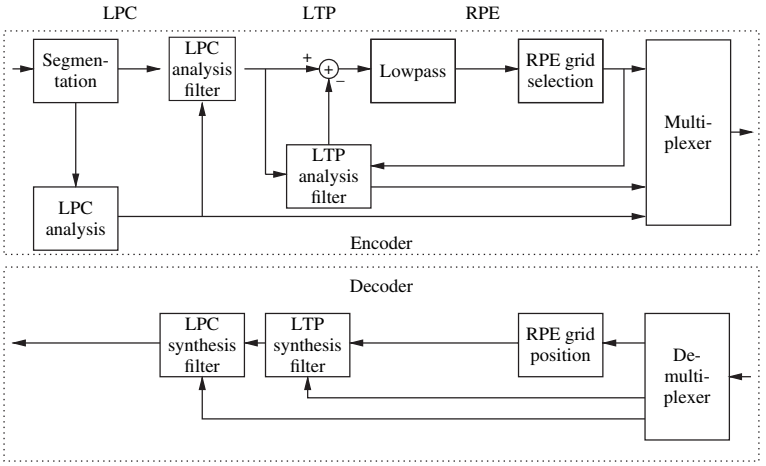


FIGURE 5 GSM-FR system (RPE-LTP-RELP with $r = 3$).

realizable with then available hardware, and with the best quality achievable thereby–reckoned as "not too good", nowadays.

This still quite limited quality as well as the tremendous growth of user numbers caused a wave of research towards improvements. They relied, firstly, on an enormous increase in computational power: nowadays, a 32-fold complexity and more can be envisaged. Secondly, the enhancements were based on some of the above-named tools, but above all, thirdly, on three *decisive* ideas described below.

2.4 The Breakthrough Basis

Vector quantization: Firstly, the VQ principle [31] turned out to be a very powerful concept. Work both on the theory of an N-dimensional extension of the pdf-based scalar optimum quantizer [58] and its limitations [7] and on practical codebook design [57] was helpful. VQ is applied to both receiver-excitation and parameters. Especially, VQ of the vector **lsf** achieves very good quality at only 20 . . . 24 bits per 20 ms-frame, i.e. a rate of 1 . . . 1.2 kbit/s [63].

Improved codebooks via efficient training, especially for $d(k)$ or $e(k)$, enhanced quality. Simple entries (0 and ± 1 only) or sparse vectors (adopting the RPE and MPE principles of Section 2.2) reduced computations and led to realizability. The simple (shape) vectors were combined with separate gains ("shape-gain VQ") and also efficiently transformed into more complex varieties by linear operations ("vector-sum (VS) VQ", "algebraic codebooks").
Analysis by synthesis: Secondly, and above all, however, a better optimization criterion was decisive: instead of minimizing the

MSE of the residual approximation, the distortion of the receiver-output signal (synthesized in the transmitter!) was reduced to a minimum. Furthermore, this minimization enclosed also the LTP into its closed loop, viewed as a second, (pitch-) adaptive codebook. Again, not an optimal pitch estimation is then found, but the most suitable period according to the criterion.

The underlying idea is (a) as simple as many good ideas, and (b) more than 40 years old ([4,35], see also [67]), but became realizable only with today's electronics. Actually, it means that the computation of the inverse filter $H(z)$ (plus the pitch predictor, if applicable) and the scalar or vector quantization of the residual signal cannot be regarded as independent steps. Rather, the best combination of all quantized (!) parameters and excitations has to be searched for - "best" with respect to the reconstructed, full-length signal. In practice, however, finite-length blocks are considered (with only some inclusion of the past), and the (LPC-) filter analysis is made "off-loop" as in Section 2.1 for a given input-signal block, after some pre-emphasis filtering by $H_G^{-1}(e^{j\Omega})$. But (within realizability:) "all" possible excitations are indeed applied to the full synthesis process, including all delayed versions of preceding excitations, i.e. possible (pitch-) periodic components.

Perceptual weighting: Furthermore, this "analysis-by-synthesis" technique included a perceptual weighting (as a crude perception model). For the above "best-match", the simple MSE is not an appropriate criterion: it maximizes the SNR only, which does not reflect a listener's quality impression for low-rate coders. Therefore, the difference between the synthesized and the original speech signals is spectrally weighted prior to the mean-square calculation and minimization. The weighting is carried out by a transfer function

$$W(z) = H(z)/H(z/\gamma),\ \gamma \approx 0.9. \tag{11}$$

$H(z)$ is the "inverse-vocal-tract" filter of Eq. (10). With $\gamma < 1$, $H(z/\gamma)$ corresponds to a "damped" version of $H(e^{j\Omega})$. Thus, $|W(e^{j\Omega})|$ varies with minima near the formant frequencies F_v and maxima between them. Thereby, distortions are less important close to strong spectral components, and emphasized otherwise: masking effects are coarsely modelled.

2.5 CELP Coders

Analysis by synthesis as well as perceptual weighting techniques can be applied in various systems, including LP-based RPE and MPE coders. Most successful, however, was the combination of analysis by synthesis, perceptual weighting by $W(z)$ (see Eq. (11)), vector or *c*odebook *e*xcitation, and LP: such coders are termed

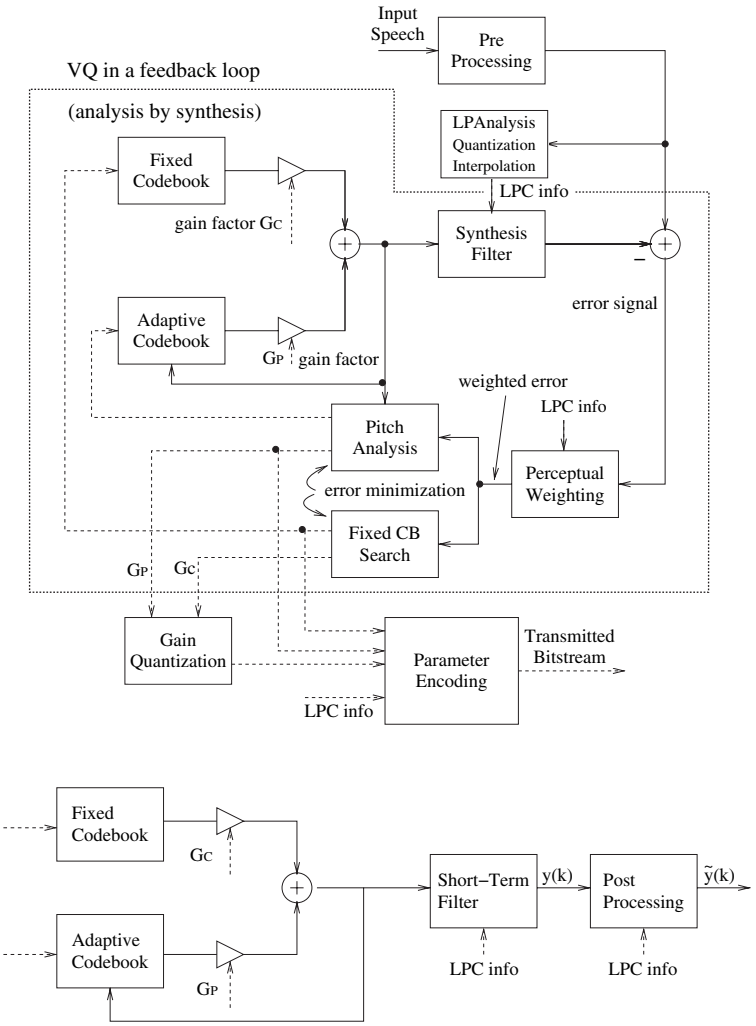


FIGURE 6 Typical CELP coder and decoder (ITU G.729 ACELP).

CELP systems [69]. Their numerous variants can be understood from the above considerations. Figure 6 shows, as a typical codec, one of the ITU standards [47]. It uses the approach of a simple-entry codebook with matrix manipulation, giving rise to its denotation as "algebraic CELP" (ACELP). This is one of many popular CELP systems in use nowadays. Some of them are listed below [39,71,77].

ITU G.729:	Algebraic CELP coder (ACELP); $f_B = 8\,\text{kbit/s}$; high quality; high complexity (18 MOps) [47]
GSM-HR:	ETSI half-rate coder; vector-sum excitation (VSELP); $f_B = 5.6\,\text{kbit/s}$; quality $\leq$ GSM-FR; high complexity (20 MOps) [13]
GSM-EFR:	ETSI-enhanced full-rate coder; ACELP similar to [47]; $f_B = 12.2\,\text{kbit/s}$; enhanced quality and error protection compared to GSM-FR; high complexity (18 MOps) [14]
ITU proposal:	ITU coder for high quality at low rates; "eX-CELP"; $f_B = 4\,\text{kbit/s}$; very high complexity (not yet fixed) [73]

ITU G.728:	ITU codec with low delay (2 ms, "LD-CELP");
	high-order prediction ($n = 50$!), backward adaptation;
	no LTP; very short codewords ($N = 5$);
	$f_B = 16\,\text{kbit/s}$;
	high quality; high complexity (20 MOps) [46]

MPEG-4 CELP:	MPEG-4 system with scalable data rate; ACELP with MPE/RPE variants; $f_B = 4 \ldots 24\,\text{kbit/s}$; variable quality; high to very high complexity (simple narrow-band coding: ≈ 20 MOps, enhanced narrow-band version: ≈ 40 MOps, enhanced wide-band coding: $\approx 50 \ldots 70 \ldots$ MOps) [32]
GSM-AMR:	ETSI adaptive multi-rate system; ACELP; variable rates $f_B = 4.75 \ldots 12.2\,\text{kbit/s}$, variable quality; up to high complexity (17 MOps) [15]
GSM WB-AMR:	ETSI wide-band AMR system: ACELP; low to high rates $f_B = 6.6 \ldots 23.85\,\text{kbit/s}$, up to very high quality; up to very high complexity (35 MOps) [76]

The two last ones adapt to changing transmission channels and/or user requirements, with 8 and 9 different rates, respectively; the latter is to be applied also in the enhanced GSM system and in UMTS. The notation "eX-CELP" refers to an "extension" of the CELP principles by combining it with other techniques and, above all, by varying sub-codebooks and even parts of the algorithm depending on a signal-segment classification [72]. The quality of these systems has to be compared to that of the common ISDN telephone and that of the wide-spread GSM-FR codec, i.e. to a relatively "high" and a "somewhat reduced" quality. A good CELP system may indeed yield "high quality", much better than RPE-LTP-RELP. But also the complexities have to be taken into account: while the ITU standards G.711 (PCM) and G.726 (ADPCM) need almost no computation, ETSI's GSM needs some 3.5 MOps (= Mega-Operations per second) already, while a good CELP coder may need 5–10 times as much or even more; 3.5 MOps present no problem for a simple digital signal processor (DSP), more advanced DSPs can handle 20 MOps easily, and cost start to really increase at 100 MOps – although hardware performance is still growing quickly.

2.6 Frequency-Domain Coding

In the early GSM preparation phase, most proposals aimed at "medium rates" with a different approach [38], namely, by means of "continuous" sub-band or "block-oriented" transform coders (SBC, TC). In fact, both are systems with analysis-synthesis filterbanks, whether realized by parallel band-pass filters or by filterbank-equivalent DFTs or DCTs; the true differences were their down-sampling factors r and corresponding numbers $M = r$ of frequency channels/components: $M = \ldots 8 \ldots 16 \ldots$ in SBC, $M = \ldots 128 \ldots$ in TC.

Figure 7 displays a typical TC system with adaptive quantization and bit allocation following the spectral envelope

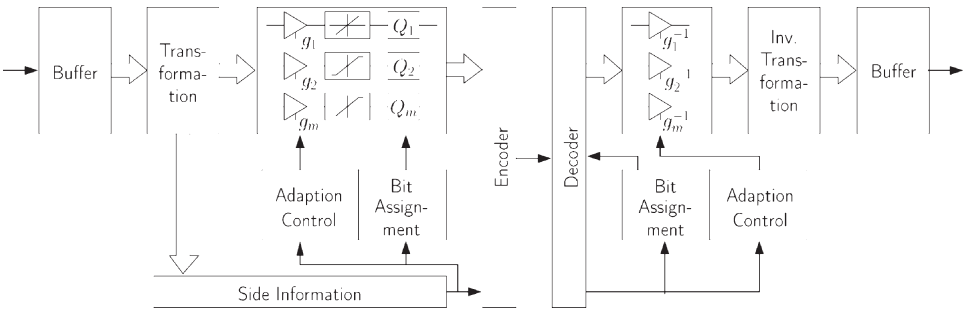


FIGURE 7 Adaptive transform codec (ATC).

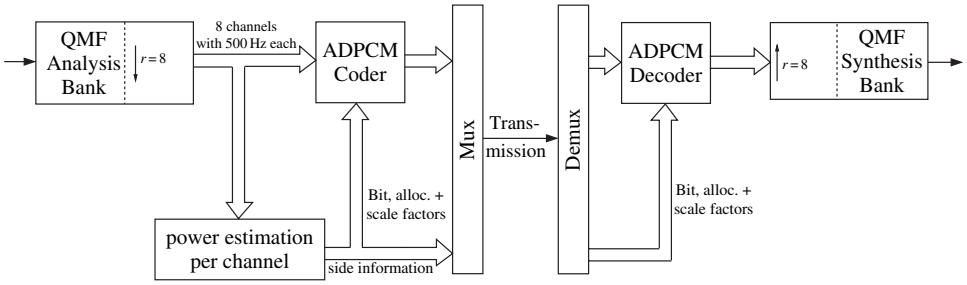


FIGURE 8 Sub-band codec (SBC) with adaptive bit allocation and ADPCM.

(ATC). This was realized in [24] with $f_B = 14\,\text{kbit/s}$, a quality close to GSM-FR, and a similar complexity. Figure 8 shows an SBC system with lower quality, but also lower complexity [55].

Of course, ideas applied in CELP systems could be transferred to frequency-domain coders: VQ is readily used for both main and side information – to spectral-value sets and to a spectral-envelope description for the bit allocation, based on a vector of averaged amplitudes as in [24,60] or on LPC results like **a** or **lsf**. Perceptual weighting could lead to a "damped envelope" in bit allocation, modelling (partial) masking again. Many such ideas were examined, but (for speech in contrast to audio coding!) frequency-domain coders were much less successful than the predictive systems discussed above, although there is no simple, generally accepted reason.

2.7 Other Ways Towards Closing the Gap

Describing speech segments as a finite sum of sinusoids with suitable amplitudes (coded directly or via an envelope, e.g. by means of LPC coefficients), frequencies and phases was intensively investigated in the late 1980s [74]. A trade-off between quality, complexity and rate was possible by choosing more or less refined descriptions especially of the frequencies and phases. A simplified version of this "sinusoidal modelling" (SM) applies a (pitch) frequency grid; this "harmonic-coding" (HC) scheme can be augmented by allowing narrow-band noise in unvoiced sections. SM and HC in their original forms can be seen as frequency-domain approaches; together with their "relatives" from Section 2.6, they have been of less interest in the past 10 years. From here, however, the way is not far to interesting hybrid time–frequency codecs with mixed or multi-band excitation (MELP, MBE, IMBE, EMBE), aiming at rates near 4 kbit/s and below [77].

2.8 Wide-Band Speech Coding

There has been some work on ATC or SBC for wide-band speech, especially when combined with music. At present, however, CELP systems, with an enlarged prediction order $n > 10$, dominate the scene. A variant – ACELP, similar to that in Figure 6 – is included in the list of Section 2.5, as a part of the ETSI WB-AMR standard.

A detail in this coder [6], already seen in some earlier development [11,64], is worth mentioning: the "split-band" principle of [44] reappears. But instead of separating frequencies below and above 4 kHz, an asymmetrical splitting takes place: 0 . . . 6 kHz (6.4 kHz in [6]) are directly (ACELP-) coded. The remaining band is replaced by noise (or aliased parts of the lower band, similar as in BB-RELP, see Section 2.2), which is filtered by $H^{-1}(z)$; only a gain factor needs to be transmitted. The gain is quite sensitive, but its word length can be reduced by linear prediction. The idea behind this was already used in an approach to "wide-band speech coding without *any* upper-band information": in [8,9], the components above 4 kHz were not transmitted, but substituted in the receiver as above, with a spectral weighting and gain term estimated from the lower band, using CELP techniques.

3 • Speech Transmission over Noisy Channels

3.1 Preliminary Remarks

Speech transmission is usually applied in conversational applications (telephony), which means that the tolerable delay is rather low. This limits the number of source samples that may be encoded commonly to, e.g. 160 in the European mobile radio standards [12–15].

As stated above, speech encoding is carried out in two steps: first, each block of source samples is decomposed into a set of parameter-vectors $x^1, x^2 \ldots$ which consist, e.g. of the LPC coefficients or the residual signal. In the second step the vectors x^j are quantized and the output bit vectors i^j are generated. This scenario is depicted in Figure 9. The two steps may be carried out jointly if the "analysis-by-synthesis" coding principle (as in CELP) is used. This, however, does *not* change the principle of

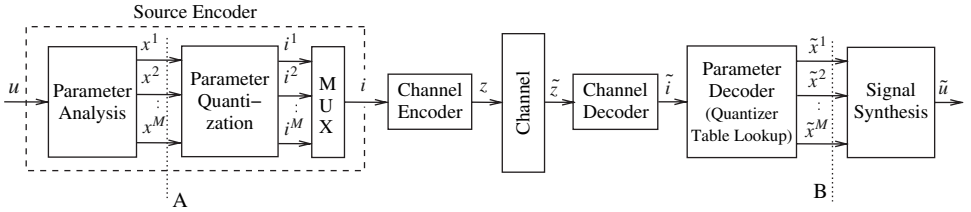


FIGURE 9 Basic model of a speech transmission system.

decomposing the actual source signal into parameter descriptions which are quantized and transmitted.

A channel code is applied to protect the source bits i against channel noise. Very often convolutional channel codes [56] are used, mainly because soft-information (unquantized matched-filter outputs) from the channel can be easily exploited by moderate-complexity decoding algorithms (Viterbi algorithm or BCJR algorithm [3]), which leads to strong gains over hard-decision decoding [56].

The actual goal of the system design is to obtain a decoder output signal $\tilde{u}$ that is as similar as possible to the input speech signal u. If the "similarity" shall be obtained in terms of *auditory* speech quality, it is very hard (if not impossible) to find a mathematically tractable measure for this, especially if it must have moderate complexity and has to operate on relatively short blocks of speech samples. But even if we used the simple mean-squared reconstruction error of the speech signal as a distortion measure, it would still be impossible to analytically optimize the system by variational techniques, because source codecs are complex algorithms that are hard to describe analytically (in the sense of a "transfer function").

There are two concepts to circumvent this problem. Firstly, one can use listening tests to get rid of the problem that we need an analytically tractable quality measure. The problem with this approach is that one must implement an idea which is mainly based on a "guess" and then it must be proven by listening tests that the scheme works. Clearly, this way of system design is extremely time-consuming and only a very small subset of potential options can be checked. Nevertheless, these methods have led to significant improvements in speech quality.

Another way is to optimize the speech quality *indirectly* by minimizing the bit error rate after channel decoding or by minimizing the reconstruction error of the speech codec *parameters* (i.e. minimization of the mean-squared error between the points "A" and "B" in Figure 9). Although there is no guarantee that the best possible quality is obtained, the concept is reasonable and it opens the way for mathematically tractable optimizations without the need to conduct listening tests.

Although we have a number of different parameters in Figure 9 we will, for brevity of notation, only consider one representative parameter signal below, and therefore we will omit superscripts. In what follows we will present a collection of methods to improve the quality for noisy transmission of coded speech signals.

3.2 Some Approaches for Quality Improvements

Comments on the separation principle: The foundation for the structure of today's communication systems is information theory: it guarantees that (possibly by use of infinitely large block lengths) one may realize the transmitter by a cascade of a source and a channel encoder with a binary interface in between (the same holds for the corresponding receiver), at which independent and uniformly distributed bits are exchanged – without any loss in performance compared with the best possible "one-step" encoder. The basic notion is to apply channel coding to achieve error-free bits at a bit rate that equals the channel capacity. The latter describes the amount of information that can be reliably transmitted over the channel. A source code is applied to reduce the bit rate required to represent the source signal down to the channel capacity. Clearly, this reduction is only possible at the price of some distortion when the source alphabet is continuous as in the case of speech and audio signals.[1]

If the separation principle is applied in a practical system with *limited block lengths* the following problems occur:

- It is impossible to realize channel codes that have a residual bit error rate of zero. The reason is that, as long as the block length is limited, in some blocks more random channel errors may occur than can be corrected.
- Practical source encoders do not produce independent and uniformly distributed bits. This is again due to the limited block length. If, for instance, we encode a correlated signal, there is always correlation left between adjacent blocks and, hence, between the quantizer indices.
- The distortion at the source decoder output depends on which bit is in error, because any practical quantizer for a limited block length has more or less "significant" bits. In case of parametric coding such as CELP, even the codec parameters we extract from the signal are not equally important for the (auditory) quality: a bit error in the quantizer index for the signal power leads to severe quality degradations while a flipped sign-bit of the excitation signal is sometimes not even noticeable.

In spite of these "non-idealities," source and channel coding are separate processing steps in all practical systems (with some important modifications discussed below). Besides the

[1]It is common practice to consider the amplitudes of speech and audio signals as continuous, although in an implementation the samples are typically represented by 16 bits (linear PCM) and, thus, they are pre-quantized before the actual source compression takes place.

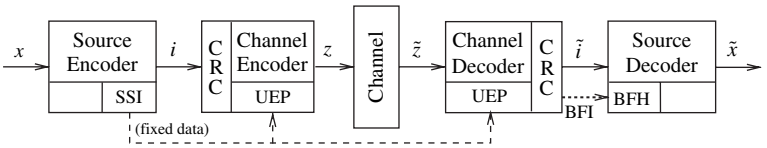


FIGURE 10 Unequal error protection (UEP) for which source significance information (SSI) is used. Error concealment with a cyclic redundancy check (CRC) that at the receiver sets the bad frame indicator (BFI) initiating bad frame handling (BFH).

separate history of source and channel coding motivated by information theory, the practical reason is that the separation allows to construct "partial" encoders/decoders that, compared with asymptotic results, sometimes may have only moderate performances[2] but their complexities are tolerable. Moreover, the system design is greatly simplified.

Unequal error protection and error concealment: Figure 10 shows some system modifications that account for the drawbacks of the separation principle stated above. Implicitly, a channel with discrete input alphabet (application of digital modulation) is assumed, because the channel codeword z, a bit vector, is directly fed into the channel. Some of the source-encoder output bits cause stronger quality degradations than others if they are in error but are used for decoding anyhow. Source significance information (SSI) can be derived from that and be used by the channel-coding scheme to apply stronger error protection to the more sensitive bits.

The residual redundancies in the source-encoder output bits can be exploited to conceal bit errors, e.g. by repetition of the old parameters from the previous block (bad frame handling (BFH)). Clearly, frame repetitions only make sense, if the codec parameters are correlated in time. This is the case, e.g. for the signal power, the LPC coefficients or the "pitch-lag". For the initiation of the error concealment, an error detection (bad frame indication, BFI) is required, which is usually realized by an error-detecting channel code (cyclic redundancy check (CRC)). The strongly sensitive bits of the source encoder are CRC-encoded (Class-1-bits) and afterwards the CRC code word is channel-encoded together with the remaining data bits. Thus, in combination with UEP, the bits of the source encoder are divided into

[2]It should be mentioned that in the "waterfall-region" Turbo Codes [5] and the developments based thereon brought the performance of channel coding schemes with fairly long block lengths very close to the information theoretical bounds, if a residual bit error rate of, say, 10^{-6} is accepted.

several classes which are protected by adjustable amounts of channel-code redundancy. Fortunately, the more sensitive bits are typically the stronger correlated ones at the same time, i.e. UEP and error concealment can be combined very efficiently. Hence, both schemes are frequently used, e.g. in mobile radio standards [17–20].

Source-controlled channel decoding: The idea of source-controlled channel decoding is to exploit the residual redundancies systematically as a-priori information in the decision procedure for the data bits in channel decoding to reduce the bit error rate [1,33]; the basic setup is depicted in Figure 11. Source-controlled channel decoding can be implemented very efficiently for the decoding of binary convolutional channel codes and it can be extended to soft-output channel decoders (APRI-SOVA, [33]), which can be combined with estimation-based source decoders described below. For a correct weighting of the a-priori information and the soft channel outputs, a channel state information (CSI, e.g. noise variance) must be available.

A drawback is that the redundancies are usually exploited on bit basis[3] (when, as in many cases, binary channel codes are used), although the source encoder often emits correlated *indices* that consist of several bits. Thus, a large part of the redundancies is removed if only the marginal distributions for the individual bits are used.

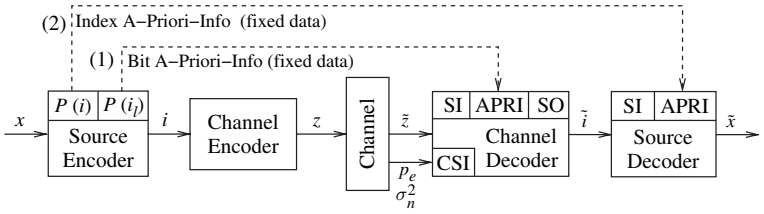


FIGURE 11 (**1**) Source-controlled channel decoding with soft inputs (SI) and outputs (SO) using channel state information (CSI) and source-bit a-priori information (APRI); (**2**) estimation-based source decoding using index a-priori information.

[3]In [36] a channel-decoding scheme for binary convolutional codes is given that is able to optimally exploit index-based a-priori information, if all the bits of an index are adjacently mapped into the input of the channel encoder. However, the method may not be applicable because usually bit-reordering/interleaving is employed at the channel-encoder input to account for different sensitivities of the bits or to improve the distance properties of the concatenated code.

Estimation-based source decoding: In conventional systems, simple procedures, e.g. quantizer table-lookups, are frequently applied for source decoding. If, however, source encoding is not perfect and residual redundancies – e.g. correlations in time – are left in the data bits, this a-priori knowledge can be exploited for better source decoding by performing estimations of the source-encoder input [23,54,59,66,68,70]; such a system is also depicted in Figure 11. In contrast to source-controlled channel decoding, it is now possible to exploit the full correlations on *index* basis and the soft-in/soft-out (SISO) channel decoder merely acts as a device that improves the reliability of the virtual channel "seen" by the source decoder.

Frequently, we want to minimize the squared error in the reproduction $\tilde{x}_k$ of the parameter values x_k. For this, the estimator

$$\tilde{x}_k = \sum_{\forall i} \underbrace{\Pr(I_k = i|\tilde{\tilde{i}}_k, \tilde{\tilde{i}}_{k-1}, \ldots)}_{\text{index APP}} \cdot x_q(i) \tag{12}$$

is a suitable decoding rule [23,70]. The decoder output $\tilde{x}_k$ is computed as a weighted average of all quantizer reproducer values (or vectors) $x_q(i)$. The weighting factors are the a-posteriori probabilities (APPs) of the quantizer indices (for which we use the random variable I_k at time k and the realizations i). The APPs are conditioned on the current and all previous channel decoder outputs $\tilde{\tilde{i}}_k, \tilde{\tilde{i}}_{k-1}, \ldots$, in order to exploit the memory of the source. Hence, the main problem is how to compute these APPs. If we assume a first-order Markov chain to model the autocorrelations of the quantizer indices we obtain [23,66]

$$\underbrace{\Pr(I_k = i|\tilde{\tilde{i}}_k, \tilde{\tilde{i}}_{k-1}, \ldots)}_{\text{index APP}} = A_k \cdot \underbrace{p(\tilde{\tilde{i}}_k|I_k = i)}_{\text{channel term}} \cdot \sum_{\forall i'} \underbrace{\Pr(I_k = i|I_{k-1} = i')}_{\text{first-order Markov model}} \cdot \underbrace{\Pr(I_{k-1} = i'|\tilde{\tilde{i}}_{k-1}, \tilde{\tilde{i}}_{k-2}, \ldots)}_{\text{"old" index APP}}, \tag{13}$$

where A_k is a normalizing constant that ensures that $\Pr(I_k = i|\tilde{\tilde{i}}_k, \tilde{\tilde{i}}_{k-1}, \ldots)$ sums up to one over all possible indices i. Equation (13) is a recursion that allows to compute the index APPs from the "old" APPs at the previous time instant,[4] from the Markov transition probabilities $\Pr(I_k = i|I_{k-1} = i')$ of the statistical model for the quantizer indices and from the channel term. The latter weights the statistical knowledge about the source with

[4] At time $k = 0$ we start with the unconditional probability distribution of the quantizer indices.

the probability density function of the channel-decoded currently received channel information $\tilde{i}_k$. Hence, a soft-output channel decoder is needed [3] that provides reliability information about the channel decoded data bits and not only hard decisions. As in the previous section, bit-based a-priori information may be used to aid channel decoding, but it is not clear how the a-priori information shall be optimally divided between the channel decoder and the source decoder. This problem gave rise to the idea of "true" joint source-channel decoding.

Joint source-channel decoding: Many schemes in the area of "joint source-channel decoding" use residual redundancies to improve *either* channel decoding *or* source decoding (but usually not both, see also [22]). At the end of the previous section it was discussed that it is unclear how to divide residual correlation between source and channel decoding. Hence, it is just consequent to completely join source and channel decoding. If only one parameter is considered that is individually channel-encoded, and we are again interested in a good estimation in the minimum mean-squared error sense, we obtain [26,27]

$$\tilde{x}_k = \sum_{\forall i} \underbrace{\Pr(I_k = i | \tilde{z}_k, \tilde{z}_{k-1}, \ldots)}_{\text{index APP}} \cdot x_q(i) \tag{14}$$

with

$$\underbrace{\Pr(I_k = i | \tilde{z}_k, \tilde{z}_{k-1}, \ldots)}_{\text{index APP}} = A_k \cdot \underbrace{p(\tilde{z}_k | Z_k = CC(i))}_{\text{"channel" term}} \cdot \sum_{\forall i'} \underbrace{\Pr(I_k = i | I_{k-1} = i')}_{\text{first-order Markov model}} \cdot \underbrace{\Pr(I_{k-1} = i' | \tilde{z}_{k-1}, \tilde{z}_{k-2}, \ldots)}_{\text{"old" index APP}} \tag{15}$$

where the notation $Z_k = \mathrm{CC}(i)$ corresponds to the channel codeword Z_k that results from the quantizer index i. Note that the index APPs and the channel term, in contrast to (13), now contain the received channel words $\tilde{z}_k$ (instead of $\tilde{i}_k$) and the transmitted channel codeword Z_k (instead of $I_k = i$) that results from the channel code by some hypothesis i for the quantizer index. As long as we channel-encode the indices individually, there is no major difference in complexity compared with the method given by (13).[5] But in practice, we will have a set of quantizer indices (for the LPC coefficients, the lag information and others) that are jointly channel encoded. Therefore, the

[5] Some results for joint source-channel decoding for speech transmission by a CELP speech codec in the GSM half-rate channel are given in [25].

received channel codeword depends on the whole set of source-codec indices and, hence, we would have to compute the channel term in (15) for all possible bit combinations of all quantizer indices: an infeasible task if we have, e.g. 10 indices with a total of 200 data bits. Although these issues limit the applicability of (15), the formula provides an interesting and unconventional view of channel coding: its effect on the source decoder output $\tilde{x}_k$ is in no way different from other kinds of redundancies like the residual correlations of the quantizer indices. Both types of redundancy form weighting factors for the APPs of the data indices in (15). This insight into source-channel coding paved the way for iterative source-channel decoding.

Iterative source-channel decoding: In [27,29,65] iterative source-channel decoding is introduced. Although the sketch of the system in Figure 12 is similar to the systems in Figure 11, the theoretical concept is different: the transmitter is interpreted as a serially concatenated coding scheme; the constituent "codes" are formed by the implicit residual redundancies within the source encoder indices and by the explicit redundancy of the channel code.

The interpretation as a concatenated coding system calls for the use of the turbo principle [5,34] for decoding: as in all iterative decoding schemes so-called APP decoders (a-posteriori probability decoders) for both constituent "codes" must be available that are able to exchange extrinsic information on the data bits within the decoding iterations. The extrinsic information contains only the *new* part of information that is generated from *one* type of redundancy. While APP decoders for the frequently used binary convolutional channel codes are well known from literature [3,34], they can also be formulated on the basis of estimation-based decoding for the source-coding part of the system [27,28]. Among the realizable decoding schemes, iterative source-channel decoding works best for a given transmitter but at the same time it is the most complex approach, as several iterations of the component decoders are needed.

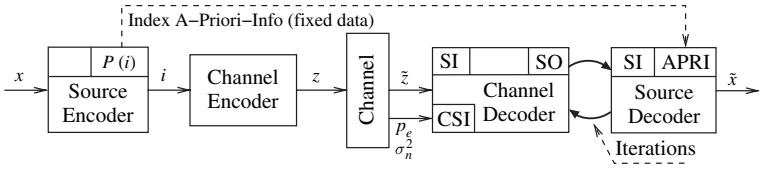


FIGURE 12 Iterative source-channel decoding with a soft-in (SI) soft-out (SO) channel decoder that requires channel state information (CSI) and a soft-in source decoder using a-priori information (APRI).

An intermediate discussion: In practice, it strongly depends on the encoding scheme if the application of advanced receiver structures such as iterative source-channel decoding is recommendable: if the source encoder works very well, it will leave only a small amount of redundancy in its output bits. Hence, it may be questionable if the effort for advanced joint decoding schemes really pays off as the potential quality gain might be small. Moreover, any decoding scheme can only be optimal for one given transmitter but *if* we are free to choose the encoder too, it is not really clear which scheme is optimal, especially in terms of auditory quality.

We recall from the earlier discussion of the separation principle that, ideally, the channel code makes available the amount of information that can be reliably transmitted over the channel. As the channel capacities in mobile radio are strongly time-variant, we actually have to vary the channel code rate, too. But if we are forced (e.g. by a standard) to use a fixed channel code rate we can only choose it for the "average case". This means, however, that the scheme is not optimal in many cases and that we also should consider more flexible systems including encoder modifications. In the discussion above we actually did that already, when we added a channel code to at least detect the errors that cannot be corrected and we used unequal error protection. Below we will discuss more elaborate schemes that can also be adapted to varying channels statistics.

Channel-optimized vector quantization: The idea of channel-optimized vector quantization (COVQ) [21,51,78] is to use the statistics of the source and the channel in the design of the source encoder (quantizer codebook) and in the encoding algorithm. COVQ is one of the few concepts which do not separate at all into source and channel coding. The scheme assumes a limited block length which is fixed for each source vector (i.e. fixed bit rate) but there are no further restrictions. Given the system constraints, COVQ does the best job possible with strong gains over source-optimized VQ.

Due to the complexity of encoding and the memory requirements for the codebook, COVQ is only applicable, if the number K of bits representing the quantizer indices is smaller than, e.g. 10. Therefore, COVQ cannot be used directly for speech, audio or image coding. In our setup, COVQ could be used to quantize the parameters of a source codec. Note that as long as the encoder and the decoder use the same codebooks, COVQ could be also used within an analysis-by-synthesis loop of an advanced CELP speech codec. If the channel conditions change, it is possible to adapt COVQ by use of an appropriate codebook that is designed

for the current bit error rate. This requires, however, that the bit error rate is known at the transmitter, which is only realizable if the channel statistics change slowly (that is not the case for fast fading on a mobile radio link). Another drawback is that one would, in principle, need a new COVQ codebook for each channel condition, i.e. a huge amount of storage would be necessary. It was observed in [48] that the performance degradation due to moderately varied channel properties is small, so that a small set of codebooks is enough to cover the range of potential channel variations. Moreover, in [30] a simplification is introduced, which often allows to almost achieve the performance of COVQ while only one codebook and a channel-adaptive scaling factor for the fixed codevectors are required.

A major problem with parameter-individual COVQs is that at a bit rate of, say, one bit per sample, we could apply efficient error control coding for a long block of, e.g. 200 data bits. That would be much more efficient than the implicit error control implemented by each "small" parameter-individual COVQ. Therefore, COVQ will probably not find its way into speech coding with moderate block lengths, but there might be low-delay speech coding applications in which the block length is too small for efficient channel coding. An example is line-multiplexers for transatlantic telephony which use the low-delay CELP codec [46].

Adaptive rate-allocation techniques: From the discussion above it is clear that for varying channel conditions we should vary the bit rates for source and channel coding. As this rate adaptation requires signalling between the transmitter and the receiver, it cannot be carried out extremely fast, i.e. it is mainly useful to combat relatively slow fading in mobile radio.[6]

The basic idea is also clear from the discussion above: if the channel causes strong impairments, the source-coding bit rate is lowered and the number of redundancy bits used for channel coding is enlarged. This improves the error correction capabilities at the price of lower quality of source coding. If the channel noise is low, we choose a small number of redundancy bits and we increase the bit rate for source coding. As standardized for the adaptive multirate (AMR) system in GSM [15,16,20] and also for UMTS, the idea can be implemented by a set of coding schemes (for an introduction, see [40]).

---

[6] A way to combat *fast* fading is to use interleaving (which means increased delay) to spread the bits of a channel codeword over longer time periods, so that, in spite of the quickly time-varying channel, we obtain almost constant channel statistics after de-interleaving. The method is widely used in mobile radio.

Rate-allocation techniques can be combined with the other source-channel (de)coding methods described above, i.e. a good system will contain a mixture of rate-allocation techniques, bad frame handling, unequal error protection and possibly estimation-based/iterative source-channel decoders.

4 • Further Work, Results and Literature

4.1 Speech Coding

Good telephone quality will soon be achieved at $f_B = 4\,\text{kbit/s}$, while good wide-band quality is available at $f_B = \ldots 12 \ldots \text{kbit/s}$. Good telephone speech with $f_B = \ldots 2.4 \ldots \text{kbit/s}$ is currently studied. For instance, SM/HC ideas reappear in the MPEG-4 subsystem called "harmonic vector-excited coder" (HVXC); pitch-dependent waveform-interpolation concepts as proposed in [50] are investigated; models of the hearing process are used in the LP excitation [2]; also phoneme classification is applied [10].

4.2 Joint Source-Channel Coding

During the last decade, various methods have been derived that strongly improve the quality of coded speech transmission over noisy channels. The more elaborate methods such as estimation-based decoding will have to be implemented efficiently to make their potential available in practice. Techniques for rate adaptation, depending on the quality demands and on the current channel conditions, are important fields of future work. Another crucial aspect is the evolution from traditional circuit-switched to packet-switched transmission techniques, which are used, e.g. in the Internet.

4.3 Further Reading

A huge variety of literature, including books, are available. The choice given here is, of course, limited – and subjective. This holds also for the following recommendations:

A good entrance into the details of waveform-preserving coder approaches applied to speech (and video) is found in the "standard" [49], a broader overview on speech processing in [62], a less broad but more detailed and recent one in [77]. Two other good references for speech coding are [41,52]. The books also contain good introductory as well as more advanced aspects of error concealment and joint source-channel coding. A comprehensive treatment of channel coding can be found in [56]. All books and overview papers are, of course, rich sources for further references.

References

1. Alajaji, F., Phamdo, N., Fuja, T. (1996) Channel Codes that Exploit the Residual Redundancy in CELP-Encoded Speech. IEEE Trans. Speech Audio Process., **4**, 325–336

2. Ambikairajah, E., et al. (2001) Wideband-Speech and Audio Coding Using Gammatone Filterbanks. Proc. IEEE Int. Conf. Acoustics Speech Signal Process., Salt Lake City, **II**, 773–776
3. Bahl, L.R., Cocke, J., Jelinek, F., Raviv, J. (1974) Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate. IEEE Trans. Inform. Theory, **20**, 284–287
4. Bell, C.G., Fujisaki, H., Heinz, J.M., Stevens, N.K., House, A.S. (1961) Reduction of Speech Spectra by Analysis-by-Synthesis Techniques. J. Acoust. Soc. Am., **33**, 1725–1736
5. Berrou, C., Glavieux, A. (1996) Near Optimum Error Correcting Coding and Decoding: Turbo-Codes. IEEE Trans. Commun., **44**, 1261–1271
6. Bessette, B., Jaervinen, K., et al. (2001) Techniques for High-Quality ACELP Coding of Wide-Band Speech. Proc. EUROSPEECH, Aalborg, **3**, 1993–1996
7. Brehm, H., Stammler, W. (1987) Description and Generation of Spherically-Invariant Speech-Model Signals. EURASIP Signal Process., **12**, 119–141
8. Carl, H. (1994) Examination of Different Speech-Coding Methods and an Application to Band-Width Enhancement for Narrow-Band Speech Signals (in German). Diss. Ruhr-Univ., Bochum.
9. Carl, H., Heute, U. (1994) Bandwidth Enhancement of Narrow-Band Speech Signals. Proc. EUSIPCO, Edinburgh, 1716–1719
10. Ehnert, W., Heute, U. (1997) Variable-Rate Speech Coding: Replacing Unvoiced Excitations by Linear-Prediction Residues of Different Phonemes. Proc. GRETSI, Grenoble, 993–996
11. Erdmann, Ch., Vary, P., et al. (2001) A Candidate Proposal for a 3GPP Adaptive Multi-rate Wide-Band Speech Codec. Proc. IEEE Int. Conf. Acoustics Speech Signal Process., Salt Lake City, **II**, 757–760
12. ETSI (1999) Rec. GSM 06.10: Digital Cellular Telecommunication System (phase 2+) (GSM); Full-Rate Speech Transcoding
13. ETSI (1999) Rec. GSM 06.20: Digital Cellular Telecommunications System (phase 2+) (GSM); Half-Rate Speech Transcoding
14. ETSI (1999) Rec. GSM 06.60: Digital Cellular Telecommunications System (phase 2+) (GSM); Enhanced Full-Rate (EFR) Speech Transcoding
15. ETSI (1999) Rec. AMR 06.90: Digital Cellular Telecommunications System: Adaptive Multi-rate (AMR) Speech Transcoding
16. ETSI (1999) Rec. GSM 05.03: Digital Cellular Telecommunications System (phase 2+) (GSM): Channel Coding
17. ETSI (1999) Rec. GSM 06.11: Digital Cellular Telecommunications System (phase 2+) (GSM): Substitution and Muting of Lost Frames for Full-Rate Speech Channels
18. ETSI (1999) Rec. 06.21: Digital Cellular Telecommunications System (phase 2+) (GSM): Substitution and Muting of Lost Frames for Half-Rate Speech Traffic Channels
19. ETSI (1999) Rec. 06.61: Digital Cellular Telecommunications System (phase 2+) (GSM): Substitution and Muting of Lost Frames for Enhanced Full-Rate (EFR) Speech Traffic Channels

20. ETSI (1998) Rec. 06.91: Digital Cellular Telecommunications System (phase 2+) (GSM): Substitution and Muting of Lost Frames for Adaptive Multi-rate (AMR) Speech Traffic Channels
21. Farvardin, N., Vaishampayan, V. (1991) On the Performance and Complexity of Channel-Optimized Vector Quantizers. IEEE Trans. Inform. Theory, **37**, 155–160
22. Fazel, T., Fuja, T. (2003) Robust Transmission of MELP-Compressed Speech: An Illustrative Example of Joint Source-Channel Decoding. IEEE Trans. Commun., **51**, 973–982
23. Fingscheidt, T., Vary, P. (1997) Robust Speech Decoding: A Universal Approach to Bit Error Concealment. Proc. IEEE Int. Conf. Acoust. Speech Signal Process., **3**, 1667–1670
24. Gluth, R., Guendel, L., Heute, U. (1986) ATC – A Candidate for Digital Mobile-Radio Telephony. Proc. Nordic Semin. Dig. Land-Mob. Rad. Commun., Stockholm, 230–235
25. Görtz, N. (1999) Joint Source-Channel Decoding by Channel-Coded Optimal Estimation (CCOE) for a CELP Speech Codec. Proc. EUROSPEECH, **3**, 1447–1450
26. Görtz, N. (2000) Joint Source-Channel Decoding by Channel-Coded Optimal Estimation. Proc. 3rd ITG Conf. Source Channel Cod., VDE-Verlag, 267–272
27. Görtz, N. (2001) On the Iterative Approximation of Optimal Joint Source-Channel Decoding. IEEE J. Select. Areas Commun., **19**, 1662–1670
28. Görtz, N. (2003) Optimization of Bit-Mappings for Iterative Source-Channel Decoding. Proc. Int. Symp. Turbo Codes Relat. Top., 255–258
29. Görtz, N., Heute, U. (2000) Joint Source-Channel Decoding with Iterative Algorithms. Proc. EUSIPCO, **1**, 425–428
30. Görtz, N., Kliewer, J. (2003) Memory-Efficient Adaptation of Vector Quantizers to Time-Varying Channels. EURASIP Signal Process., **83**, 1519–1528
31. Gray, R.M. (1984) Vector Quantization. IEEE ASSP Mag., **1**, no. 2, 4–29
32. Grill, B., Edler, B., et al. (1998) Information Technology – Very Low Bitrate Audio-Visual Coding, Part 3: Audio. ISI/IEC FCD 14496–3, Subpart 1
33. Hagenauer, J. (1995) Source-Controlled Channel Decoding. IEEE Trans. Commun., **43**, 2449–2457
34. Hagenauer, J., Offer, E., Papke, L. (1996) Iterative Decoding of Binary Block and Convolutional Codes. IEEE Trans. Inform. Theory, **42**, 429–445
35. Halle, M., Stevens, N.K. (1959) Analysis by Synthesis. Proc. Semin. Inform. Theory Speech Compression **II**, paper no. D7
36. Heinen, S., Geiler, A., Vary, P. (1998) MAP Channel Decoding by Exploiting Multilevel Source A-Priori Knowledge. Proc. ITG-Fachtgg. "Cod. für Quelle, Kanal und Übertrgg.", 89–94
37. Hess, W. (1983) Pitch Determination of Speech Signals. Springer, Berlin

38. Heute, U. (1988) Medium-Rate Speech Coding – Trial of a Review. Speech Commun., **7**, 125–149
39. Heute, U. (1994) Speech Coding: Approaches, Trends, Standards (in German). Proc. ITG Conf. Source Channel Cod., Munich, 437–448
40. Hindelang, T., Kaindl, M., Hagenauer, J., Schmauz, M., Xu, W. (2000) Improved Channel Coding and Estimation for Adaptive Multirate (AMR) Speech Transmission. Proc. IEEE 51st Vehic. Technol. Conf., 1210–1214
41. Kleijn, W.B., Paliwal, K.K. (Editors) (1995) Speech Coding and Synthesis. Elsevier
42. Itakura, F. (1975) Line-Spectral Representation of Linear-Prediction Coefficients of Speech Signals. JASA, **57**, Suppl. S35A
43. ITU/CCITT (1988) Recommendation G.711: "Coding of Analogue Signals by Pulse-Code Modulation (PCM) of Voice-Frequencies"
44. ITU-T (1988) Recommendation G.722: "7 kHz Audio Coding within 64 kbit/s. Fascicle III.4, Blue Book, 269–341
45. ITU-T (1992) Recommendation G.726: "40,32,24,16 kbit/s Adaptive Differential Pulse-Code Modulation (ADPCM)"
46. ITU-T (1992) Recommendation G.728: "Coding of Speech at 16 kbit/s Using Low-Delay Code-Excited Linear Prediction"
47. ITU-T (1995) Recommendation G.729: "Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic Code-Excited Linear Prediction (CS-ACELP)"
48. Jafarkhani, H., Farvardin, N. (2000) Design of Channel-Optimized Vector Quantizers in the Presence of Channel Mismatch. IEEE Trans. Commun., **4**, 118–124
49. Jayant, N.S., Noll, P. (1984) Digital Coding of Waveforms. Prentice Hall, Englewood Cliffs, NJ
50. Kleijn, B., Haagen, J. (1995) A Speech Coder Based on Decomposition of Characteristic Waveforms. Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process., Detroit, 508–511
51. Kumazawa, H., Kasahara, M., Namekawa, T. (1984) A Construction of Vector Quantizers for Noisy Channels. Electron. Eng. Jpn, **67-B**, 39–47
52. Kondoz, A. (1994) Digital Speech for Low Bit Rate Communication Systems. Wiley
53. Lacroix, A. (2004) Speech-Production – Acoustics, Models, and Applications. In: Blauert, J. (Ed.): Communication Acoustics, Chapter 13. Springer, Berlin Heidelberg.
54. Liu, F.H., Ho, P., Cuperman, V. (1993) Joint Source and Channel Coding Using a Non-linear Receiver. Proc. IEEE Int. Conf. Commun., **3**, 1502–1507
55. Lazzari, V., Montagna., R., Sereno, D. (1988) Comparison of Two Speech Codecs for DMR Systems. Speech Commun., **7**, 193–207
56. Lin, S., Costello, D.J. (2nd edition 2004) Error Control Coding. Pearson Prentice-Hall
57. Linde, Y., Buzo, A., Gray, R.M. (1980) An Algorithm for Vector-Quantizer Design. IEEE Trans. Commun. **28**, 84–95

58. Max, J. (1960) Quantizing for Minimum Distortion. IRE Trans. Inform. Theory, **6**, 7–12
59. Miller, D.J., Park, M. (1998) A Sequence-Based Approximate MMSE Decoder for Source Coding over Noisy Channels Using Discrete Hidden Markov Models. IEEE Trans. Commun., **46**, 222–231
60. Noll, P., Zelinski, R. (1977) Adaptive Transform Coding of Speech Signals. IEEE Trans. Acoustics Speech Signal Process., **25**, 299–309
61. Noll, P. (1993) Wideband-Speech and Audio Coding. IEEE Commun. Mag., **Nov**., 34–44
62. O'Shaughnessy, D. (1987) Speech Communication – Human and Machine. Addison-Wesley, New York
63. Paliwal, K.K., Atal, B.S. (1993) Efficient Vector Quantization of LPC Parameters at 24 bits/frame. IEEE Trans. Speech Audio Process, **1**, 3–14.
64. Paulus, J. (1997) Coding of Wide-Band Speech Signals at Low Data Rates (in German). Diss. RWTH, Aachen
65. Perkert, R., Kaindl, M., Hindelang, T. (2001) Iterative Source and Channel Decoding for GSM. Proc. IEEE Int. Conf. Acoustics Speech Signal Process., **4**, 2649–2652
66. Phamdo, N., Farvardin, N. (1994) Optimal Detection of Discrete Markov Sources over Discrete Memoryless Channels–Applications to Combined Source-Channel Coding. IEEE Trans. Inform. Theory, **40**, 186–193
67. Rabiner, L.R., Schafer, R.W. (1978) Digital Processing of Speech Signals. Prentice-Hall, Englewood Cliffs, NJ
68. Sayood, K., Borkenhagen, J.C. (1991) Use of Residual Redundancy in the Design of Joint Source/Channel Coders. IEEE Trans. Commun., **39**, 838–846
69. Schroeder, L.R., Atal, B.S. (1985) Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates. Proc. IEEE Int. Conf. Acoustics Speech Signal Process., Paris, **I**, 25.1.1–25.1.4.
70. Skoglund, M. (1999) Soft Decoding for Vector Quantization over Noisy Channels with Memory. IEEE Trans. Inform. Theory, **45**, 1293–1307
71. Spanias, A.S. (1994) Speech Coding: A Tutorial Review. Proc. IEEE **82**, 1541–1582
72. Thyssen, J., et al. (2001) EX-CELP: A Speech-Coding Paradigm. Proc. IEEE Int. Conf. Acoustics Speech Signal Process., Salt Lake City, **II**, 689–692
73. Thyssen, J., et al. (2001) A Candidate for the ITU-T 4 kbit/s Speech-Coding Standard. Proc. IEEE Int. Conf. Acoustics Speech Signal Process., Salt Lake City, **II**, 681–684
74. Trancoso, I.M., et al. (1988) Harmonic Coding – State of the Art and Future Trends. Speech Commun., **7**, 239–245
75. Tremain T.E. (1982) The Government Standard Linear-Predictive Coding: LPC-10. Speech Technol., **1**, 40–49
76. Varga, I. (2001) Standardization of the Adaptive Multi-rate Wideband Codec. Proc. ITG Conf. Source Channel Cod., Berlin, 341–346

77. Vary, P., Heute, U., Hess, W. (1998) Digitale Sprachsignalverarbeitung. Teubner, Stuttgart
78. Zeger, K., Gersho, A. (1998) Vector Quantizer Design for Memoryless Noisy Channels. Proc. IEEE Int. Conf. Commun., **3**, 1593–1597
79. Zwicker, Fastl (1990). Psychoacoustics – Facts and Models. Springer, Berlin Heidelberg

46

Echo Cancellation

Walter Kellermann

Multimedia Communications and Signal Processing, University Erlangen-Nuremberg, Erlangen, Germany, e-mail: wk@lnt.de

1 • Echo Cancellation and Suppression

Echoes as perceived with speech communication terminals in telecommunication networks typically result from two sources (Figure 1): Firstly, the so-called *electrical echoes* or *line echoes* are usually caused by hybrid circuits, as employed for connecting four-wire links and two-wire links [1]; secondly, *acoustic echoes* result from acoustic feedback from the loudspeakers into the microphones of the terminal equipment with widely varying feedback gain depending on the terminal and the acoustic scenario. In both cases, such echoes should be removed from the signal $x(n)$ that is transmitted to the far-end terminal. The signal $x(n)$ can generally be written as a sum of three components

$$x(n) = x_s(n) + x_v(n) + x_r(n), \tag{1}$$

where in the acoustic echo scenario, $x_s(n)$ represents the component due to the local talker's original speech $s(n)$, $x_v(n)$ is the loudspeaker signal echo and $x_r(n)$ results from noise sources $r(n)$ in the acoustic environment. In the electrical echo scenario, $x_r(n)$ is often neglected and $x_s(n)$ corresponds to the speech from the local terminal. While *echo suppression* aims at attenuating the residual echo $x_v(n)$ within $z(n)$ by a time-varying gain or filtering of $y(n)$, *echo cancellation* aims at creating an echo estimate $\hat{x}_v(n)$

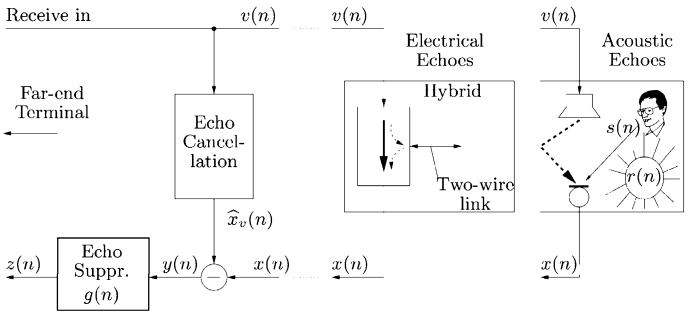


FIGURE 1 Cancellation and suppression of line echoes and acoustic echoes.

which minimizes $e(n) = x_v(n) - \hat{x}_v(n)$ subject to a given error criterion. Obviously, $e(n)$ is identical to y and as such can be directly observed if $x_s(n) = x_r(n) = 0$ can be assumed.

1.1 Echo Suppression

In its simplest version, an echo suppressor operates as a voice-activated switch: Whenever the far-end speaker is active, the echo suppressor applies zero gain $g(n) = 0$, so that feedback of $x_v(n)$ in $z(n)$ is completely suppressed. Obviously, this reduces full-duplex communication to a half-duplex mode, as during *double-talk*, i.e. when both $v(n)$ and $s(n)$ are active, the gain must be zero and the desired signal $s(n)$ will not be transmitted. To mitigate this undesirable effect, the gain control will aim at a compromise between attenuation of the echo $x_v(n)$ and gain for the desired signal component $x_s(n)$. Typically, short-time estimates for the levels of $x_v(n)$, $x_s(n)$ and $x_r(n)$ are extracted from $x(n)$ to support a proper choice of $g(n)$. For this, short-term variation of the speech signal's envelope, stationarity assumptions for the background noise $r(n)$ and correlation between $v(n)$ and $x_v(n)$ can be exploited. Fast and reliable detection of double-talk in $x(n)$ is crucial to subjective quality to avoid suppression of initial low-energy fricatives in $x_s(n)$, which is especially difficult for noisy signals, i.e. with high levels of $x_r(n)$. Besides linear gain control, nonlinear processing for echo suppression has also been introduced in various forms, most notably, centre-clipping [2]. Over the years, the control of echo suppression has reached a high level of sophistication to overcome its structural deficiencies [3]. As a common refinement in current systems, the gain control $g(n)$ is applied individually in frequency subbands or DFT (*Discrete Fourier Transform*) bins, especially for *residual echo suppression*, i.e. as a complement to acoustic echo cancellation (see below). For that, not only the above traditional echo suppression techniques but also the advanced noise reduction

schemes (see below) are employed, handling residual echoes with similar techniques as background noise [4,3].

1.2 Echo Cancellation

Echo cancellation is commonly viewed as a system identification problem where the echo path is mostly modelled by a linear finite impulse response (FIR) filter.

For electrical echoes, the duration of the corresponding impulse reponse varies strongly, mostly depending on the delays for the hybrids where the echoes are caused. With an increasing amount of digital signal processing introducing extra delay in modern telecommunication networks, link delays are increasing and, in addition to the long-standing satellite links with one-way delays in the order of 250 ms, more and more links involve large delays, especially if packet-switched networks and mobile phones are involved. As the echo path of the actual hybrid is well modelled by a relatively short impulse response (covering not more than 50 ms), the impulse response modelling an entire link is usually sparse, i.e. many samples exhibit only small values.

The acoustic echo path is essentially determined by the reverberation in the local acoustic environment, which is commonly captured by the reverberation time T_{60}, which is the time interval for the sound level to decay by 60 dB in a room after the source is muted [5]. This means for attenuating the acoustic echo by x dB, roughly $L = x \cdot f_s \cdot T_{60}/60$ FIR filter coefficients are needed, where f_s is the sampling frequency [3]. Thus, in a well-damped office room, e.g. $L = 25 \cdot 8\,\text{kHz} \cdot 300\,\text{ms}/60 = 1,000$ coefficients are needed for basic telephone bandwidth if 25 dB echo attenuation should be achievable (Figure 2).

1.2.1 The System Identification Problem

The identification of a linear FIR filter with impulse response $\hat{h}_k(n)$ of length L, with $k = 0, \ldots, L-1$ indexing the impulse response samples, is commonly solved via minimizing a mean square error criterion[1]

$$\varepsilon\{e^2(n)\} \rightarrow \min, \tag{2}$$

where ε is the expectation operator and

$$e(n) = \hat{h}_k(n) \times v(n) - x_v(n). \tag{3}$$

[1]For notational convenience, only real-valued signals are considered here, the straightforward generalization to complex signals is given, e.g. in [6].

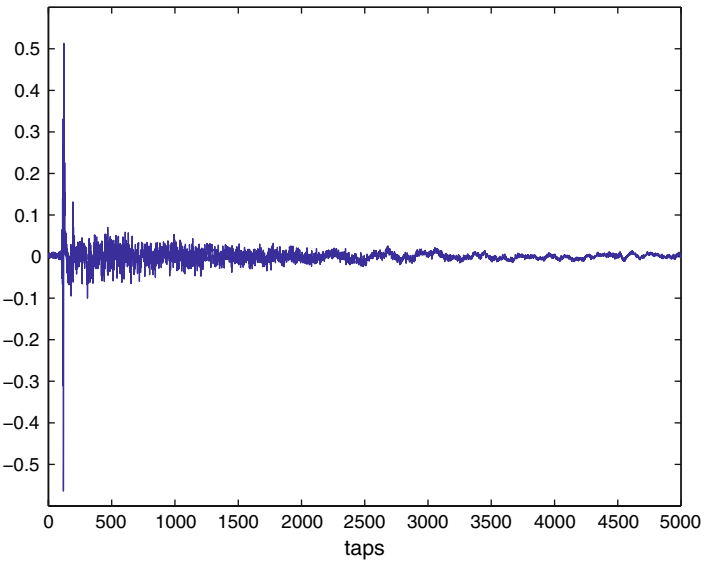


FIGURE 2 Impulse response of acoustic echo path in an office ($5.5\,\mathrm{m} \times 3\,\mathrm{m} \times 2.8\,\mathrm{m}$, $T_{60} \approx 300\,\mathrm{ms}$, sampling frequency $f_s = 12\,\mathrm{kHz}$).

Assuming (at least wide-sense) stationary processes $v(n), x(n)$ and a time-invariant impulse response $\hat{h}_k(n) = \hat{h}_k$, the solution is given by the Wiener filter [6]

$$\hat{\mathbf{h}} = \mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{r}_{x\mathbf{v}} \tag{4}$$

where $\hat{\mathbf{h}} = \left(\hat{h}_0, \ldots, \hat{h}_{L-1}\right)^{\mathrm{T}}$ is the vector of the FIR filter coefficients, $\mathbf{R}_{\mathbf{vv}}$ is the $L \times L$ autocorrelation matrix and $\mathbf{r}_{x\mathbf{v}}$ is the $L \times 1$ crosscorrelation vector, with elements $\mathbf{R}_{\mathbf{vv},ij} = \varepsilon\{v(n-i)v(n-j)\}$ and $\mathbf{r}_{x\mathbf{v}} = \varepsilon\{x(n)v(n-j)\}(i,j = 0, 1, \ldots, L-1)$, respectively $((\cdot)^{\mathrm{T}}$ denotes transposition).

In the given context, no signal ensembles for computing expectations are available, but only single finite observations, which requires deterministic algorithms. Corresponding to Eq. 2, we formulate a least squares criterion as the weighted squared error signal $e^2(n)$ for a set of observable samples:

$$J(n) = \sum_{k=0}^{n} w(k,n)e^2(n-k) \rightarrow \min. \tag{5}$$

Here, $w(k,n)$ denotes a windowing function with $0 \leq w(k,n) \leq 1$. This leads to a least squares solution of the form [6,7]

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{R}}_{\mathbf{vv}}^{-1}(n)\hat{\mathbf{r}}_{x\mathbf{v}}(n), \tag{6}$$

where the required estimates $\hat{\mathbf{R}}_{\mathbf{vv}}(n)$, $\hat{\mathbf{r}}_{x\mathbf{v}}(n)$ obviously depend on the window $w(k,n)$ in Eq. 5.

In the echo cancellation context, one has to account for the time-variance of the echo path, which especially for acoustic echoes may be drastic and unpredictable. Thus, *adaptive filters* are employed which try to identify the above optimum filters, by using update equations of the form:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \Delta\hat{\mathbf{h}}(n). \tag{7}$$

For defining the update term, two philosophies are in common use: For one, gradient techniques iteratively approximate the minimum of the cost function, while others directly aim at the optimum solution at time n as given by Eq. 6.

For the well-known and widely studied *Least Mean Square (LMS) algorithm* [8,6], the update term $\Delta\hat{\mathbf{h}}(n)$ is obtained as the negated gradient of Eq. 5 for $w(k,n) = \delta(k)$ (with $\delta(k)$ as the unit impulse [9]), multiplied by a stepsize $\mu/2$, and reads with $\mathbf{v}(n) = (v(n), \ldots, v(n-L+1))^{\mathrm{T}}$

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu\mathbf{v}(n)e(n). \tag{8}$$

This algorithm was suggested already in the first publication on acoustic echo cancellation [10] and its normalized version, the *Normalized Least Mean Square (NLMS) algorithm* ($\mu = \alpha/(\mathbf{v}^{\mathrm{T}}(n)\mathbf{v}(n)), 0 \le \alpha \le 2$) is still the most commonly used algorithm in real-world applications:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \alpha\frac{\mathbf{v}(n)e(n)}{\mathbf{v}^{\mathrm{T}}(n)\mathbf{v}(n)}. \tag{9}$$

The popularity of the NLMS algorithm is based not only on its simplicity and low complexity (roughly $2L$ multiplications/additions per sampling interval) but also on its robustness against changing signal statistics. The price for the simplicity and robustness of the LMS/NLMS-type algorithms is a relatively slow convergence for correlated signals $v(n)$, such as speech in the given context. Among the many variations of the LMS algorithms, the *Proportionate NLMS* algorithm [11,7] is of special interest for sparse impulse responses (as common for electrical echoes) because it allows for fast adaptation of the largest coefficients of the impulse response.

On the other hand, the *Recursive Least Squares (RLS) algorithm*[12] guarantees the optimum solution of Eq. 6 at each time instant n. Here, the update term recursively incorporates new observations $v(n)$ and $e(n)$ and adapts the solution to the

changing cost function value $J(n)$. Choosing an exponential weighting, $w(k,n) = \lambda^{n-k}, \lambda < 1$, the update equation can be written as [6]:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \hat{\mathbf{R}}_{\mathbf{vv}}^{-1}(n)\mathbf{v}(n)e(n), \tag{10}$$

with

$$\hat{\mathbf{R}}_{\mathbf{vv}}(n) = \lambda \hat{\mathbf{R}}_{\mathbf{vv}}^{-1}(n-1) + v(n)\mathbf{v}^{\mathrm{T}}(n). \tag{11}$$

Although the inversion of the autocorrelation matrix at each sampling instant n is usually circumvented by the matrix inversion lemma [6], the computational complexity of the RLS is still of $\mathcal{O}(L^2)$ which is prohibitive for long FIR filters. The so-called *Fast RLS* algorithms or *Fast Transversal Filters* (*FTF*) exploit the transversal filter structure versions to further reduce computational load [12]. However, robustness is a crucial issue for RLS algorithms: Regularization for $\hat{\mathbf{R}}_{\mathbf{vv}}(n)$ or its inverse has to account for strongly correlated signals (leading to a large eigenvalue spread) as well as for speech pauses, and, moreover, the choice of λ needs to balance tracking capability versus stability of convergence [13].

The *Affine Projection (AP) algorithm* [14] can be viewed as a compromise between LMS-type algorithms and RLS algorithms. The detrimental effect of signal correlation on LMS-type algorithms, as, e.g. given for autoregressive (AR) processes $v(n)$ such as speech, is removed by the AP algorithm up to a given AR process order K at the cost of an increased computational complexity of $\mathcal{O}(KL)$. As with the RLS algorithm, fast versions have been proposed [15,16], which require additional stabilization measures.

To reconcile the requirements of computational efficiency and fast and robust convergence for long adaptive FIR filters, a transformation of the filter adaptation in Eq. 10 seems attractive so that ideally the inversion of the autocorrelation matrix reduces to the inversion of a diagonal matrix. Obviously, the signal-dependent Karhunen–Loève transform [6] would be an ideal choice. In practice, signal-independent transforms into a frequency subband domain or into the DFT domain proved to be reasonably close and efficient approximations. Two classes of practically relevant concepts resulted from this idea:

Frequency subband filtering. The oversampled frequency subband approach [17,18,19] decomposes $v(n)$ and $x(n)$ into N frequency subbands (Figure 3) such that aliasing after downsampling by a factor $R < N$ in each subband is sufficiently small while overall analysis/synthesis filterbanks are still sufficiently close to

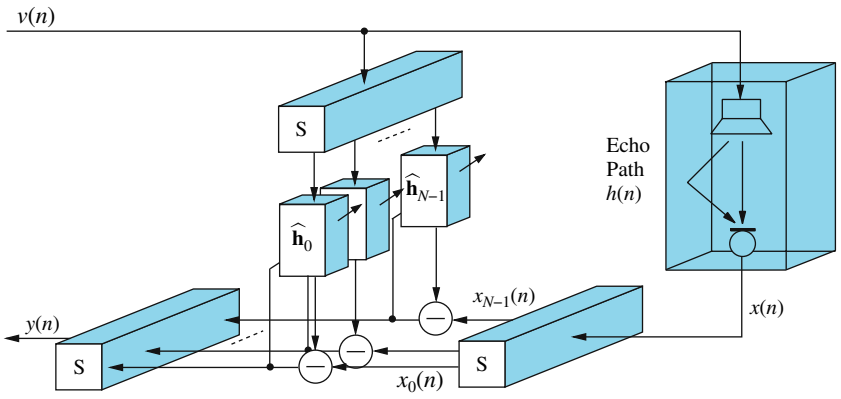


FIGURE 3 Frequency subband acoustic echo cancellation.

perfect reconstruction [20]. Then, adaptive subband filters can operate independently in each frequency subband. These filters are shorter than the equivalent fullband filter due to the reduced sampling rate and also assure a faster and more robust overall convergence for low-complexity adaptation algorithms such as the NLMS [17]. A significant computational gain is achieved over the corresponding fullband system due to the reduced sampling rate by R in the subbands, which implies both reduced clock rate for filtering and reduced filter length. In principle, all adaptive filtering algorithms for fullband signal processing can be implemented in each subband, with the additional advantage that data can be shared among different subbands, if desired, e.g. for control strategies. The filtering cost for analysis and synthesis is minimized if efficient uniform filterbanks based on polyphase structures [20] are used so that only the additional signal delay due to the filterbanks remains as a potential disadvantage in practical applications.

Frequency-Domain Adaptive Filtering (FDAF). This second class of transform-domain adaptive filtering is directly based on the DFT [21,22] and was initially motivated by the advantages of performing linear filtering by multiplication of DFT bins instead of convolving time-domain sequences (*Fast Convolution*). This computational advantage carries over to adaptive filtering and is accompanied by an almost-diagonalization of the autocorrelation matrix $\hat{\mathbf{R}}_{\mathbf{vv}}$ in the update equation, Eq. 10, if an RLS-based formulation is used [23,7]. The block delay associated to processing successive DFT data frames can be reduced at the cost of increased computational load by using a large overlap of the DFT frames or by using *multidelay filters (MDF)* [24,25] (also,

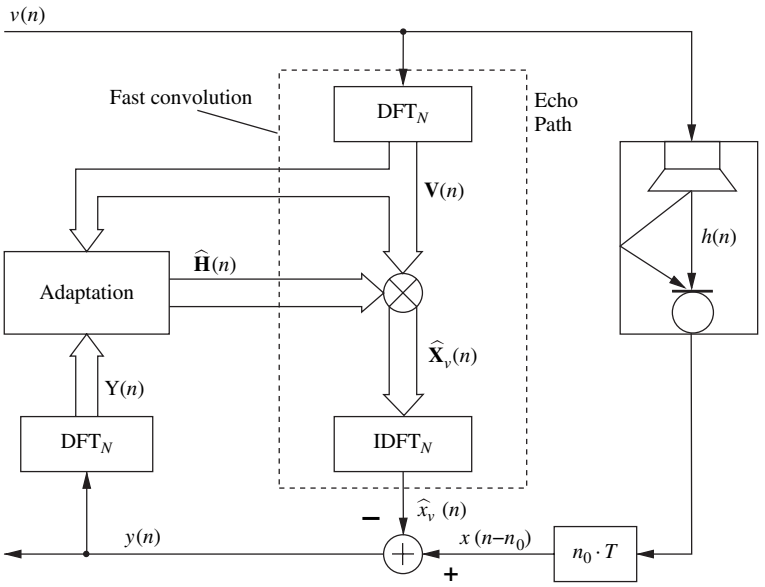


FIGURE 4 Acoustic echo cancellation using FDAF.

partitioned block frequency-domain adaptive filters (PBFDAF)), where the impulse response of the echo path model is partitioned into several segments, so that the DFT length is reduced. As a generalization to the MDF, the *extended multidelay filter (EMDF)* accounts for the signal correlation between partitions and thereby approximates both the low complexity and the fast convergence of the FDAF without the large delay, which is especially attractive for very high-order echo path models [26] (Figure 4).

1.2.2 Controlling the Adaptation

Using the above adaptive filtering algorithms, the echo path can only be identified as long as the signal components $x_s(n)$ and $x_r(n)$ do not interfere significantly with the adaptation, so that $y(n) = x(n) - \hat{x}_v(n)$ is essentially equal to the system identification error $e(n)$. However, typically permanent background noise $x_r(n)$ and intermittent and unpredictable local talker activity $x_s(n)$ call for additional control measures. Aside from adaptation under permanent noise, the distinction between a change of the echo path and the onset of local talker activity, which both lead to large signal values for $x(n)$, presents the main problem: In the first case, adaptation should be aggressive, whereas in the second case, it should be stalled. This so-called *double-talk detection* problem [23,7] can be addressed by deriving some estimate for the short-time crosscorrelation between $v(n)$ and

$x(n)$, and by a two-path model which compares the effect of the updated filter to the previous version [27]. For limiting divergence in case of detection errors, robust statistics has been introduced as an elegant and effective extension to various of the above adaptive filtering algorithms [23,7,28]. On the other hand, especially for (N)LMS-based algorithms empirical stepsize control methods for maximizing convergence speed while minimizing the risk of divergence have been developed to a high level of sophistication [3].

1.2.3 Complementing the Linear FIR Filter Model

The adaptive FIR filter will in general not be able to model the echo path perfectly at all times, not only because of limited tracking capability of the adaptation algorithms but also because the implemented filters may be too short, or the echo path is not strictly linear. Often, an additional echo suppression mechanism is implemented for the residual echo (see above). In case of severe nonlinearities in the echo path as they occur, e.g. in mobile telephone terminals due to saturated amplifiers and overloaded loudspeakers, efficient nonlinear echo path models are desirable [29,30,31].

1.3 Multichannel Acoustic Echo Cancellation

Since stereo reproduction has been invented, users appreciate multichannel sound reproduction. For hands-free telecommunications and human/machine dialogues, this introduces additional fundamental problems for acoustic echo cancellation, as is illustrated for the stereo case in Figure 5. Obviously, each of the microphones at the local terminal simultaneously records echoes of both loudspeaker signals $v_1(n), v_2(n)$. To identify the two echo path models $\hat{\mathbf{h}}_1(n), \hat{\mathbf{h}}_2(n)$ with only one error signal derived from $x_1(n)$ requires that $v_1(n)$ and $v_2(n)$ contain uncorrelated components [32,33]. In fact, the optimum stereo acoustic echo cancellation system can be written in the same way as Eq. 6 with the input data and filter coefficient vectors of the single-channel case replaced by a concatenation of the two-channel input data and filter coefficient vectors for the two channels, e.g. $\mathbf{v}(n) = (\mathbf{v}_1^{\mathrm{T}}(n), \mathbf{v}_2^{\mathrm{T}}(n))^{\mathrm{T}}$. Accordingly, all the above adaptive filtering algorithms can be formulated for the P-channel case [7]. The usually strong crosscorrelation between the reproduction channel leads however to a much greater matrix condition number for the resulting PL × PL autocorrelation matrix $\mathbf{R}_{\mathbf{vv}}$ which now also contains the crosscorrelations between the channels in the off-diagonal $L \times L$ block matrices. In the typical stereo communication scenario of Figure 5, the two reproduction channels are recorded from a single source, so that the microphone signals $v_1'(n), v_2'(n)$ are essentially convolutional products

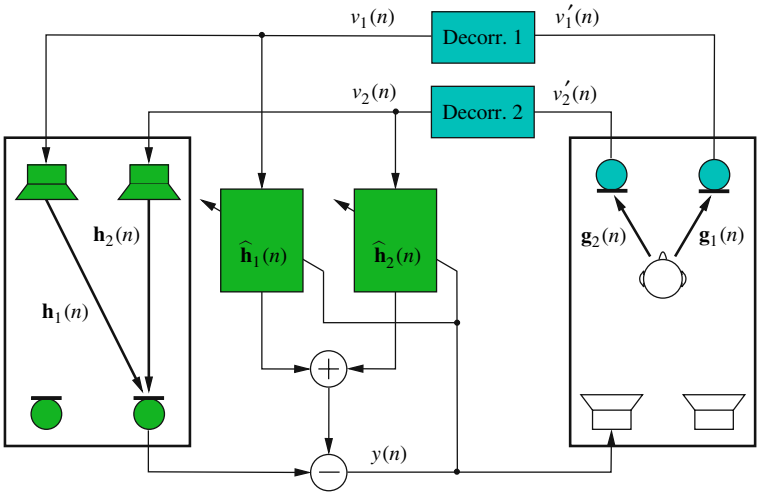


FIGURE 5 Scenario for stereo acoustic echo cancellation.

derived from a single signal source and therefore strongly correlated. To reduce this correlation, time-varying filtering, nonlinear distortion and/or injection of noise can be applied to $v_1'(n), v_2'(n)$, which must be different for each channel. Obviously, the cost is some distortion in the signals $v_1(n), v_2(n)$ which should be minimized. At the same time, this distortion should provide significant energy for uncorrelated components to sufficiently support the system identification. For the given options, time-varying allpass filtering [34] must be used diligently if spatial realism should not be impaired, noise injection should be limited to the masking thresholds of the auditory system [35] and nonlinear distortion should also remain inaudible [36].

For computationally inexpensive time-domain algorithms like the NLMS, the increased condition number for $\mathbf{R}_{\mathbf{vv}}$ directly reduces the convergence speed. Thus, the demand for diagonalization of the correlation matrix becomes even stronger for multichannel systems, while at the same time the number of filter coefficients is multiplied by the channel number, so that an RLS implementation in the time-domain becomes increasingly unrealistic. As a consequence, both the frequency subband concept and the DFT domain approach become even more attractive and have been studied extensively [7, chapters 5–9]. While the subband approach has been refined to product maturity for the stereo teleconferencing application [7], the multichannel FDAF algorithm has been practically verified even for surround systems with five reproduction channels [37].

References

1. M.M. Sondhi and D.A. Berkley. Silencing echoes on the telephone network. *Proceedings of the IEEE*, 68(8):948–963, August 1980.
2. O.M.M. Mitchell and D.A. Berkley. A full-duplex echo suppressor using center-clipping. *Bell System Technical Journal*, 50(3): 1619–1630, May/June 1971.
3. E. Hänsler and G. Schmidt. *Acoustic Echo and Noise Control: A Practical Approach*. Wiley, Hoboken, NJ, 2004.
4. S. Gustafsson, R. Martin, and P. Vary. Combined acoustic echo cancellation and noise reduction for hands-free telephony. *Signal Processing*, 64(1):21–32, 1998.
5. H. Kuttruff. *Room Acoustics*. Spon Press, London, 4th edition, 2000.
6. S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, NJ, 4th edition, 2002.
7. J. Benesty, T. Gänsler, D.R. Morgan, M.M. Sondhi, and S.L. Gay. *Advances in Network and Acoustic Echo Cancellation*. Springer, Berlin, 2001.
8. B. Widrow and M.E. Hoff. Jr. Adaptive switching circuits. In *IRE Wescon Conf. Rec., Part 4*, pp. 96–104, 1960.
9. A.V. Oppenheim and R.W. Schafer. *Digital Signal Processing*. Prentice Hall, Englewood Cliffs, NJ, 1975.
10. M.M. Sondhi. An adaptive echo canceler. *Bell System Technical Journal*, 46:497–511, March 1967.
11. D.L. Duttweiler. Proportionate normalized least-mean-squares adaptation in echo cancelers. *IEEE Transactions on Speech and Audio Processing*, 8(5):508–518, September 2000.
12. J.M. Cioffi and T. Kailath. Fast, recursive-least-squares transversal filters for adaptive filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 32(2):304–337, April 1984.
13. D.T.M. Slock and T. Kailath. Numerically stable fast transversal filters for recursive least squares adaptive filtering. *IEEE Transactions on Signal Processing*, 39(1):92–114, January 1991.
14. K. Ozeki and T. Umeda. An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. *Electronic Communications Japan*, 67-A(5):19–27, January 1984.
15. M. Tanaka, Y. Kaneda, S. Makino, and J. Kojima. Fast projection algorithm and its step size control. In *Proceedings of Internation Conference on Acoustics, Speech, and Signal Processing*, pp. 945–948, Detroit, MI, May 1995. IEEE.
16. S.L. Gay and S. Tavathia. The fast affine projection algorithm. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pp. 3023–3026, Detroit, MI, May 1995. IEEE.
17. W. Kellermann. Analysis and design of multirate systems for cancellation of acoustical echoes. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pp. 2570–2573, New York, April 1988. IEEE.
18. M.M. Sondhi and W. Kellermann. Echo cancellation for speech signals. In S. Furui and M.M. Sondhi, editors, *Advances in*

Speech Signal Processing, chapter 11, pp. 327–356. Marcel Dekker, New York, 1991.
19. S. Gay and J. Benesty, editors. *Acoustic Signal Processing for Telecommunications*. Kluwer, Norwell, MA, 2000.
20. P.P. Vaidyanathan. *Multirate System and Filter Banks*. Prentice Hall, Englewood Cliffs, NJ, 1993.
21. D. Mansour and A.H. Gray Jr. Unconstrained frequency domain adaptive filter. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(5):726–734, October 1982.
22. J.J. Shynk. Frequency-domain and multirate adaptive filtering. *IEEE Signal Processing Magazine*, 9(1):14–37, January 1992.
23. S.L. Gay and J. Benesty, editors. *Acoustic Signal Processing for Telecommunication*. Kluwer, Boston, MA, 2000.
24. J.-S. Soo and K.K. Pang. Multidelay block frequency domain adaptive filter. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(2):373–376, February 1990.
25. E. Moulines, O.A. Amrane, and Y. Grenier. The generalized multidelay adaptive filter: Structure and convergence analysis. *IEEE Transactions on Signal Processing*, 43(1):14–28, January 1995.
26. H. Buchner, J. Benesty, and W. Kellermann. An extended multidelay filter: Fast low-delay algorithms for very high-order adaptive systems. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pp. 385–388, Hongkong, China, April 2003 IEEE.
27. K. Ochiai, T. Araseki, and T. Ogihara. Echo canceler with two echo path models. *IEEE Transactions on Communications*, 25(6):589–595, June 1977.
28. H. Buchner, J. Benesty, and W. Kellermann. An outlier-robust extended multidelay filter with application to acoustic echo cancellation. In *Proceedings of International Workshop on Acoustic Echo and Noise Control*, pp. 19–22, Kyoto, Japan, September 2003.
29. A. Stenger and W. Kellermann. Adaptation of a memoryless preprocessor for nonlinear acoustic echo cancelling. *Signal Processing*, 80:1747–1760, Sep 2000.
30. F. Küch and W. Kellermann. Nonlinear line echo cancellation using a simplified second-order Volterra filter. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, May 2002. IEEE.
31. F. Kuech and W. Kellermann. Partitioned block frequency-domain adaptive second-order volterra filter. *IEEE Transactions on Signal Processing*, 53(2):564–575, February 2005.
32. M.M. Sondhi, D.R. Morgan, and J.L. Hall. Stereophonic echo cancellation: An overview of the fundamental problem. *IEEE Signal Processing Letters*, 2(8):148–151, August 1995.
33. J. Benesty, D.R. Morgan, and M.M. Sondhi. A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation. *IEEE Transactions on Speech and Audio Processing*, 6(2):156–165, March 1998.

34. Y. Joncours and K. Sugiyama. A stereo echo canceller with preprocessing for correct echo-path identification. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pp. 3677–3680, Seattle, WA, 1998. IEEE.
35. T. Gänsler and P. Eneroth. Influence of audio coding on stereophonic acoustic echo cancellation. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pp. 3649–3652, Seattle, WA, 1998. IEEE.
36. D.R. Morgan, J.L. Hall, and J. Benesty. Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation. *IEEE Transactions on Speech and Audio Processing*, 9(6), September 2001.
37. H. Buchner, J. Benesty, and W. Kellermann. Multichannel frequency-domain adaptive filtering with application to acoustic echo cancellation. In J. Benesty and Y. Huang, editors, *Adaptive Signal Processing: Application to Real-World Problems*, chapter 3. Springer, Berlin, January 2003.

47

Noise Reduction and Interference Cancellation

Walter Kellermann

Multimedia Communications and Signal Processing, University Erlangen-Nuremberg, Erlangen, Germany, e-mail: wk@LNT.de

Noise reduction and interference cancellation differ structurally from echo cancellation insofar as no undisturbed reference signal (such as the loudspeaker signal for AEC) is available, neither for the desired signal nor for the unwanted noise or interference component. It can be viewed as a signal separation problem where estimates of desired and undesired signal components have to be found and separated from each other.

Based on the number of used sensor (i.e. microphone) signals $x_\nu(n)$, we distinguish between single-channel and multichannel techniques. With single-channel techniques typically no reliable phase information for the various signal components can be extracted, whereas multichannel techniques implicitly or explicitly exploit phase differences between the various sensor signal components in order to perform spatial filtering or to separate signal components by their different spatial coherence.

1 • Single-Channel Techniques

Assuming a microphone signal $x(n)$ which contains the desired (speech) signal component $s(n)$ plus some unwanted noise or interference $r(n)$ $(x_s(n) = s(n), x_r(n) = r(n)$ in Figure 1 of Chapter 46), we seek an optimum filter for suppressing the noise $r(n)$ without distorting $s(n)$ unduly. If the optimization criterion is the minimum mean square error (MSE) between the filter output and the desired signal $s(n)$ (equivalent to maximum signal-to-noise power ratio (SNR)) for at least weakly stationary und mutually orthogonal signals $s(n)$, $r(n)$, Wiener filter theory [1] yields the corresponding optimum filter, which in its noncausal version is given in the frequency domain for discrete-time signals by

$$H_{\mathrm{W}}(\mathrm{e}^{\mathrm{j}\omega}) = \frac{S_{ss}(\mathrm{e}^{\mathrm{j}\omega})}{S_{rr}(\mathrm{e}^{\mathrm{j}\omega}) + S_{ss}(\mathrm{e}^{\mathrm{j}\omega})} = 1 - \frac{S_{rr}(\mathrm{e}^{\mathrm{j}\omega})}{S_{rr}(\mathrm{e}^{\mathrm{j}\omega}) + S_{ss}(\mathrm{e}^{\mathrm{j}\omega})}$$
$$= 1 - \frac{S_{rr}(\mathrm{e}^{\mathrm{j}\omega})}{S_{xx}(\mathrm{e}^{\mathrm{j}\omega})}. \tag{1}$$

Here, $S_{xx}(\mathrm{e}^{\mathrm{j}\omega})$, and $S_{ss}(\mathrm{e}^{\mathrm{j}\omega})$, and $S_{rr}(e^{j\omega})$ are the power spectral densities (PSDs) of $x(n)$, $s(n)$, and $r(n)$, respectively, and ω is the normalized frequency.

Alternatively, under the same assumptions for $s(n)$, $r(n)$, a spectral subtraction seems reasonable, $S_{ss}(\mathrm{e}^{\mathrm{j}\omega}) = S_{xx}(\mathrm{e}^{\mathrm{j}\omega}) - S_{rr}(\mathrm{e}^{\mathrm{j}\omega})$, to obtain the PSD of the desired speech [2]. Formulating this as a filtering operation yields

$$S_{ss}(\mathrm{e}^{\mathrm{j}\omega}) = \left(1 - \frac{S_{rr}(\mathrm{e}^{\mathrm{j}\omega})}{S_{xx}(\mathrm{e}^{\mathrm{j}\omega})}\right) \cdot S_{xx}(\mathrm{e}^{\mathrm{j}\omega}) =: |H_{\mathrm{SS}}(\mathrm{e}^{\mathrm{j}\omega})|^2 \cdot S_{xx}(\mathrm{e}^{\mathrm{j}\omega}), \tag{2}$$

so that obviously $\left|H_{\mathrm{SS}}(\mathrm{e}^{\mathrm{j}\omega})\right| = \sqrt{H_{\mathrm{W}}(\mathrm{e}^{\mathrm{j}\omega})}$. Integrating both versions, a *generalized Wiener filter* can be defined as follows:

$$H_{\mathrm{WG}}(\mathrm{e}^{\mathrm{j}\omega}) := \left[1 - \left(\frac{S_{rr}(\mathrm{e}^{\mathrm{j}\omega})}{S_{xx}(\mathrm{e}^{\mathrm{j}\omega})}\right)^{\beta}\right]^{\alpha}. \tag{3}$$

Aside from Wiener filtering $(\alpha = 1, \beta = 1)$ and PSD subtraction $(\alpha = 0.5, \beta = 1)$, this includes also the so-called spectral magnitude subtraction $(\alpha = 1, \beta = 0.5)$ [3,4] and forms the basis for several other variations [5].

In noise reduction for speech enhancement, the PSDs $S_{rr}(\mathrm{e}^{\mathrm{j}\omega})$ and $S_{xx}(\mathrm{e}^{\mathrm{j}\omega})$ must be estimated on a short-term basis due to the nonstationarity of the speech signal $s(n)$, so that for a length-N data window taken at time n, only the discrete frequencies

$\omega_\nu = 2\pi\nu/N, \nu = 0, \ldots, N-1$, must be considered. For estimating the noise PSD, two methods are well established: Either it is estimated only during speech pauses, which requires voice activity detection, or it is continually estimated by tracking spectral minima [6]. In both cases, the desired noise estimates must be computed by averaging measured short-time PSDs over time, whereas the PSD estimate of $x(n)$ cannot be averaged in the same way, as the nonstationarity of $s(n)$ must be followed. This mismatch of time constants for the PSD estimation leads to the so-called *musical tones*, which are decisive for subjective quality. Usually some amount of noise power suppression is sacrificed to better avoid musical tones. Besides reducing α, β, the introduction of a minimum spectral gain $\hat{H}_{\min}(n, \omega_\nu)$ (*spectral floor*) is common, so that the time-varying estimated generalized Wiener filter can be written in the DFT domain as

$$\hat{H}_{\mathrm{WG}}(n, \omega_\nu) = \max \left\{ \left[1 - \left(\frac{\hat{S}_{rr}(n, \omega_\nu)}{\hat{S}_{xx}(n, \omega_\nu)} \right)^{\beta} \right]^{\alpha}, \hat{H}_{\min}(n, \omega_\nu) \right\}, \quad (4)$$

where $\hat{S}(n, \omega_\nu)$ represents the νth DFT bin of the corresponding PSD estimate at time n. Obviously, the larger the $\hat{H}_{\min}(n, \omega_\nu)$, the less the noise attenuation.

While the above filters are all based on broadband optimization criteria, another widely used scheme directly minimizes the mean square spectral magnitude error for each DFT bin or its logarithm [7,8] assuming normally distributed DFT bin values for both speech and noise. This approach involves careful smoothing of SNR estimates and, with proper parametrization, has been found to be relatively robust against musical noise.

The optimum choice of the various parameters strongly depends on underlying assumptions for the noise component and for the acceptable level of musical tones. As a general tendency, the less stationary the noise is, the more fragile the noise reduction methods are. This, especially for human listeners, recommends larger values for $\hat{H}_{\min}$ to preserve the desired speech integrity and maintain naturalness of the background noise while accepting higher noise power levels.

Note also that the above filters (Eqs. (3) and (4)) are all linear-phase filters due to the even symmetry of the underlying PSD estimates. This implies that the phase of the estimated desired signal is still the phase of the measured noisy signal $x(n)$, except for some additive linear-phase term.

In addition to removing the noise PSD, significant research on speech enhancement was dedicated to exploiting structural knowledge about the speech signal to improve perceptual speech

quality. One of the key ideas is here to exploit the harmonic structure of speech and emphasize spectral components corresponding to the pitch frequency and its higher harmonics [9] or the poles of an estimated all-pole model for speech production [10]. The latter can then also be connected to pattern matching techniques as common in speech recognition (see [11] and references therein).

2 • Multichannel Techniques

The advantages of multichannel techniques over single-channel techniques all originate from the exploitation of spatial diversity as given by the simultaneous sampling of sound fields at different locations. Essentially, the available spatial information provides another domain for separating desired from undesired signals aside from time and frequency.

According to the setup of Figure 1, the M microphone signals $x_\mu(n)$ contain differently filtered (or delayed) versions $x_{\mu,s}(n)$ of a desired source signal $s(n)$ and undesired noise components $x_{\mu,r}(n)$. The filtering by $\mathbf{w}_\mu(n)$ usually aims at extracting the desired signal from the set of microphone signals and at eliminating all undesired components $x_{\mu,r}(n)$ while allowing for some propagation and processing delay $n_0 \geq 0$, so that $y(n) \approx s(n - n_0)$. For this, ideally, all the desired source components should be time-aligned and added constructively, whereas all undesired components should be cancelled out.

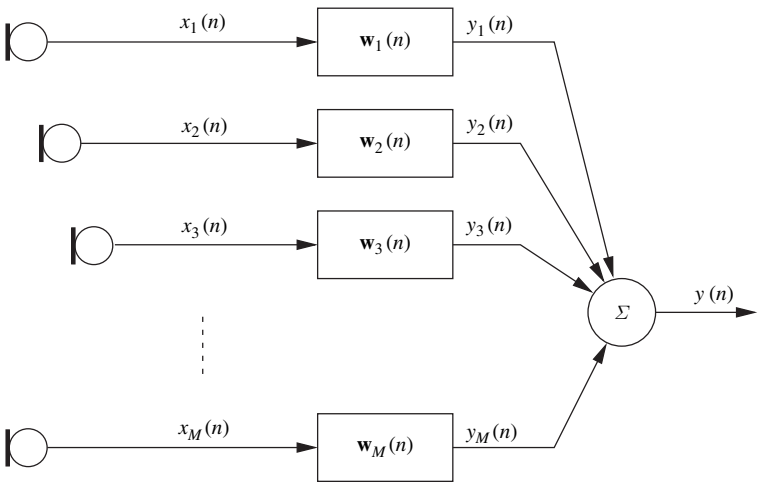


FIGURE 1 Multichannel recording by a microphone array for noise reduction and interference cancellation.

2.1 Data-Independent Beamforming

In a basic version, the differences between propagation delays for the direct paths from source to microphones are estimated – or assumed based on a priori knowledge – and then compensated by $\mathbf{w}_\mu$ prior to summation. It can be shown that with such a *delay and sum beamformer* the power of noise and interference which appears uncorrelated and with identical level at the M sensors can be suppressed by a factor M, commonly expressed as *white noise gain* of $10 \log_{10} M$ dB. For a given *look direction*, the filters $\mathbf{w}_\mu$ are here time invariant and only realize a (possibly fractional) delay, whereas in the more general case of a *filter and sum beamformer* they will also exhibit frequency-dependent and possibly time-variant magnitude characteristics.

A characteristic example for the spatial selectivity for a signal-independent *filter and sum beamformer* is shown in Figure 2. It shows the amplitude gain factor of a linear array using $M = 8$ sensors with a uniform spacing of $d = 4\,\text{cm}$ as a function of frequency and the angle of incidence of a plane wave (*far-field assumption*) in a horizontal half-plane. The look direction of the beamforming is toward *broadside*, that is, the array provides a distortionless response for plane waves impinging perpendicularly to the array axis (at 0 degrees), whereas plane waves from other directions are suppressed depending on their angle of incidence. In this example, some fundamental properties of beamforming for wideband audio signals using microphone arrays are illustrated (see also [12]): As beamforming corresponds to a spatial sampling, the spatial resolution is determined by the number of microphones and the spacing of the microphones relative to the signal wavelengths. Therefore, for a given array

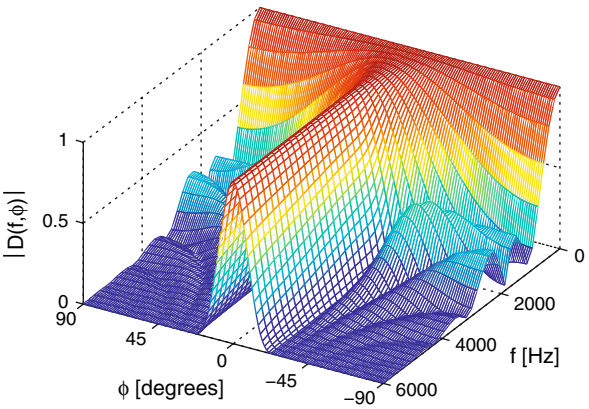


FIGURE 2 Spatial directivity of a linear array with filter and sum beamformer ($M = 8$ microphones with spacing 4 cm).

configuration, the spatial selectivity at low frequencies is less pronounced than for higher frequencies, as is clearly reflected in Figure 2. In order to avoid a second beam of increased sensitivity (*spatial aliasing*) at high frequencies in this half-plane, the microphone spacing may not exceed half a wavelength.

Due to the wideband nature of speech and audio signals which often span two to three orders of magnitude of wavelengths, traditional beamforming methods focussing on narrowband signals [13] have been extended to aim at frequency-independent spatial selectivity [14], which is reasonably well approximated in Figure 2 for the mainlobe for frequencies above 2,000 Hz.

In order to overcome the low spatial resolution at low frequencies, nested arrays can be used which realize different apertures with different spatial sampling for each octave of bandwidth [12]. Often, however, enlarging the aperture is not acceptable for the given application, so that differential (*superdirective*) beamforming becomes attractive which essentially relies on finite difference approximations of spatial derivatives of the sound field [15]. The approximation is valid as long as the acoustic wavelength is large relative to the sensor spacing and therefore can be used for conventional array geometries at the low frequencies. The price for the relatively small apertures is the sensitivity to spatially uncorrelated noise which so far limits the applicability for practical systems to third-order derivatives [15]. Another advanced concept for differential beamforming with small apertures (notably circular or spherical arrays) foresees a decomposition of the acoustic wavefield into eigenfunctions. Instead of the plain microphone signals, the contributions to the various modes are combined to achieve spatial selectivity, i.e. to realize a modal beamformer [16]. As an important advantage over classic conventional beamforming techniques, here, the spectral selectivity and the spatial selectivity can be designed independently.

2.2 Data-Dependent Beamforming

While the above beamformers are data-independent in that they do not exploit any knowledge about the actual signal properties and the positions of interferers or noise sources, data-dependent beamformers exploit such knowledge to optimally suppress noise and interference. Two optimization philosophies are generally followed: One is based on a mean square error (MSE) criterion aiming at maximizing SNR as a multichannel version of the single-channel Wiener filter. As opposed to the single-channel Wiener filter, the spatial domain is exploited to better estimate noise and interference or desired signal characteristics by measuring the spatial co-herence of the signals

[17,18]. In this multichannel Wiener filter approach, suppression of spatially incoherent noise can also efficiently be realized as a single-channel postfilter [19], with some unavoidable distortion of the desired signal as a consequence. The multichannel Wiener filter can also be the basis of subspace methods which essentially try to decompose the multichannel covariance matrices into two orthogonal subspaces for the desired signal and noise, respectively, by, e.g. using general singular value decomposition [18].

An alternative concept introduces a constraint in order to preserve the integrity of the desired signal, and, as in the popular case of the *linearly constrained minimum variance* (LCMV) beamformer [13], minimizes the output signal power subject to this constraint. As a special case, the *minimum variance distortionless response* (MVDR) beamformer demands a distortionless response for the assumed direction of the desired signal. To avoid the constrained optimization in practical applications, the *generalized sidelobe canceller* (GSC) [20] separates the problem into a signal-independent beamformer and a so-called blocking matrix for assuring the constraint, and an unconstrained optimization of the interference canceller (see Figure 3). The blocking matrix has to "block" all desired signal components, so that at its output only noise components appear, which are then used by the interference canceller to estimate the remaining noise component in the fixed beamformer's output. If desired signal components leak through the blocking matrix they will be interpreted as noise and subtracted from the fixed beamformer's output leading to highly undesirable cancellation of the desired signal. In practice, for robustness against source movements and acoustic echoes of the desired source signal, the blocking matrix is preferably adaptive, requiring a carefully designed adaptation control [21,22].

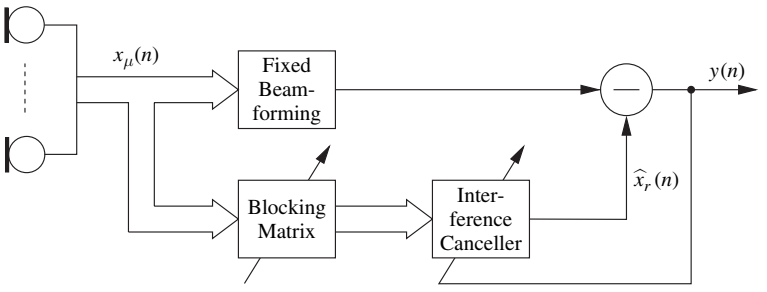


FIGURE 3 GSC with adaptive blocking matrix and adaptive interference canceller.

2.3 Adaptive Beamforming

In practical systems, signal-dependent optimum beamformers have to be adapted to the nonstationary charateristics of the involved speech and noise signals and the time-variant environment. The adaptive filtering algorithms as outlined for acoustic echo cancellation (from LMS to RLS, in time, frequency subband, and DFT domain) can directly be applied to beamforming once the the reference signals (filter inputs) and the desired signal have been defined. This is obvious for the multichannel interference canceller in the GSC, where the blocking matrix outputs play the same role as the loudspeaker inputs in the multichannel echo cancellation system, in the sense that they constitute reference input signals for identifying an optimum multichannel filter. Other adaptive beamformer algorithms face essentially the same estimation challenge as single-channel noise reduction: Reference signals that contain only noise (or desired signal) are difficult to obtain. However, if reliable spatial information on source directions is exploited properly, these references will always lead to better system performance than in the single-channel case [17,18,22].

2.4 Source Localization

If the above spatial filtering techniques require spatial knowledge on source positions which is not given a priori, localization algorithms are needed. A widely accepted classification [23] divides the current methods into "steered beamformers", "high-resolution spectral estimation", and "TDOA (*time difference of arrival*)-based methods". While "steered beamformers" essentially scan the spatial region of interest to measure the response power to localize a source, spectral estimation techniques may include statistical signal models and use eigenanalysis. TDOA-based methods are still dominant in the given applications: They rely on relative delays between two sensor signals which are typically estimated by the *generalized crosscorrelation* [24]. As a more advanced concept, the *adaptive eigenvalue decomposition* [25] extracts the TDOA estimates from estimated impulse responses for the propagation paths from the sources to the microphones and can be expected to perform better in moderately reverberant environments. In general, reflections and reverberation affect all of the above localization algorithms severely, as these signal components are acting as interference correlated to the desired component.

3 • Blind Signal Separation of Convolutive Mixtures

While the above multichannel techniques assume some knowledge about the position of one desired source and try to suppress signals from all other directions, blind source separation (BSS) effectively performs a blind beamforming by

identfying linear filters that minimize statistical dependency between several desired signals [26–28]. Thereby, several desired source signals can be extracted simultaneously from the given convolutive mixtures at the microphones, which record a sum of all source signals each convolved with the corresponding acoustic impulse response from the source to the microphone.

Comparing the signal paths in the 2-by-2 case in Figure 4 to the GSC in Figure 3 the structural similarity can be illustrated: BSS can be viewed as performing interference cancellation for each output signal, with the respective other output being the interferer. A fundamental difference to known interference cancellers is given by the optimization criterion that is used for identifying the filters $\mathbf{w}_{ij}(n)$. Often, only second-order statistics are used, i.e. entries of crosscorrelation matrices between output values $y_i(n)$ are minimized, and useful solutions are found because the speech signals to be separated are nonwhite and nonstationary [27,29]. In the most general case, mutual information between the output channels is minimized, which then also requires higher order statistics to additionally exploit the nongaussianity of the signals [29]. Two families of algorithms developed for separation of convolutive speech mixtures can be distinguished: DFT-domain algorithms reduce the convolutive mixtures to scalar mixtures in each DFT bin and then solve the separation problem separately in each bin. While conceptually simple, the alignment of the separation results of several bins to the respective sources requires additional repair mechanisms, typically introducing an additional constraint [27,28]. On the other hand, time-domain algorithms fully acknowledge the convolutional character of the mixture, usually at the cost of greatly increased computational load. However, similar to fast convolution, some time-domain algorithms may be implemented in the DFT-domain with computational advantage and thereby become practicable [29].

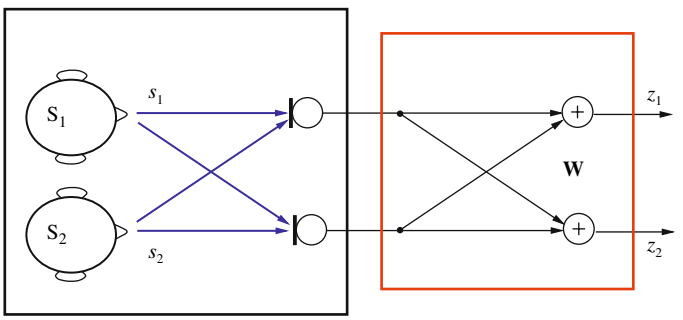


FIGURE 4 Blind source separation for the 2-by-2 case.

References

1. S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, NJ, 4th ed., 2002.
2. S. Boll. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27:113–120, 1979.
3. S.L. Gay and J. Benesty (eds). *Acoustic Signal Processing for Telecommunication*. Kluwer, Boston, MA, 2000.
4. P. Vary. Noise suppression by spectral magnitude estimation – mechanism and theoretical limits. *Signal Processing*, 8(4):387–400, July 1985.
5. J.S. Lim (ed.). *Speech Enhancement*. Prentice Hall, Englewood Cliffs, NJ, 1983.
6. R. Martin. Spectral subtraction based on minimum statistics. *Signal Processing VII: Theories and Applications*, pp. 1182–1185, Edinburgh, Scotland, September 1994. EURASIP.
7. Y. Ephraim and D. Malah. Speech enhancement using a minimum meansquare error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 32: 1109–1121, 1984.
8. Y. Ephraim and D. Malah. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33:443–445, 1985.
9. J.H. Chen and A. Gersho. Adaptive postfiltering for quality enhancement of coded speech. *IEEE Transactions on Speech and Audio Processing*, 3(1):59–71, January 1995.
10. J.S. Lim and A.V. Oppenheim. Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, 67(12): 1286–1604, December 1979.
11. S.F. Boll. Speech enhancement in the 1980s: noise suppression with pattern matching. In M.M. Sondhi and S. Furui (eds). *Advances in Speech Signal Processing*, pp. 309–323. Marcel Dekker, New York, 1991.
12. J.L. Flanagan, D.A. Berkley, G.W. Elko, J.E. West, and M.M. Sondhi. Autodirective microphone systems. *Acustica*, 73:58–71, 1991.
13. B.D. Van Veen and K.M. Buckley. Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–24, April 1988.
14. D.B. Ward, R.A. Kennedy, and R.C. Williamson. Constant directivity beamforming. In M.S. Brandstein and D. Ward (eds). *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 13, pp. 3–18. Springer, Berlin, May 2001.
15. G. Elko. Differential microphone arrays. In Y. Huang and J. Benesty (eds). *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, pp. 2–65. Kluwer, 2004.
16. J. Meyer and G. Elko. Spherical microphone arrays for 3d sound recording. In Y. Huang and J. Benesty (eds). *Audio Signal Processing*

for Next-Generation Multimedia Communication Systems, pp. 67–89. Kluwer, 2004.
17. J. Bitzer and K.U. Simmer. Superdirective microphone arrays. In M.S. Brandstein and D. Ward (eds). *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 2, pp. 19–38. Springer, Berlin, May 2001.
18. S. Doclo and M. Moonen. Gsvd-based optimal filtering for multi-microphone speech enhancement. In M.S. Brandstein and D. Ward (eds). *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 6, pp. 111–132. Springer, Berlin, May 2001.
19. K.U. Simmer, J. Bitzer, and C. Marro. Post-filtering techniques. In M.S. Brandstein and D. Ward (eds). *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 3, pp. 39–60. Springer, Berlin, May 2001.
20. L.J. Griffiths and C.W. Jim. An alternative approach to linear constrained adaptive beamforming. *IEEE Transactions on Antennas and Propagation*, 30(1):27–34, January 1982.
21. O. Hoshuyama and A. Sugiyama. A robust adaptive beamformer with a blocking matrix using constrained adaptive filters. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pp. 925–928. IEEE, 1996.
22. W. Herbordt and W. Kellermann. Adaptive beamforming for audio signal acquisition. In J. Benesty (ed.). *Adaptive Signal Processing: Application to Real-World Problems*. Springer, Berlin, January 2003.
23. J.H. DiBiase, H.F. Silverman, and M.S. Brandstein. Robust localization in reverberant rooms. In M.S. Brandstein and D. Ward (eds). *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 8, pp. 156–180. Springer, Berlin, May 2001.
24. C.H. Knapp and G.C. Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(8):320–327, August 1976.
25. Y. Huang, J. Benesty, and G.W. Elko. Adaptive eigenvalue decomposition algorithm for realtime acoustic source localization system. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, USA. IEEE, March 1999.
26. J.-F. Cardoso. Blind signal separation: statistical principles. *Proceedings of the IEEE*, 86(10):2009–2025, October 1998.
27. L. Parra and C. Spence. Convolutive blind separation of non-stationary sources. *IEEE Transactions on Speech and Audio Processing*, 8(4):320–327, May 2000.
28. S. Makino. Blind source separation of convolutive mixtures of speech. In J. Benesty (ed.). *Adaptive Signal Processing: Application to Real-World Problems*. Springer, Berlin, January 2003.
29. H. Buchner, R. Aichner, and W. Kellermann. Blind source separation for convolutive mixtures: a unified treatment. In Y. Huang and J. Benesty (eds). *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, chapter 10, pp. 255–296. Kluwer, Boston, MA, 2004.

48

Terminals and Their Influence on Communication Quality

Walter Kellermann

Multimedia Communications and Signal Processing, University Erlangen-Nuremberg, Erlangen, Germany, e-mail: wk@lnt.de

With the advent of affordable signal processing hardware, acoustic interfaces of many telecommunication terminals developed from loudspeakers and microphones with some gain control into complex systems integrating powerful signal processing for enhancement of communication quality. The requirements for signal processing to satisfy the human user vary strongly depending on the acoustic interface and the communication scenario. Several perceptually motivated criteria (which are also partially incorporated in recommendations issued by standardizing organizations, such as ITU, ETSI, ISO-MPEG) must be accounted for: Echoes in telecommunication networks should generally be suppressed as much as possible, and the later the echoes arrive, the more crucial is their cancellation/suppression [1]. The echo cancellation requirement can be somewhat relieved for double-talk situations, assuming that voice activity of a communication partner masks the echo of one's own voice.

For conversations among humans, signal delays due to processing and transmission should be minimized in order not to impair perceived communication quality. However, while until

the 1980s only a few milliseconds were tolerated by the established standards for the mostly analog wired networks, the driving forces for digital mobile phones, low-bitrate video conferencing, and voice over IP pushed the limits to larger delays, and nowadays one-way delays on the order of 100 ms are widely accepted.

When eliminating noise from microphone signals, it is not just the amount of noise power which should be minimized, but ideally the signal characteristics of the original noise should be maintained. In any case, artifacts of the noise reduction method must be avoided, as they will call for the listener's special attention and are all the more noticeable, even if they are of little average power. In these cases, even adding some inconspicuous *comfort noise* [2] for masking such artifacts will improve perceptual quality. This technique can also be used for masking residual echoes. Moreover, perception and acceptance of undesired signal components from a remote communication partner depends also on the listening environment. On the other hand, with increasing noise level, a speaker will automatically raise his speech level (*Lombard effect*) and thereby enhance the SNR, which benefits the communication efficiency. If communication is supported by video or still images, the spatial alignment between visual and acoustic perception becomes necessary for user acceptance and, therefore, spatial audio is highly desirable for speech reproduction.

1 • Acoustic Interfaces

Obviously, the closer the microphones and loudspeakers are to the human's mouth and ears, respectively, the less problematic are the corresponding transmission channels and the less is the signal processing needed to assure acceptable communication quality. For instance, for headsets and telephone handsets the quality is essentially determined by the quality of the transducers and their integration in the housing.

For some handsets, especially for handhelds in mobile telephony, the acoustic coupling between loudspeaker and microphone, however, calls already for acoustic echo cancellation, and in case of low-voltage amplifiers and low-cost loudspeakers, the according echo path often exhibits significant nonlinearities [3,4].

In seamless hands-free communication terminals, as depicted for the general multichannel case in Figure 1, every microphone signal generally contains not only the desired signal but also its reverberation, plus additional noise and undesired interfering sources of the local environment, plus acoustic echoes from the loudspeaker(s) broadcasting the remote

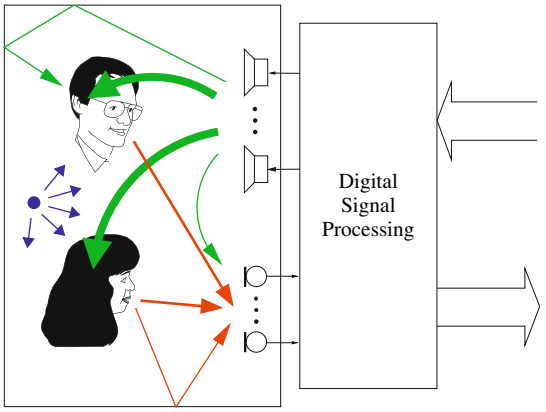


FIGURE 1 General scenario for seamless hands-free speech communication.

communication partners' signals. Along with increasing distance of the desired source(s), the gain for the desired signal(s) must be increased which automatically increases the recorded level of all undesired signal components as well. Similarly, for sound reproduction, the loudspeaker volume will be increased along with the distance between loudspeaker(s) and listener(s) which additionally increases the acoustic feedback into the microphones.

Signal processing systems should fulfill several tasks in the hands-free communication scenario: On the recording side, all desired sources should be captured separately without reverberation, the desired signals should be free from interference and noise, and, often, the source positions should be identified. As discussed above, the techniques of echo cancellation, noise reduction, beamforming, and blind source separation are capable of meeting these goals to a certain extent. So far, dereverberation remains the most difficult task: While beamforming will suppress some reflections from others than the desired direction and the filtering of blind source separation may or may not result in some dereverberation, the actually desired blind equalization of the acoustic transfer functions is still out of reach for realistic time-varying acoustic scenarios.

On the sound reproduction side, the listener, ideally, should hear a clean undistorted signal from the remote partner, while the ambient noise in the listening environment should be suppressed. With multichannel reproduction schemes, typically, spatial information for the remote sources should be represented as well. Current techniques mostly focus on the latter aspect, either by creating a *sweet spot* in the listening environment, where spatial information can be perceived correctly (stereo systems or extensions to five or more channels), or by binaural processing which

can be used for sound reproduction via headsets and crosstalk-canceling loudspeaker systems [5]. Multichannel sound reproduction has also been extended to using dozens to hundreds of channels for synthesizing predescribed wavefields (*ambisonics* and *wavefield synthesis*) [6]. With these concepts, equalization of the listening room acoustics and active cancellation of noise sources and interferers in the listening environment may come into reach.

Considering typical terminals for hands-free telecommunication, the prevalent representatives are desktop telephones incorporating hands-free mode, hands-free telephones in cars, self-contained mobile teleconferencing terminals as used in normal offices especially equipped with teleconferencing rooms and, to a much lesser extent, teleconferencing equipment as used for auditoria [7].

While for the hands-free mode in standard telephone terminals mere echo suppression is still common, echo cancellation and noise reduction are almost always implemented in hands-free equipment for cars where the distance from the mouth to the microphone is often more than half a meter and the ambient noise level is high. Here, microphone arrays with up to four sensors are already in use allowing for multichannel noise reduction and beamforming. This is not only desirable for telecommunications, but also for speech dialogue systems which otherwise suffer severely from the high ambient noise levels in cars. So far, echo cancellation in cars is mostly implemented for monaural sound reproduction, but as most cars are equipped with high-quality stereo or multichannel reproduction systems, their use for telecommunications becomes highly desirable, although this calls for multichannel acoustic echo cancellation (see Sect. 46.1.3).

For mobile teleconferencing equipment, low-cost directional microphones with small apertures are desirable, and therefore differential microphone arrays [8] or just satellite microphones [9] are used which are placed close to the talkers. Satellite microphones for each participant are also often used for dedicated teleconferencing rooms, but here, microphone arrays with an aperture size on the order of the video screen size are already a realistic option to capture several talkers without requiring individual microphones. The advantage of microphone arrays becomes even more evident, when a large number of participants should be covered as, e.g., for people asking questions in a classroom or an auditorium.

If instead of single microphones a – possibly adaptive – microphone array is used to capture the various sources, a combination

of AEC and beamforming becomes necessary, whose interactions have been studied in [10].

Especially in teleconferencing with video modality and multiple talkers on each side, multichannel sound reproduction is necessary to align the acoustic to the visual perception, which requires multichannel AEC either in its general or in some special form [9,11 (Chaps. 5,7)]. As the background noise level is generally very low in such environments and therefore residual echoes are well audible even at very low levels, residual echo suppression is often applied to assure echo attenuation beyond 30 dB which cannot be provided by echo cancellers alone.

References

1. S. Gay and J. Benesty, editors. *Acoustic Signal Processing for Telecommunications*. Kluwer, Norwell, MA, 2000.
2. E. Hänsler and G. Schmidt. *Acoustic Echo and Noise Control: A Practical Approach*. Wiley, Hoboken, NJ, 2004.
3. A. Stenger and W. Kellermann. Adaptation of a memoryless preprocessor for nonlinear acoustic echo cancelling. *Signal Processing*, 80:1747–1760, Sep 2000.
4. F. Küch and W. Kellermann. Nonlinear line echo cancellation using a simplified second-order Volterra filter. *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, USA, May 2002. IEEE.
5. C. Avendano. Virtual spatial sound. In Y. Huang and J. Benesty, editors, *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Chap. 13, pages 345–371. Kluwer, Boston, MA, 2004.
6. S. Spors, H. Teutsch, A. Kuntz, and R. Rabenstein. Sound field synthesis. In Y. Huang and J. Benesty, editors, *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Chap. 11, pages 323–344. Kluwer, Boston, MA, 2004.
7. J.L. Flanagan, D.A. Berkley, G.W. Elko, J.E. West, and M.M. Sondhi. Autodirective microphone systems. *Acustica*, 73:58–71, 1991.
8. P. Chu. Superdirective microphone array for a set-top videoconferencing system. *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pages 235–238, Munich, Germany, April 1997. IEEE.
9. A. Nakagawa, S. Shimauchi, Y. Haneda, S. Aoki, and S. Makino. Channel-number-compressed multichannel acoustic echo canceller for high-presence teleconferencing system with large display. *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pages 813–816, Istanbul, Turkey, June 2000. IEEE.
10. W. Kellermann. Acoustic echo cancellation for beamforming microphone arrays. In M.S. Brandstein and D. Ward, editors, *Microphone*

Arrays: Signal Processing Techniques and Applications, Chap. 13, pages 281–306. Springer, Berlin, May 2001.

11. J. Benesty, T. Gänsler, D.R. Morgan, M.M. Sondhi, and S.L. Gay. *Advances in Network and Acoustic Echo Cancellation*. Springer, Berlin, 2001.

49

Networks and Their Influence on Communication Quality

H.W. Gierlich

HEAD acoustics GmbH, Herzogenrath, Germany

1 • PSTN (Public Switched Telephone Networks)

The dominating influences of traditional TDM-type networks on the communicational quality are

- Frequency-dependent attenuation in the analog parts of the network
- Delay
- Echo performance problems due to insufficient trans-hybrid loss of analog systems

In analog networks no synchronization problems occur; TDM-type networks work isochronously; in all situations the speech channel is associated to a physical channel in the network. The delays in the networks are typically well controlled due to proper transmission planning. Within the network of one operator, conversational speech quality problems due to high delay are unlikely as long as terrestrial transmission systems are used within the network. As soon as satellite transmission is used, e.g., for transcontinental communication, the communication dynamics may be impaired due to long delay which is typically around 240 ms one-way transmission time.

Factors impairing the listening speech quality are found in analog networks: the line characteristics (length and type of lines) have to be taken into account, since they may have impact on loudness ratings, frequency responses, noise, sidetone performance and trans-hybrid losses (see [5,11]).

When a delay greater than 15 ms is introduced in a connection, echo cancellation has to be provided. In the PSTN, echo cancellation is typically concentrated in international exchanges since within national or continental connections the network delay is typically below 15 ms. The echo cancellation is intended to cancel echoes produced by the hybrid of analog terminals. The maximum echo loss to be provided is generally determined by the maximum delay which may be inserted in a connection. However, most modern network echo cancellations conform to ITU-T Recommendation G. 168 [7] which assumes 250 ms one-way transmission time and provides adequate performance limits. If echo suppression is used (sometimes still found in older networks), instead of proper echo cancellation, the speech communication quality may be impaired significantly. Due to switching, the speech signal may be clipped partially, no double talk is possible and the background noise performance is significantly impaired by the network (see description in Section 44.3).

Different speech levels due, for example, to different national level settings or different attenuation in analog systems, can be expected and may degrade the performance of echo cancellers and other signal processing in the network and the terminals.

Speech coding is not typically found in PSTN networks, with the exception of DCME systems (digital circuit multiplication equipment, see [9]). If DCME is inserted in a PSTN network, all considerations concerning speech quality found in Section 44 apply.

2 • Packet-Based Transmission

Packet-based networks differ considerably from circuit-switched networks. In packet-based networks the communication quality may be influenced by

- Speech coding and packetization
- Delay which may be high and unpredictable
- Delay variation during a call
- Packet loss and packet loss concealment
- Echo performance problems due to insufficient echo cancellation
- VAD (voice activity detection) and comfort noise insertion

The (coded) input signal (speech) is segmented into packets typically of fixed length. Typical packet lengths are in the range 5–30 ms. Packetization itself introduces a minimum delay of the packet length introduced into the transmission. Packets from different sources (e.g., terminals or subscribers) are sequenced for transmission. This process, as well as any buffering or routing in a connection, may add additional delay. While in isochronous networks the delay of a connection is constant and predictable, the delay in IP networks may vary between different connections depending on network load, routing, bandwidth, switching and other effects. In addition, jitter may be introduced by the IP connection. Jitter occurs when the delays for different packets are different. Delay and jitter are not predictable in IP networks; therefore, they are described by statistical parameters, such as average delay (including delay distribution) and packet loss distribution.

At the IP network termination points (in gateways or the IP terminals) the receiver has to take care of the collection of the packets and their correct ordering. Jitter buffers are used in order to collect and sequence the packets received in such a way that an almost error-free transmission is achieved.

Due to the time-variant nature of IP networks (see Chapter 43.2.2) time-variant delay and jitter as well as packet loss have to be taken into account, and speech quality measurements have to be conducted for the various network conditions to be expected. It is advisable to make appropriate simulations in a lab environment in order to get an estimate of the speech quality range to be expected under real use conditions (see ITU-T Y. 1541 [10]).

In most packet-based systems, echo cancellation is provided, either in the terminal or in the network. In IP connections gateways have to be equipped with echo cancellers providing sufficient echo loss for all expected delays in a connection. The maximum echo loss to be provided is determined by the maximum delay which may be inserted in a connection. Since under worst-case conditions the maximum one-way transmission delay may be higher than 300 ms, the minimum echo loss which has to be provided is 55 dB [6]. Consequently, any IP terminal connected to an IP network has to provide this amount of echo loss. Different speech levels, due to computer terminals with insufficiently controlled loudness ratings or equipment with improperly implemented codecs for example, may degrade the performance of echo cancellers and other signal processing in an IP network. Speech coding especially in combination with packet loss and jitter may add further impairments to the

transmitted speech signal. Speech coders with integrated packet loss concealment algorithms (PLC), e.g., [8], may improve the listening speech quality.

Frequently, silent periods, when the signal level is below a defined threshold, are recognized using voice activity detection (VAD) and are not transmitted in order to save bandwidth or power. At the far end, silent periods are typically replaced by low-level noise (comfort noise). In background noise situations, this procedure may lead to background noise modulations which may be annoying for the user. Furthermore, voice activity detection may cause clipping of the actual speech signal.

Further information about VoIP networks and configurations and QoS requirements can be found in [4] and [11].

3 • Mobile Networks

In mobile networks, the main parameters contributing to the communication speech quality are

- The radio channel and its transmission impairments
- Speech coding
- Delay introduced by the speech coding, the channel coding and the transmission in the mobile network
- Bit errors, frame loss and frame loss concealment
- Handover between different cells
- Echo performance of the echo cancellers used for interconnection to the PSTN and echo cancellation of the mobile terminal's echo
- Performance of additional signal processing introduced in the network and intended to automatically adjust signal levels or to reduce background noise

The radio network and its cellular structure is very sensitive to transmission errors. Disturbances introduced by the radio channel may result in bit errors or frame loss. As shown in Chapter 43.2.3, the different elements of the mobile network show a different sensitivity to transmission errors.

3.1 Listening Speech Quality

In the mobile network, the listening speech quality is limited by the codecs used [1,2]. However, even if no degradation by the mobile network is observed, the speech quality provided by the speech coder may be decreased due to environmental influences, especially background noise. Noise cancellation in mobile terminals may help to reduce the background noise, but will degrade the speech sound quality. Furthermore, the listening

speech quality may be degraded by a number of different impairments introduced by the network. In the radio network, co-channel interference (C/I), adjacent channel interference (C/A), noise limitations (C/N) and other effects may lead to increasing bit error rates and to frame errors. Increasing BER and FER lead to less information about the coded speech and thus the listening speech quality will degrade. The various techniques used in the radio network to increase the quality of the radio environment also have impact on the listening speech quality. The advantage of DTX (discontinuous transmission) is the reduction of the interference in the radio channel since the signal is only transmitted during speech activity. This may lead to less BER and FER and thus to an increase of the listening quality of the speech transmitted. On the other hand, however, the speech quality may be impaired by speech clipping (due to non-ideal speech activity detection) and comfort noise insertion used at the far end in order to substitute the background noise (which is not transmitted). Frequency hopping may increase the listening speech quality since the effect of multipath fading and interference can be reduced. Dynamic power control may result in a further increase in the protection against fading dips and interference and therefore may contribute to less BER and, as a result of lower BER, to higher listening speech quality. The quality is always impaired during handover. When a user "leaves" one cell and enters another one (handover), the speech is interrupted briefly, and speech frames are used to transmit signaling information.

3.2 Conversational Speech Quality

Due to the delay introduced by mobile networks for source coding, channel coding and radio transmission, the conversation dynamics may be impaired, especially in connections where the mobile network is interconnected to other networks introducing additional transmission delay. Furthermore, due to the delay introduced by the mobile network, echo cancellation must be provided in the mobile switching centers to cancel echoes produced, for example, by analog PSTN connections. This echo cancellation is required to protect the mobile terminal user from echoes generated by the PSTN. To avoid audible acoustical echoes the echo loss provided by a mobile terminal has to be at least 46 dB [3]. Some mobile terminals, however, do not provide the required echo attenuation in all operational modes; therefore, in addition to the echo cancellers facing the PSTN, echo cancellers facing the mobile users' terminals are integrated into some mobile networks. These are intended to cancel acoustical echoes remaining from mobile terminals. It should be recognized that these echo cancellers have to operate on a non-linear and time-variant echo path. The complete

radio channel including the speech codec is part of the echo path. Due to the mostly uncorrelated input and output signal (see Figure 1 in Chapter 50), the adaptive FIR filter is not suitable to model the echo path. As a consequence, these speech echo cancellers typically rely on the non-linear processor and operate as a type of echo suppressor. This operating mode leads to speech clipping during double talk and thus to an insufficient performance. Furthermore, strongly varying speech levels in the network may impair the efficiency of the echo cancellers.

In addition, a variety of non-linear and/or time-variant signal processing, such as AGC (automatic gain control) or compression techniques, may be used in mobile terminals in order to enhance the speech quality in the different acoustical conditions, but this processing may interact with the network signal processing (see Chapter 50).

References

1. 3GPP: TS 46.010: Full Rate Speech Encoding, Third Generation Partnership Project (3GPP), 2002
2. 3GPP: TS 46.051: GSM Enhanced Full Rate Speech Processing Functions: General Description, Third Generation Partnership Project (3GPP), 2002
3. ETSI EG 201 377-2: Specification and Measurement of Speech Transmission Quality; Part 2: Mouth-to-Ear Speech Transmission Quality including Terminals, 2004
4. ETSI TS 101 329-5: End-to-End Quality of Service in TIPHON Systems, Part 5: Quality of Service (QoS) Measurement Methodologies, 2002
5. ITU-T Recommendation G.107: The E-model, A Computational Model for Use in Transmission Planning, International Telecommunication Union, Geneva, 2003
6. ITU-T Recommendation G. 131: Control of Talker Echo, International Telecommunication Union, Geneva, 2000
7. ITU-T Recommendation G. 168: Digital Network Echo Cancellers, International Telecommunication Union, Geneva, 2004
8. ITU-T Recommendation G.711: Appendix I: A High Quality Low-Complexity Algorithm for Packet Loss Concealment with G.711, International Telecommunication Union, Geneva, 1999
9. ITU-T Recommendation G. 763: Digital Circuit Multiplication Equipment Using G.726 ADPCM and Digital Speech Interpolation, International Telecommunication Union, Geneva, 1998
10. ITU-T Recommnedation Y.1541: Network Performance Objectives for IP-Based Services, International Telecommunication Union, Geneva, 2001
11. Richards, D.L.: Telecommunication by Speech, Wiley, New York, 1973, ISBN 0 470 71949 4

50

Interaction of Terminals, Networks and Network Configurations

H.W. Gierlich

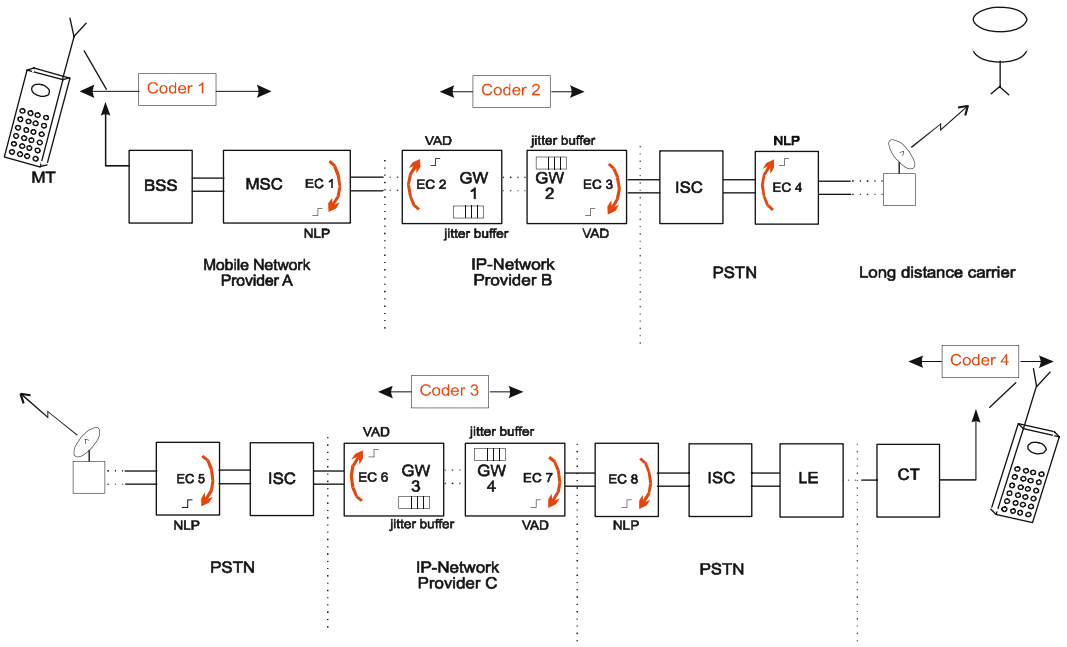
HEAD acoustics GmbH, Herzogenrath, Germany

The interaction of terminals and different network configurations may be many-fold and is discussed in this chapter only with reference to some examples. Typically, it is the task of standardization to avoid, or at least to minimize, interaction problems which might occur in different network configurations. Due to the liberalized telecommunications market, however, the variety of configurations found in today's telecommunication is increasing rapidly. Non-linear and time-variant signal processing is used in terminals as well as in networks and network components. Also different types of networks (analog, TDM, IP, mobile networks) are interconnected globally and different network providers may be involved in one connection.

Figure 1 gives an overview of a possible configuration which is used to explain different types of interaction problems which may occur.

1 • Tandeming of Speech Coders

Different types of speech coders may be used in different parts of the connection. Depending on the speech coding, the degradation of listening speech quality may vary significantly. In



BSS	Base Station System
MSC	Mobile Switching Center
ISC	International Switching Center
PSTN	Public Switched Telephone Network
LE	Local Exchange
GW	IP-Gateway
CT	Cordless Telephone
EC	Echo Canceller
NLP	Non-Linear Processor
VAD	Voice Activity Detection

FIGURE 1 Example configuration for the interconnection of terminals and networks.

general, tandeming of all types of CELP coders using a parametric description of the speech signal (Chapter 45.2) will lead to stronger degradation as compared to the tandeming of waveform coders. Reference [1] gives an overview of the expected quality degradation for different types of tandemed coders. The influence of tandeming on the listening speech quality has to be derived subjectively. For transmission planning, the impairments are provided in terms of equipment impairment factors (*I*e) based on the E-model [2]. The E-model assumes that impairments are additive on a psychological scale, but this is not really valid when tandeming different types of speech coders. Since this interaction problem is known, more effort is put into the standardization of

tandem-free operation. However, due to the different topologies of networks interconnected and due to missing or improperly implemented signaling information about the type of coders used in a connection, such tandem-free operation can be realized typically only in one network. The realization of universally accepted signaling information about the type of coders seems to be not achievable.

2 • Delay

Figure 1 illustrates that when interconnecting multiple networks and different types of terminals, the transmission delay is increased with the insertion of each additional network and the type of signal processing. Although the individual components of a connection might well be within a specified range, the overall transmission delay might be unacceptably high. This is to be expected especially in cases where mobile and IP networks are interconnected. Therefore, in terminals as well as in networks, the general rule to keep the delay as low as technically possible should be followed during the design process. Otherwise, the communication quality might be impaired significantly so that successful interaction between the conversation partners is no longer possible. Furthermore, the increase of the delay increases the requirements on echo loss significantly [3].

3 • Tandeming of Speech Echo Cancellers

Echo cancellation may be found in many places in the network but the general rule should be that echo cancellation should be performed as closely as possible to the echo source itself. A four-wire network itself typically does not produce any echo, so the echo cancellation should be realized in the terminals which produce acoustical echo. During recent years, this has been more clearly recognized, and the requirements have been specified by standards [4]; however, it is still not guaranteed that a sufficient amount of echo loss is provided in all conversational situations. Today's telecommunications requirements often do not address the problem of time-variant behavior of echo cancellers in different conversational situations in the appropriate standards. This situation leads to the insertion of additional echo cancellers in networks, even where the echo cancellation should have been completely provided by the terminal itself (Figure 1, ECs 2, 4, 6).

One more reason for deployment of speech echo cancellers in a network is the analog network which, for the end point connections to the customer, uses two-wire connections. These end

points are terminated by a hybrid [5] and the trans-hybrid loss is always limited, causing echo. Echo cancellation therefore has to be provided for these connections where the transmission delay exceeds about 15 ms. In the past, this was typically required only for lengthy international connections, and the echo cancellers were deployed only in the international exchanges [6,7]. With proper network planning, it was possible to avoid the use of echo cancellers in the shorter national networks, at least for smaller countries. In today's network configurations, these echo cancellers are still used, but the PSTN may be interconnected to other types of networks rather than directly to the conversation partner (Figure 1). When assuming that an IP network exists in the connection between the conversation partners, both gateways in the IP network are equipped with echo cancellers to reduce possible echoes from their end points. As seen in Figure 1, many echo cancellers (ECs 1, 3, 5, 7, 8 and ECs 2, 4, 6) are facing the same direction. Under this condition, only a proper design of the network speech echo canceller will avoid echo problems: The echo canceller closest to the echo source needs to sufficiently reduce the echo produced by the terminal under all conditions. The echo cancellers in the network facing the same direction should be mostly transparent. The ideal solution would be proper signaling between the networks for disabling the echo cancellers which, however, cannot be guaranteed. Therefore, a proper control of the "internal" echo cancellers is the only possibility to avoid effects like echo generation by the digital filter in the echo canceller (due to a "non-existent echo path"), temporal clipping (due to improperly activated non-linear processing) and others.

4 • Tandeming VAD and NLP

Yet another tandeming problem can be seen in the example shown in Figure 1. Voice activity detection in the gateways in combination with non-linear processing in the echo cancellers may lead to a variety of speech clipping effects. Different thresholds may lead to interrupted speech when low-level speech signals are transmitted. Different types of comfort noise insertion used in combination with different VAD- and NLP-implementations may lead to interrupted speech signals modulated with various types of comfort noise and – in the presence of background noise – also to background noise modulations. Such effects may be very annoying and should be avoided.

5 • Other Signal Processing

Recent standardizations [8] also describe the so-called voice enhancement devices in the network. The signal processing is intended to equalize speech levels found in a network, to reduce background noise which might have been picked up by the terminal and transmitted to the network, or even to enhance the "clarity" of the voice. The use of such equipment in a network may cause problems in cases where similar signal processing is already implemented in the terminal. For example, tandemed noise reduction will lead to speech sounding robotic and often to background noise modulation.

All the effects described are not properly modeled by the network planning tools used today [2]. Therefore, only careful transmission planning by experts which are aware of the problems described, in combination with careful system design, can avoid or minimize interaction problems.

References

1. ITU-T Recommendation G.113: Transmission Impairments due to Speech Processing, International Telecommunication Union, Geneva, 2001
2. ITU-T Recommendation G.107: The E-model, A Computational Model for Use in Transmission Planning, International Telecommunication Union, Geneva, 2003
3. ITU-T Recommendation G. 131: Control of Talker Echo, International Telecommunication Union, Geneva, 2000
4. ANSI-TIA/EIA 810-A: Transmission Requirements for Narrowband Voice over IP and Voice over PCM Digital Wireline Telephones, Dec. 19, 2000 ETSI EG 201 377-1: Specification and Measurement of Speech Transmission Quality; Part 1: Introduction to Objective Comparison Measurement Methods for One-Way Speech Quality Across Networks, 2003
5. Richards, D.L.: Telecommunication by Speech, Wiley, New York, 1973, ISBN 0 470 71949 4
6. Schütte, U.: Übertragungstechnische Planung von Netzen für Sprachdienste, Schiele & Schön, 1992, ISBN 379 4905 326
7. ITU-T Recommendation G.101: The Transmission Plan, International Telecommunication Union, Geneva, 2003
8. ITU-T Recommendation G. 169: Automatic Level Control Devices, International Telecommunication Union, Geneva, 1999

PART VIII
NOISE

Ichiro Yamada

Aviation Environment Research Center, Tokyo, Japan

51 Sound Levels of Noise Sources

Hideki Tachibana 933

Sound Power Level • Sound Energy Level

52 Outdoor Measurements

Hideki Tachibana, Kohei Yamamoto, Yasushi Miki, Yasushi Takano, and Ichiro Yamada 949

Measurement of Acoustical Properties of Ground Surfaces – *Hideki Tachibana* • Measurement of Shielding Effect of Noise Barriers – *Kohei Yamamoto and Yasushi Miki* • Measurement of Sound Attenuation for Long-Distance Propagation Outdoors – *Yasushi Miki* • Noise Sources in Road and Railway Transport – *Yasushi Takano* • Detection of Aircraft Sound Events – *Ichiro Yamada*

53 Measurements of Physical Characteristics Using Sound at Audible or Lower Frequencies

Ippei Torigoe 983

Acoustic Micrometer • Acoustic Volume Meter • Acoustic Surface Area Meter • Acoustic Rate Gyro • Acoustic Impedance Flowmeter • Chopper Microphone • Kinematic Viscosity Sensor

54 Community Response to Noise

Sandy Fidell 1005

Statistical Character of Environmental Noise in Populated Areas • Predicting the Annoyance of Environmental Noise • Predicting Sleep Disturbance • Predicting the Likelihood of Noise-Induced Complaints • Summary

51

Sound Levels of Noise Sources

Hideki Tachibana

Institute of Industrial Science, University of Tokyo, Japan

1 • Sound Power Level

For assessment of the acoustic emission from sound sources such as machinery and equipment, sound pressure level is often measured at a specified point at a certain distance from a source. The result of this kind of measurement indicates the acoustical state (sound pressure level) at the measurement point at a particular time and place and can be used for simple comparison of noise emission by machinery and equipment of the same kind and operating under the same conditions in a similar acoustical environment.

However, when we want to know the total acoustic emission from sound sources, it is necessary to measure the sound power level if the sound source is steady or continuous. When the source emits impulsive or intermittent sound, the sound energy level has to be de determined, see Figure 1.

When describing the acoustic emission of a source of steady or continuous sound, sound power (in watts) is used. The result is usually expressed as sound power level, L_W, (in decibels) according to

$$L_W = 10 \lg \left(\frac{P}{P_0} \right) \text{dB}, \tag{1}$$

where P is the sound power of the source and $P_0 = 1\,\text{pW}$ (reference value).

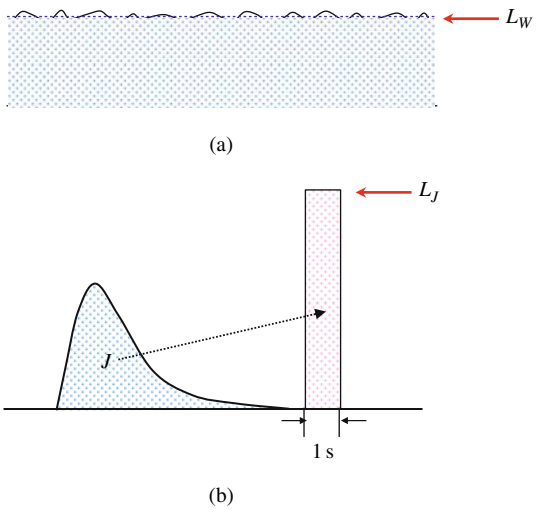


FIGURE 1 The ways of expressing acoustic emission from (**a**) steady and (**b**) impulsive/ intermittent sound sources.

There are several methods to determine sound power level. The methods can be divided into (1) the sound pressure method and (2) the sound intensity method, according to the physical quantity to be measured. The sound pressure method is further divided into I (free-field and hemi-free-field methods) and II (diffuse field method). These measurement methods are specified in ISO standards (ISO 3740 series [1–8] and 9614 series [9–11]) according to the measurement principle, the nature of the sound field and the uncertainty in the result. In this section, the methods for the determination of sound power level are described with emphasis on the measurement principle.

1.1 Sound Pressure Method I (Free-Field Method, Hemi-free-field Method)

When considering a sound source located in a free field and surrounded by a hypothetical sphere as shown in Figure 2a, the sound power, P, radiated from the sound source equals the total sound power passing outwards through the surface of the sphere, expressed as the product of the time-averaged sound intensity, $\bar{I}$, averaged over the surface of the sphere and the surface area, S, of the sphere as follows:

$$P = \bar{I} \cdot S. \tag{2}$$

In the sound field far from the sound source and over some range of frequency, the sound intensity can be approximated from the sound pressure, and Eq. (2) can be expressed as

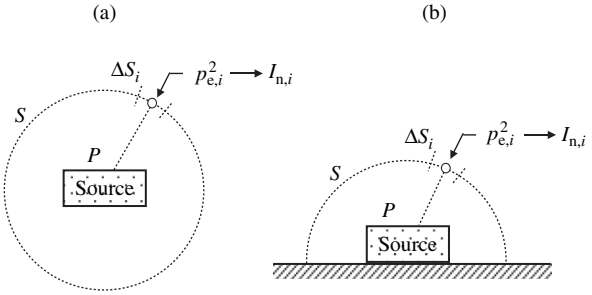


FIGURE 2 Sound pressure method I. (**a**) Free-field method; (**b**) hemi-free-field method.

$$P = \left(\frac{\overline{p_e^2}}{\rho c}\right) \cdot S, \tag{3}$$

where $\overline{p_e^2}$ is the square of the effective or root-mean-square sound pressure and the product of air density, ρ, and the speed of sound, c, is the characteristic impedance of air [N/(m^3/s)] at the prevailing temperature and static pressure of the air.

Consequently, the sound power level can be expressed as

$$L_W = \overline{L_p} + 10 \lg \left(\frac{S}{S_0}\right) \text{dB}, \tag{4}$$

where $\overline{L_p}$ is the level of the time-averaged sound pressure averaged over the surface of the sphere and $S_0 = 1\,\text{m}^2$ (reference value).

In this method, the average sound intensity over the measurement surface is approximated from the square of the time-averaged sound pressure. Thus, this method is a kind of sound pressure method (sound pressure method I), for a steady (that is, time invariant) sound source in a region free of reflecting surface; the method is called the free-field method.

When sound power is determined using a hemisphere over a reflective plane (usually, a floor) as shown in Figure 2b, the sound power of the source can be determined in the same way as described above; this method is called the hemi-free-field method. In this case, area S in Eqs. (3) and (4) is the surface area of the hemisphere.

In order to determine sound power with minimal uncertainty by the free-field or hemi-free-field method, an anechoic room or a hemi-anechoic room is needed. The room should have adequate acoustical characteristics over the range of frequency of interest as specified in ISO 3745 for precision-grade sound power measurements. In that standard, a correction is added to the measured results in order to obtain the sound power level under the standard meteorological conditions (temperature: 23 °C, atmospheric pressure: 1,013 hPa), but the numerical expression has not yet been decided. The absolute value of this correction is very small under normal meteorological condition.

When precision-grade uncertainty is not required, the sound power measurement can be made where the acoustical characteristics are not well controlled, such as outdoors or in a room with a large volume and a short reverberation time. For such measurements, ISO 3744 (engineering grade) or ISO 3746 (survey grade) is appropriate. According to those standards, a rectangular measurement surface may be used instead of a spherical or hemispherical one. In such sound fields, measurement uncertainty can be introduced because of the influence of reflections and reverberation. These sources of uncertainty have to be minimized. The environmental correction term in the applicable standards expresses the allowable extent of the contribution of reflections and reverberation.

1.2 Sound Pressure Method II (Diffuse Sound Field Method)

As another sound pressure method for determining sound power levels, the sound energy density in a diffuse sound field can be estimated from the space and time-averaged sound pressure. When a sound source is located in a diffuse sound field, for example, in a reverberation room as shown in Figure 3, the sound power P emitted from the source and that absorbed at the room boundaries (wall, ceiling and floor) may be expressed by

$$P = \frac{c \cdot \overline{E}}{4} A, \tag{5}$$

where A is the equivalent sound absorption area of the room (m^2), c is the speed of sound (m/s) and $\overline{E}$ is the time-averaged sound energy density in the room (J/m^3).

In a diffuse sound field, the space- and time-averaged sound energy density $\overline{E}$ and the square of the effective sound pressure are related as $\overline{E} = p_e^2/\rho c^2$ and Eq. (5) can be expressed as

$$P = \frac{A \cdot p_e^2}{4\rho c}. \tag{6}$$

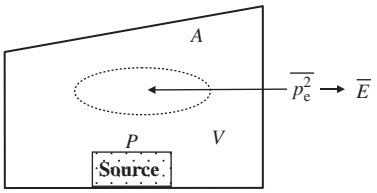


FIGURE 3 Sound pressure method II (diffuse sound field method).

Sound power level is found from

$$L_W = L_p + 10\lg\left(\frac{A}{A_0}\right) - 6\text{dB}, \tag{7}$$

where L_p is the level of the space- and time-averaged sound pressure in the room (dB) and $A_0 = 1\,\text{m}^2$.

The equivalent sound absorption area A is obtained from application of the Sabine equation and a measurement of reverberation time (s) in the room using

$$A = \frac{55.26 \cdot V}{T_{rev}}, \tag{8}$$

where V is the internal volume of the room (m^3) and the dimensional constant of 55.26 has the unit (s/m).

By method II, the result of a measurement is influenced by the sound-absorbing characteristics of the room; hence, the sound power level is best determined in bands of frequency. The measurement standard based on this principle (ISO 3741) specifies that the sound power levels shall be measured in one-third octave bands and that octave band or A-weighted sound power level shall be calculated from the one-third-octave-band sound power levels.

In ISO 3741, the following equation is adopted instead of Eq. (7):

$$L_W = \overline{L_p} + 10\lg\left(\frac{A}{A_0}\right) + 4.34\left(\frac{A}{S}\right) + 10\lg\left[1 + \frac{S \cdot c}{8V \cdot f}\right] + C - 6\text{dB}, \tag{9}$$

where S is the total surface area of the reverberation room (m^2). On the right side of Eq. (9), the third term is the correction related to the sound absorption in the reverberation room [12], the fourth term is for the correction of the sound energy that concentrates in the vicinity of the room boundaries (the Waterhouse's correction [13]). The fifth term C is for correction for meteorological condition, and its numerical expression has not yet been decided as mentioned above.

When measuring the sound power level of a sound source containing discrete spectral components by the diffuse sound field method, sufficient attention has to be paid to acoustic modes in the room. ISO 3741 gives detailed requirements for the specification of the reverberation room, use of a rotating reflector (diffuser), the number of the source positions and measurement point and the judgement of measurement uncertainty by checking the spatial deviation of sound pressure level in the room.

ISO 3743-1, ISO 3743-2 and ISO 3747 provide standard procedures for engineering-grade measurement methods based on the diffuse sound field method.

1.3 Comparison Method (Sound Source Substitution Method)

The sound power level of a sound source, L_W, can be obtained from the measurements of the sound pressure levels produced by a source under test and by a reference sound source in a reverberant sound field, respectively, by means of the following equation:

$$L_W = L_{Wr} + (\overline{L_p} - \overline{L_{pr}})\text{dB}, \tag{10}$$

where L_{Wr} is the sound power level of the reference sound source (dB), $\overline{L_p}$ is the level of the space-averaged and time-averaged sound pressure produced by the source under test (dB) and $\overline{L_{pr}}$ is the level of the space-averaged and time-averaged sound pressure produced by the reference sound source (dB).

For the method based on this measurement principle, ISO 3741 specifies a precision-grade procedure and ISO 3743-1 and ISO 3747 specify engineering-grade procedures. To determine sound power by this method, the sound power level of the reference sound source has to be accurately calibrated [12,14,15]. ISO 6926 [16] specifies the calibration procedure and the way of positioning the reference sound source in the test room. An impulsive sound source can be used as another type of reference sound source [17]; in this case, the sound energy radiated by the source has to be determined by calibration of the source of impulsive sound.

Although the sound power level of a reference sound source should be stable and constant, it is inevitable that the sound power produced by any reference sound source is influenced by the variations in radiation impedance in the sound field. Figure 4 shows the measured results of the sound power level of a loudspeaker-type sound source (B& K 4205, Figure 4a) [18] measured by the sound intensity method as described in the next clause. In this experiment, the sound source was located on a reflective floor, and the distance between the source and

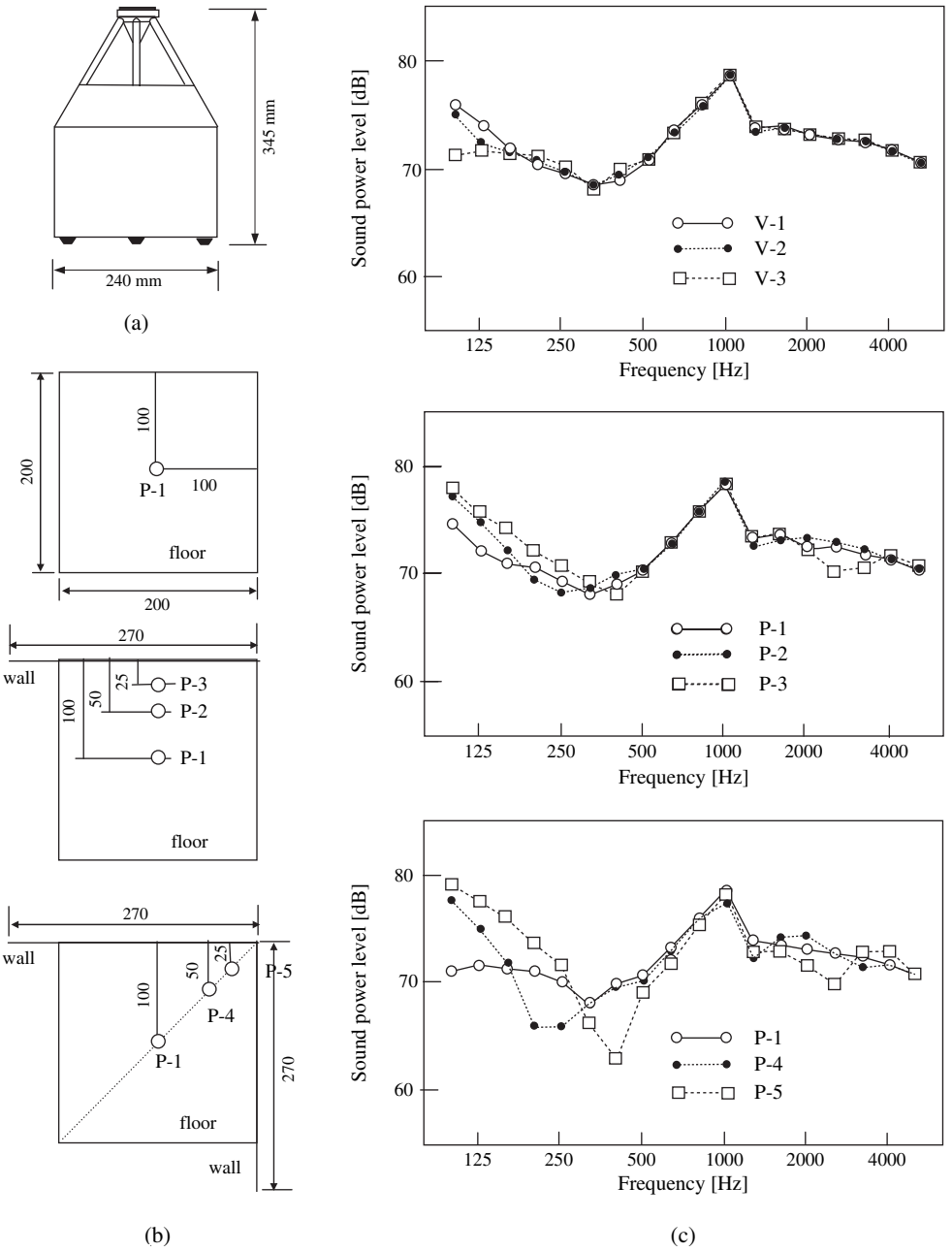


FIGURE 4 Change of sound power level of a loudspeaker-type sound source due to the differences in the boundary condition [18]. (**a**) Sound source (B&K 4205); (**b**) measurement conditions; (**c**) measurement results.

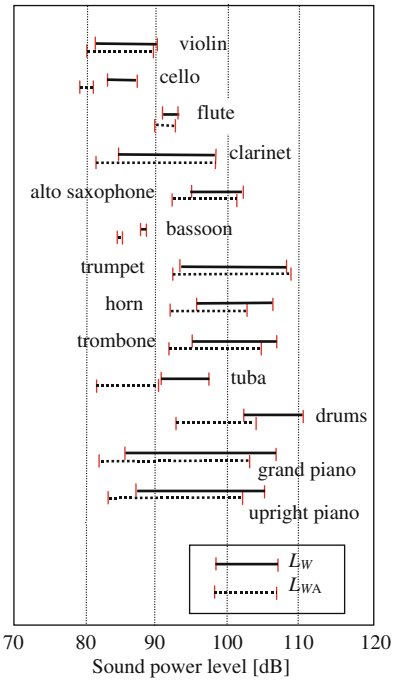


FIGURE 5 Sound power levels of musical instruments measured by the comparison method.

the reflective wall(s) was varied as indicated in Figure 4b. The results in Figure 4c clearly show that the sound power level of the source varied significantly especially at low frequencies as a result of differences in the influence of reflections on the sound power output.

As an example of application of the comparison method, Figure 5 shows the measured sound power levels produced by various kinds of musical instruments. The measurements were performed in an audition room with 0.7 s reverberation time in the mid-frequency range. As a reference sound source, an omni-directional dodecahedron loudspeaker was used for which its sound power level had been calibrated by the diffuse sound field method.

1.4 Sound Intensity Method

As indicated in Figure 6, the sound intensity over a hypothetical measurement surface can be directly measured by the sound intensity method. Consequently, the sound power produced by the sound source can be obtained by integrating the sound intensity over the area of the measurement surface. That is according to

$$L_{\mathrm{W}} = 10 \lg \left(\frac{S \cdot \overline{I_{\mathrm{n}}}}{P_0} \right) \mathrm{dB}, \tag{11}$$

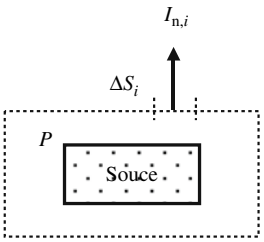


FIGURE 6 Sound intensity method.

where $\overline{I_n}$ is the surface-averaged and time-averaged sound intensity over the measurement surface (W/m^2), S is the area of the measurement surface (m^2) and $P_0 = 1\,pW$ (reference value).

In practice, the measurement surface is divided into a finite number of segments (sub-areas), and the level of the sound intensity normal to an enveloping surface is measured on each segment (see Figure 7). From the measurement results,

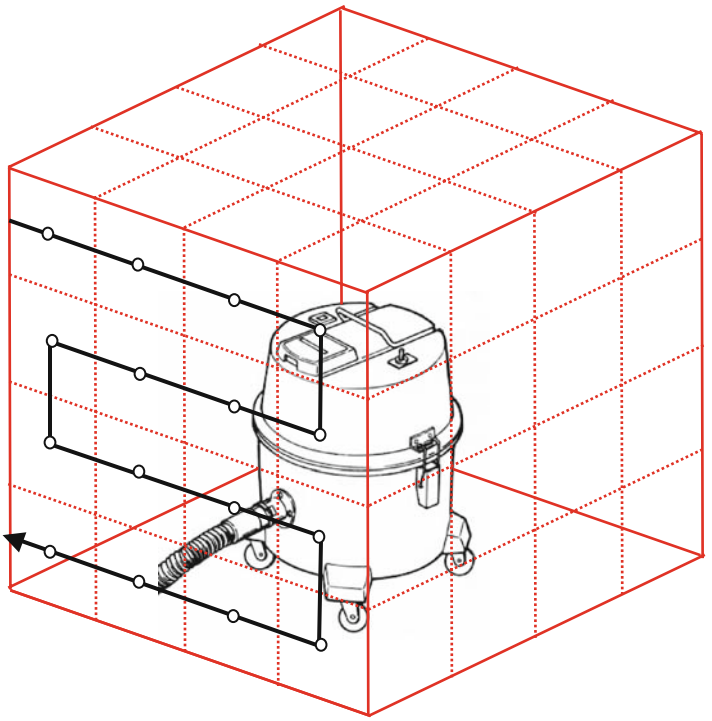


FIGURE 7 Setting a measurement surface surrounding a sound source under test.

the sound power level of the source under test is calculated according to

$$L_{\mathrm{W}} = 10 \lg \left[\frac{\sum_i \left(10^{L_{\mathrm{I_n},i}/10} \cdot S_i \right)}{S_0} \right] \mathrm{dB}, \tag{12}$$

where S_i is the area of the ith segment (m^2), $L_{\mathrm{I_n},i}$ is the mean of the level of the normal sound intensity over the ith segment (dB), and $S_0 = 1\,\mathrm{m}^2$ (reference value).

For sound power determined from measurements of sound intensity, the sound power emitted from the sound source is obtained by vector-based integration of the sound intensity over the measurement surface, and therefore the extraneous noises can be cancelled in principle.

In the case of measurement of a complex sound source as shown in Figure 8, the sound power emitted from the separate sound sources can be determined by choosing a closed measurement surface enveloping each sound source and integrating the sound intensity over each measurement surface [19]. However, when there is significant sound absorption inside a measurement surface, the influence of extraneous noises cannot be cancelled, and the uncertainty of the resulting determination of sound power level may be unacceptably large.

There are two ways to measure the normal sound intensity on a measurement surface: (1) measure the sound intensity at discrete measurement points uniformly distributed over the measurement surface (discrete-point method) and (2) measure the spatial average of the instantaneous sound intensity over the measurement surface by scanning the surface with an intensity probe that is moved continuously at a constant speed (scanning method). The first method is specified in ISO 9614-1 (precision, engineering and survey grades); the latter is specified in ISO

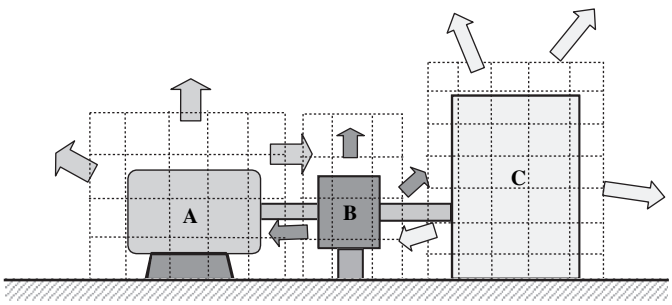


FIGURE 8 An example of a complex noise source.

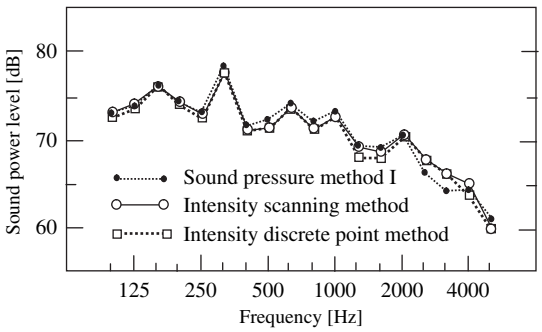


FIGURE 9 Sound power level of a vacuum cleaner measured by the three methods.

9614-2 (engineering grade) and ISO 9614-3 (precision grade). In these standards, various kinds of sound field indicators are specified and have to be measured and checked to estimate the measurement uncertainty [20]. When a precision-grade measurement is performed by the discrete-point method, the number of measurement points can be enormous. The scanning method can be conveniently applied as an alternative. Even with manual scanning, sound intensity can be measured with minimal uncertainty if the operator is well experienced.

As an example of actual measurement, Figure 9 shows the sound power level of a shop vacuum cleaner measured in a hemi-anechoic room (see Figure 7). For this measurement, the sound pressure method (hemi-free-field method), the discrete-point sound intensity method (125 points distributed over a 20-cm × 20-cm area) and the scanning sound intensity method were performed. The measured sound power levels were in good agreement.

2 • Sound Energy Level

Sound power is a quantity based on the space and time average of acoustic emission from sound sources that generate sound continuously. The concept of sound power cannot be applied to impulsive or intermittent sounds. For assessment of such transient sounds, sound energy has to be measured and expressed as a sound energy level [21,22]. Sound energy level is expressed mathematically as

$$L_J = 10 \lg \left(\frac{J}{J_0} \right) \text{dB}, \tag{13}$$

where J is the sound energy emitted from the sound source (J) and $J_0 = 1\,\text{pJ}$ (reference value).

For determination of sound energy levels of impulsive or intermittent sound sources, the sound pressure and sound intensity methods, used for determination of sound power levels, may be applied, by replacing time averaging by time integration of squared sound pressure or sound intensity.

2.1 Sound Pressure Method I (Free-Field Method, Hemi-free-field Method)

Instantaneous sound intensity at a point on a hypothetical sphere surrounding the sound source under test is expressed as

$$I_{\text{n}}(t) = p(t) \cdot u_{\text{n}}(t), \tag{14}$$

where $p(t)$ is the instantaneous sound pressure and $u_{\text{n}}(t)$ is the normal component of the instantaneous particle velocity vector.

In a sound field far from the sound source at a distance that depends on frequency, the instantaneous sound intensity can be approximated from the instantaneous sound pressure and expressed as

$$I(t) = \frac{p^2(t)}{\rho c}. \tag{15}$$

Time-integrated sound intensity (termed sound intensity exposure (J/m^2)) is expressed as

$$j = \int_0^\infty I(t)\mathrm{d}t = \left(1/\rho c\right) \int_0^\infty p^2(t)\mathrm{d}t, \tag{16}$$

and the corresponding sound intensity exposure level (dB) is expressed as

$$L_j = 10 \lg \left(\frac{j}{j_0} \right) \text{dB}. \tag{17}$$

In this case, the reference value for sound intensity exposure j_0 is determined according to the following expression:

$$j_0 = \frac{J_0}{S_0} = I_0 \cdot T_0 = \frac{p_0^2}{\rho c} T_0 = 10^{-12}\,\text{J / m}^2,$$

where $J_0 = 1\,\text{pJ}$, $S_0 = 1\,\text{m}^2$, $I_0 = 1\,\text{pW/m}^2$, $T_0 = 1\,\text{s}$, and $p_0 = 20\,\mu\text{Pa}$.

Consequently, we have

$$L_j = 10 \lg \left[\frac{\int_0^\infty p^2(t)\mathrm{d}t}{p_0^2 T_0} \right] = L_{\text{pE}}. \tag{18}$$

For an impulsive sound, sound intensity exposure level, L_j, is equal to sound exposure level, L_{pE}, as specified in the International Standard for sound level metres [23] for measurements of the level of the sound exposure of an impulsive sound.

The sound energy J emitted from the sound source can be obtained by integrating the sound intensity exposure j over the measurement surface as

$$J = \int_S j \, \mathrm{d}s. \tag{19}$$

Therefore, the sound energy level of the sound source L_J can be obtained as follows:

$$L_J = 10\lg\left(\frac{\bar{j}\cdot S}{j_0\cdot S_0}\right) = \overline{L_j} + 10\lg\left(\frac{S}{S_0}\right) = \overline{L_{\mathrm{pE}}} + 10\lg\left(\frac{S}{S_0}\right), \tag{20}$$

where $\bar{j}$ is the surface average of j over the measurement surface, $\overline{L_{\mathrm{pE}}}$ is the level of the surface-averaged sound exposure (dB) over the measurement surface, S is the area of the measurement surface (m^2) and $S_0 = 1\,\mathrm{m}^2$.

Equation (20) corresponds to Eq. (4) for the sound power measured by the free-field method for the sound from a steady sound source. The method described here for determining sound energy level is specified in ISO 3745 (precision grade) and ISO 3744 (engineering grade).

2.2 Sound Pressure Method II (Diffuse Sound Field Method)

When an impulsive or intermittent sound source is located in a diffuse sound field, the following relationship exists between the instantaneous sound power and the sound energy density [21]:

$$\mathrm{P}(t) = V\frac{dE(t)}{\mathrm{d}t} + \frac{c\cdot E(t)}{4}A. \tag{21}$$

The sound energy produced up to some observation time t is found from

$$\int_0^t P(\tau)\mathrm{d}\tau = V\cdot E(t) + \left(\frac{cA}{4}\right)\int_0^t E(\tau)\mathrm{d}\tau, \tag{22}$$

where $P(\tau)$ is the instantaneous sound power of the source (J/s), $E(\tau)$ is the instantaneous sound energy density in the sound field ($\mathrm{J/m}^3$), A is the equivalent sound absorption area of the room (m^2), V is the volume of the room (m^3) and c is the speed of sound at the prevailing air temperature (m/s).

By integrating over a duration that encompasses the complete impulsive or intermittent sound, the sound energy may be calculated as follows:

$$J = \int_0^\infty P(t)\mathrm{d}t = \left(\frac{cA}{4}\right)\int_0^\infty E(t)\mathrm{d}t. \tag{23}$$

The sound energy density can be obtained from a measurement of the space- and time-averaged sound pressure in a diffuse sound field. Equation (23) reduces to

$$J = \left(\frac{A}{4\rho c}\right)\int_0^\infty p^2(t)\mathrm{d}t. \tag{24}$$

Sound energy level may then be calculated from

$$L_J = L_{\mathrm{pE}} + 10\lg\left(\frac{A}{A_0}\right) - 6\mathrm{dB}. \tag{25}$$

Equation (25) corresponds to Eq. (7) for the determination of sound power level by the diffuse field method for steady time-invariant sound sources. The method for determining sound energy level described above is specified in ISO 3741 (precision grade), ISO 3743-1 (engineering grade) and ISO 3747 (engineering grade).

2.3 Sound Intensity Method

The sound energy, J in joules, emitted from an impulsive sound source can be obtained by integrating the sound intensity exposure j over the area of a measurement surface surrounding the source as

$$J = \int_S j\cdot\mathrm{d}s \approx \sum_i j_i\cdot S_i, \tag{26}$$

where j_i is the sound intensity exposure measured at the ith measurement point ($\mathrm{J/m^2}$) and S_i is the area associated with the ith measurement point ($\mathrm{m^2}$).

Consequently, the sound energy level may be expressed, in decibels, as

$$L_J = 10\lg\left(\frac{J}{J_0}\right) = 10\lg\left(\frac{\sum_i j_0\cdot 10^{L_{j,i}/10}\cdot S_i}{J_0}\right) = 10\lg\left(\frac{\sum_i \left(10^{L_{j,i}/10}\cdot S_i\right)}{S_0}\right). \tag{27}$$

Equation (27) corresponds to Eq. (12) for the determination of sound power level by the sound intensity method. The method described by Eq. (27) has not yet been specified in ISO standards.

References

1. ISO 3740:1980, "Acoustics – Determination of sound power levels of noise sources – Guideline for the use of basic standards".
2. ISO 3741:1999, "Acoustics – Determination of sound power levels of noise sources using sound pressure – Precision methods for reverberation rooms".
3. ISO 3743-1:1994, "Acoustics – Determination of sound power levels of noise sources – Engineering methods for small, movable sources in reverberant fields – Part 1: Comparison method for hard-walled test rooms".
4. ISO 3743-2:1994, "Acoustics – Determination of sound power levels of noise sources – Engineering methods for small, movable sources in reverberant fields – Part 2: Methods for special reverberation test rooms".
5. ISO 3744:1994, "Acoustics – Determination of sound power levels of noise sources using sound pressure – Engineering method in an essentially free field over a reflecting plane".
6. ISO 3745:2003, "Acoustics – Determination of sound power levels of noise sources using sound pressure – Precision methods for anechoic and hemi-anechoic rooms".
7. ISO 3746:1995, "Acoustics – Determination of sound power levels of noise sources – Survey method using an enveloping measurement surface over a reflecting plane".
8. ISO 3747:2000, "Acoustics – Determination of sound power levels of noise sources using sound pressure – Comparison method for use in situ".
9. ISO 9614-1:1993, "Acoustics – Determination of sound power levels of noise sources using sound intensity – Part 1: Measurement at discrete points".
10. ISO 9614-2:1996, "Acoustics – Determination of sound power levels of noise sources using sound intensity – Part 2: Measurement by scanning".
11. ISO 9614-3:2002, "Acoustics – Determination of sound power levels of noise sources using sound intensity – Part 3: Precision method for measurement by scanning".
12. M. Vorländer and G. Raabe, "Calibration of reference sound sources", Acustica, 81, 247–263 (1995).
13. S. Uosukainen, "On the use of the Waterhouse correction", J. Sound Vibration, 186(2), 223–230 (1995).
14. A.J. Campanella, "Calibration of reference sound sources, a US perspective", Proc. NOISE-CON 96, 931–936 (1996).
15. M.A. Nobile, "Calibration of reference sound sources according to ANSI S12.5 and ISO/DIS 6926", Proc. NOISE-CON 96, 937–942 (1996).
16. ISO 6926:1999, "Acoustics – Requirements for the performance and calibration of reference sound sources used for the determination of sound power levels".

17. H. Tachibana, H. Yano and M. Koyasu, "Acoustic measurements using an impulsive reference source", Proc. INTER-NOISE 96, 2673–2678 (1996).
18. H. Tachibana and H. Yano, "Change of sound power of reference sound sources influenced by boundary conditions measured by the sound intensity technique", Proc. INTER-NOISE 89, 1009–1014 (1989).
19. H. Yano, H. Tachibana and H. Suzuki, "Experimental study on the accuracy of sound power determination by sound intensity using artificial complex sound sources", Proc. 17th ICA (2001).
20. H. Tachibana and H. Suzuki, "Procedures for the precision sound power determination by S.I. scanning method", Proc. 16th ICA and 135th ASA (1998).
21. H. Tachibana, H. Yano and K. Yoshihisa, "Definition and measurement of sound energy level of a transient sound source", J. Acoust. Soc. Jpn, 8(6) (1987).
22. H. Yano, H. Tachibana and M. Koyasu, "Determination of 'Sound Energy Level' for sound sources which emit single bursts of sound energy by the free field method and diffuse field method", Proc. INTER-NOISE 98 (1998).
23. IEC 61672-1:2002, "Electroacoustics – Sound level meters – Part 1: Specifications".

52

Outdoor Measurements

Hideki Tachibana[1], *Kohei Yamamoto*[2], *Yasushi Miki*[3], *Yasushi Takano*[4], *and Ichiro Yamada*[5]

[1]*Institute of Industrial Science, University of Tokyo, Japan*
[2]*Kobayasi Institute of Physical Research, Tokyo, Japan*
[3]*Computer Science, Faculty of Engineering, Takushoku University, Tokyo, Japan*
[4]*1st Department, Mechanical Engineering Research Laboratory, Hitachi, Ltd, Japan*
[5]*Aviation Environment Research Center, Airport Environment Improvement Foundation, Tokyo, Japan*

1 • Measurement of Acoustical Properties of Ground Surfaces – *Hideki Tachibana*

Airborne sound propagating outdoors from a source to a receiver is attenuated by (1) wave divergence (geometric spreading losses), (2) absorption processes in the atmosphere associated with molecular relaxation effects, (3) shielding and scattering from objects encountered along the propagation path, and (4) absorption of sound energy caused by interaction between the sound wave and the surface of the ground. This chapter describes two techniques that may be used to quantify the acoustical properties of ground surfaces for use in estimating the propagation losses associated with ground-absorption effects. Extensive signal processing is needed to implement each method.

1.1 Normal-Incidence Sound-Absorption Characteristics of a Ground Surface

For a point source of sinusoidal sound located in a free field, the sound pressure at a point at distance r from the source decreases as the distance increases according to the following relation:

$$p(r) = \frac{A\mathrm{e}^{-\mathrm{j}kr}}{r}, \tag{1}$$

where A is the sound pressure amplitude factor determined by the strength of the sound source, k is the wavenumber ($= 2\pi f/c$), f is the frequency of the sound, and c is the speed that the sound wave propagates through the air. In Eq. (1), the term describing the time variation of the sound wave has been omitted.

When a sound source (loudspeaker) is located above the ground as shown in Figure 1, the sound pressures at the two microphone positions, Mic.1 and Mic.2, have contributions from sound that is received directly from the source and from sound that is reflected from the ground. The direct and reflected contributions are expressed as follows:

$$\text{Direct sound at Mic.1}: \quad p_{1\mathrm{i}} = A\frac{\mathrm{e}^{-\mathrm{j}k(D-d-d_0)}}{D-d-d_0} = A\cdot W_1, \tag{2}$$

$$\text{Reflected sound at Mic.1}: \quad p_{1\mathrm{r}} = r_\mathrm{p}\cdot A\frac{\mathrm{e}^{-\mathrm{j}k(D+d+d_0)}}{D+d+d_0} = r_\mathrm{p}\cdot A\cdot W_2, \tag{3}$$

$$\text{Direct sound at Mic.2}: \quad p_{2\mathrm{i}} = A\frac{\mathrm{e}^{-\mathrm{j}k(D-d_0)}}{D-d_0} = A\cdot W_3, \tag{4}$$

$$\text{Reflected sound at Mic.2}: \quad p_{2\mathrm{r}} = r_\mathrm{p}\cdot A\frac{\mathrm{e}^{-\mathrm{j}k(D+d_0)}}{D+d_0} = r_\mathrm{p}\cdot A\cdot W_4, \tag{5}$$

where r_p is the complex sound-pressure reflection coefficient.

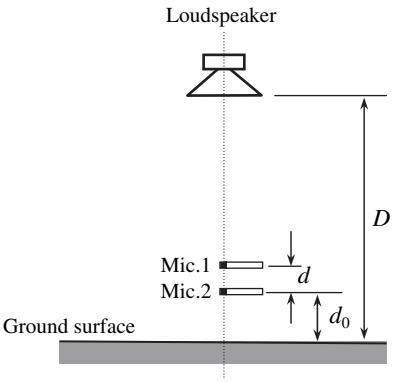


FIGURE 1 Experimental setup for measurement of normal-incidence sound-absorption characteristic of a ground surface by the 2-microphone method.

Consequently, at any instant of time, the sound-pressure signal from the two microphones is the sum of the directly incident sound and the reflected sound as given by

$$\text{Mic.1}: \qquad p_1 = p_{1\text{i}} + p_{1\text{r}} = A(W_1 + r_\text{p} \cdot W_2), \tag{6}$$

$$\text{Mic.2}: \qquad p_2 = p_{2\text{i}} + p_{2\text{r}} = A(W_3 + r_\text{p} \cdot W_4). \tag{7}$$

The reflection coefficient in Eqs. (6) and (7) is determined from

$$r_\text{p} = \frac{(H_{12}W_1) - W_3}{W_4 - (H_{12}W_2)}, \tag{8}$$

where the transfer function H_{12} or the ratio of the instantaneous sound pressures is given by

$$H_{12} = \frac{p_2}{p_1}.$$

The normal-incidence sound-attenuation coefficient α_0 can be calculated from the knowledge of the complex reflection coefficient by means of the following expression:

$$\alpha_0 = 1 - |r_\text{p}|^2. \tag{9}$$

1.2 Effective Flow Resistivity of a Ground Surface

The effective flow resistivity of the ground surface $\sigma_\text{e}\left(\text{Pa}/(\text{m}^2/\text{s})\right)$ is often used in calculations of the excess attenuation caused by the ground effects. The curve-fitting method described here may be used to estimate the effective flow resistivity.

The sound-pressure distribution over a horizontal ground surface is measured. The height of the sound source and the microphone are maintained within 1 m above the ground in order to observe the steep dip of sound pressure at a certain frequency resulting from the interference between the direct and reflected sound waves. Removing the attenuation caused by wave divergence (geometric spreading) from the results yields the measured excess attenuation caused by the ground effects.

The excess attenuation may be calculated from theoretical considerations for the same geometrical conditions as used in the field to measure excess attenuation. The following equation is used to calculate the complex specific acoustic impedance Z_s of the ground surface. This equation was proposed by Miki [1] as a modification of the empirically derived equation developed by Delany and Bazley [2]:

$$Z_\text{s} = 1 + 0.07\left(\frac{f}{\sigma_\text{e}}\right)^{-0.632} - \text{j}\left[0.107\left(\frac{f}{\sigma_\text{e}}\right)^{-0.632}\right]. \tag{10}$$

In practice, the effective flow resistivity term in Eq. (10) is varied until the calculation of the specific acoustic impedance yields a curve of excess attenuation as a function of frequency that provides a good fit to the experimentally determined excess attenuation [3,4]. The result of this process gives the best-fit value for the flow resistivity of the ground surface at the time of the measurements.

Figure 2 shows an example of the application of the method. Figure 2a shows a setup for the experimental determination of excess attenuation. A wideband random-noise signal was used as the source of the sound signal.

Figure 2b shows the result of the calculated curve fit to the measured excess attenuation data for the experimental conditions given in Figure 2a. Vertical bars through the measured data points

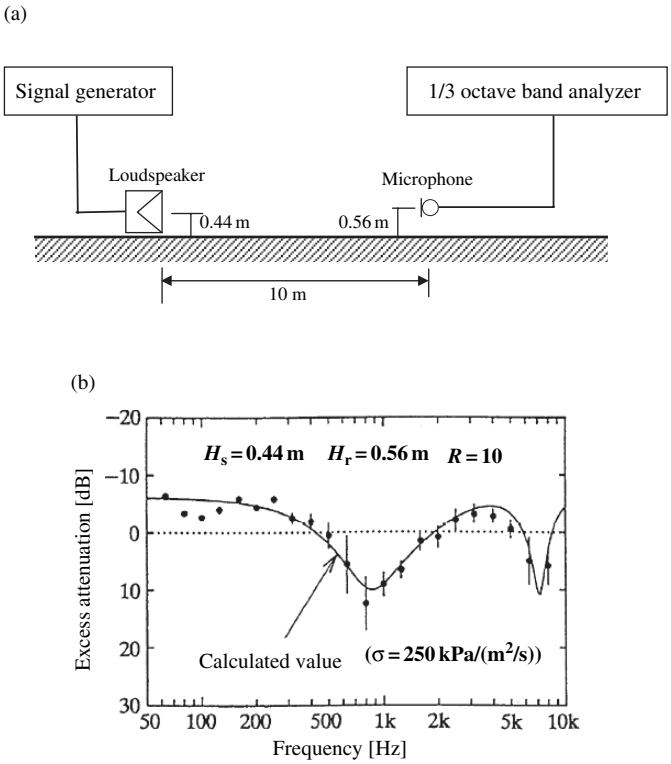


FIGURE 2 Illustration of the determination of the effective flow resistivity of a grassland surface. (**a**) Measurement of sound pressure distribution over grassland; (**b**) estimation of the effective flow resistivity of the grassland by curve-fitting method.

indicate the estimated uncertainty of the data. For this case, the best-fit curve was obtained for a flow resistivity of 250 kPa/(m^2/s).

2 • Measurement of Shielding Effect of Noise Barriers – *Kohei Yamamoto and Yasushi Miki*

Noise barriers are widely applied to highways, railways, airports, industrial areas and all other places, where noise pollution is a serious problem. There are basically two types of noise barriers; one is a reflective barrier with a hard surface made of concrete, plastic, wood, steel aluminum, glass, etc., while the other is an absorptive barrier with a soft surface covered with absorptive material such as mineral fiber, ceramics, porous concrete, aluminum fiber, perforated plastics, cement bonded wood wool or wood chips, etc. The advantage of absorptive treatment on the surface is that it reduces the sound energy of diffracted wave propagated into the shadow zone. It also reduces sound reflections between walls of a parallel barrier, resulting in suppression of noise build-up due to multiple reflections. Figure 3 shows an illustration of a reflective barrier constructed on the shoulder of a highway. The barrier is made of transparent panels with a noise reducing device attached on the top edge.

In general, the height of noise barriers ranges between 2 and 3 m, but it becomes up to 8–10 m for roads and airports dependent on how serious the noise problem is [5]. Various factors affect the performance of noise barriers: geometrical arrangement of



FIGURE 3 A highway noise barrier: a composite of transparent panel and noise reducing device.

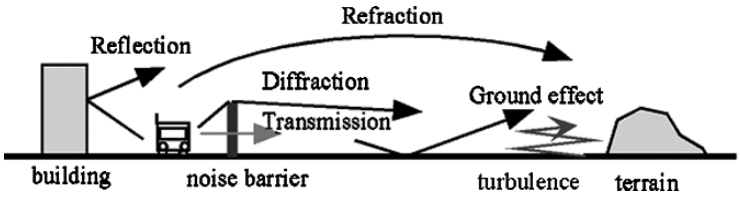


FIGURE 4 Factors affecting the effectiveness of outdoor noise barriers.

the sound source, barrier and receivers, the height of barrier, angle of sound incidence, sound frequency, type of the sound source, transmission loss of barrier wall, ground and meteorological conditions, sound reflections from the surroundings of the sound source and receivers and so on (see Figure 4). The effectiveness of barriers is tested in various stages of the installation process. In developing a barrier of new type, it is often tested by scale model experiment in a laboratory such as an anechoic or hemi-anechoic room. For a full-size trial product, it is determined at an outdoor test field. After the installation of barriers at a target site, it is verified by in situ measurement for the purpose of a routine inspection or diagnostic evaluation of barrier performance.

2.1 Measurement of Sound Attenuation of Barriers

The effect of a barrier that shields sound sources is originally evaluated in terms of sound attenuation, sometimes called excess attenuation EA. It is theoretically defined as the attenuation of sound in excess of that due to geometrical spreading in the free field as follows:

$$\mathrm{EA} = -10\ \log_{10}\left|\frac{\Phi_{\mathrm{behind}}}{\Phi_{\mathrm{free}}}\right|^2, \tag{11}$$

where Φ_{behind} and Φ_{free} are complex velocity potentials or sound pressures at a position behind the barrier and at the same position without the barrier, respectively (see Figure 5).

Sound attenuation is most of the time evaluated in a laboratory where free-field condition is satisfied. Maekawa's experiment

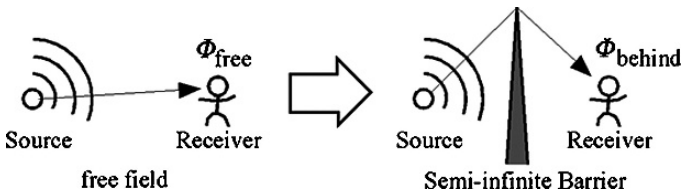


FIGURE 5 Sound attenuation by a barrier.

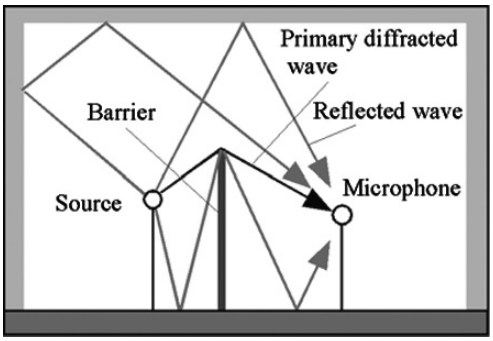


FIGURE 6 Sound paths in barrier attenuation experiments in the laboratory.

is a representative example where the determination of sound attenuation by a thin screen was carried out in an anechoic room [6].

In actual tests of sound attenuation of a barrier in the laboratory, as is shown in Figure 6, the sound wave observed by the microphone includes not only a primary wave that comes directly from the sound source after diffracted around the top edge of the barrier, but also several redundant waves that are reflected and/or diffracted from the surface of the floor and surrounding walls. In order to determine the sound attenuation of the barrier, it is necessary to extract only the primary diffracted wave from the microphone output. Under such situations, impulse sound is preferred to stationary sound as the test signal. One way to realize an impulsive test signal is to use a spark discharge as the sound source. It is, however, not always stable and non-directional. Alternative means is to use a short-duration burst of tone or band noise, but the ability of sound emission is limited dependent on the performance of electro-acoustic transducers, and the sound level observed in the shadow behind the barrier is very low compared to that without the barrier, resulting in a low signal-to-noise ratio (S/N) and an insufficient accuracy of measurement of sound attenuation.

One solution for the improvement of S/N ratio and signal stability is to equivalently realize a high gain in amplitude with the aid of a time-stretched pulse technique [7,8]. The basic idea is to disperse the sound energy of an impulse confined in a moment over a wide interval by deforming the phase relation among frequency components. The first step is to define frequency characteristics $X(k)$ of a target impulsive test signal $x(n)$; the real

part is defined as

$$\begin{aligned} \mathrm{Re}\,[X(k)] &= \mathrm{e}^{-\left(\frac{k-a}{b}\right)^c} && (k=0,1,\ldots,K/2-1) \\ &= 0 && (k=K/2) \\ &= Re\,[X(K-k)] && (k=K/2+1,\ldots,K-1) \end{aligned} \tag{12}$$

where $a=900,\ b=800,\ c=12,\ K=4,096$ and the imaginary part is assumed to be constantly zero. See Figure 7a $X(k)$ and 7b $x(n)$. Next step is to stretch $x(n)$ over a wide span by applying the following non-linear phase-shifting filter $h(n)$:

$$\begin{aligned} H(k) &= \mathrm{e}^{\mathrm{j}dk^2} && (k=0,1,\ldots,K/2-1) \\ &= 1 && (k=K/2) \\ &= H^*(K-k) && (k=K/2+1,\ldots,K-1) \end{aligned} \tag{13}$$

where $d=0.0012$. Figure 7c shows the waveform of the time-stretched signal $y(n)$, which is applied to the acoustic system

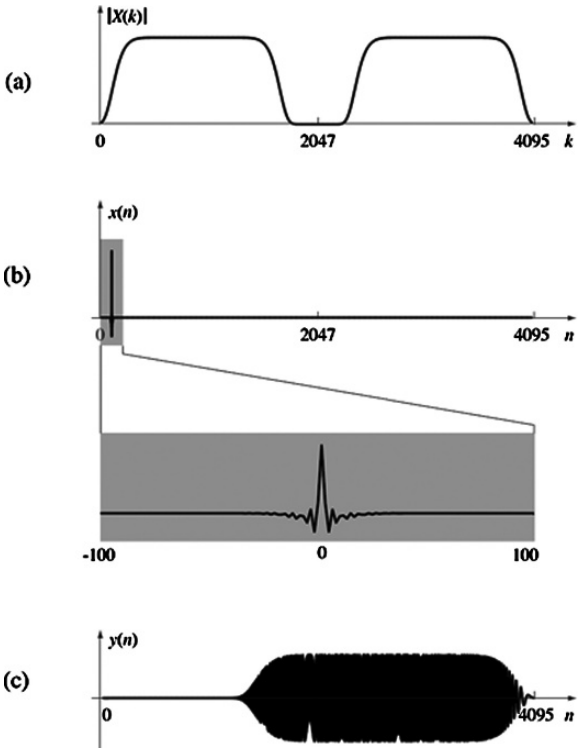


FIGURE 7 (**a**) Magnitude of frequency characteristics, (**b**) waveform of an impulsive test signal and (**c**) time-stretched signal waveform.

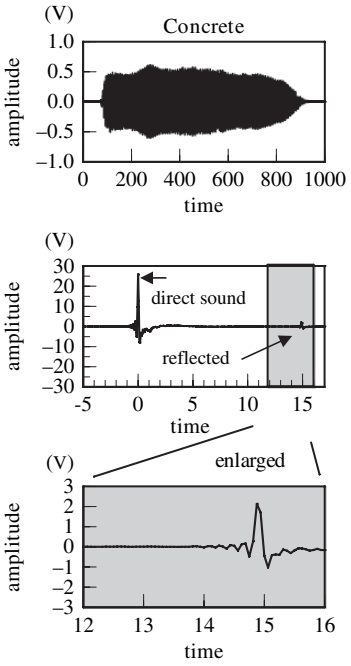


FIGURE 8 (**a**) Time-stretched signal reflected from a concrete surface and (b) demodulated waveform of the impulsive test signal.

including the target barrier. Finally, the sound observed by a microphone behind the barrier is demodulated or, in other words, time-compressed via an inverse filter $h^{-1}(n)$ to get it back to the impulsive waveform, from which we can easily extract a part of the primary diffracted wave. Figure 8 shows an illustration of demodulated waveform of reflection from a concrete surface [9]. Time-stretched pulse technique makes it possible to improve the S/N ratio by 30 dB, compared to the direct use of the impulsive test signal [8].

Another solution for the improvement of S/N ratio and signal stability is to use a binary-valued pseudo-random sequence signal, i.e., an M-sequence signal $m(t)$ as the test sound. As is well known, the auto-correlation function of $m(t)$ is close to an impulse if the period of the sequence is sufficiently long. If we calculate the cross-correlation function of $m(t)$ and the output of a target acoustic system, it becomes an estimation of the impulse response of the system. It is often called "M-sequence correlation method" or otherwise "MLS (maximum length sequence) method", which is described a little more in detail in Section 52.3, widely used for the measurement of short-range sound propagation [10–13]. Finally, however, it should be noted that, in case of long-distance sound propagation, the

time-stretched pulse technique and the *M*-sequence correlation method become ineffective because of deterioration of signal coherence due to air turbulence, etc. In such cases special techniques such as *M*-sequence modulation method and synchronized integration method become superior to the time-stretched pulse technique and *M*-sequence correlation method.

2.2 Measurement of Insertion Loss of Barriers

The definition of sound attenuation is simple, but it is not necessarily appropriate for outdoor noise barriers erected on the ground. In this case, the effectiveness of barriers is evaluated in terms of insertion loss D_{IL}, which is defined as the difference in sound-pressure level at a point between before and after the installation of a barrier as follows:

$$D_{\mathrm{IL}} = L_{\mathrm{p}}(\text{before}) - L_{\mathrm{p}}(\text{after}), \tag{14}$$

where L_{p} (before) and L_{p} (after) are sound-pressure levels at the position before and after the installation of the barrier, respectively (see Figure 9). Note that insertion loss corresponds to sound attenuation if the ground is absent.

The insertion loss represents a net efficiency of a barrier for a given position of sound sources and receivers. It generally contains effects of various conditions unique to the test site such as terrain profiles, ground conditions and surrounding reflection objects. The insertion loss of outdoor noise barriers can be determined by in situ measurement based on methods specified in ISO 10847 [14], which specifies three types of sound sources for the determination of insertion loss: a natural source, a controlled natural source and a controlled artificial source. Considering uncontrollable disturbances due to meteorological and atmospheric conditions and contamination by background noise, a controlled artificial source such as a loudspeaker is preferable to outdoor experiments. However, it should be noted that such a controlled artificial source may not be appropriate to simulate a line or area source.

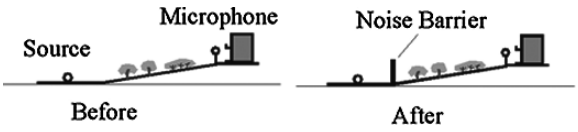


FIGURE 9 Comparison of situations before and after the installation of a noise barrier.

3 • Measurement of Sound Attenuation for Long-Distance Propagation Outdoors – *Yasushi Miki*

The measurement of outdoor sound propagation over ground is indispensable to the evaluation of noise impact of sound sources such as road traffics and factories. In the long distance from the source, however, it becomes difficult without special techniques of signal processing because of deterioration of signal coherence under circumstances of air turbulence and background noise including wind noise. This chapter describes two techniques, which may be used to estimate sound propagation under such conditions, being based on modulation–demodulation of a test sound.

3.1 *M*-Sequence Modulation Method

Suppose that two signals $x(t)$ and $y(t)$ are the input and output of a linear acoustic system with an impulse response $h(t)$, respectively. The cross-correlation function $\phi_{xy}(\tau)$ between the two signals is expressed using the auto-correlation function $\phi_{xx}(\tau)$ of the input as follows:

$$\phi_{xy}(\tau) = \int_{-\infty}^{\infty} h(u)\,\phi_{xx}(\tau - u)\,\mathrm{d}u. \tag{15}$$

If we use white noise as the input, the cross-correlation function becomes the impulse response itself. An *M*-sequence signal $m(t)$ is frequently substituted for the white noise; it is a binary-valued pseudo-random signal, whose waveform and auto-correlation function $\phi_{mm}(\tau)$ are shown in Figure 10. This technique is called the "*M*-sequence correlation method" [11(F-17),15] or otherwise "MLS method" [12]. It is widely used, needless to say, for the measurement of only short-range sound propagation. In case of long-range sound propagation outdoors, air turbulence and background noise seriously deteriorate signal coherenceand spoil the effectiveness of the method. To solve this

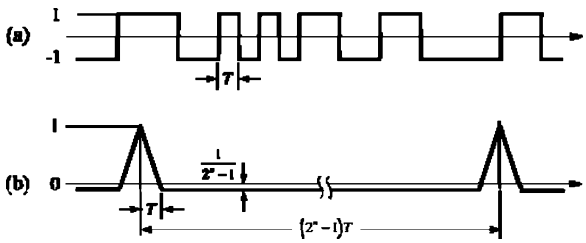


FIGURE 10 (**a**) Waveform of the *M*-sequence $m(t)$ and (**b**) auto-correlation function $\phi_{mm}(\tau)$ of the nth order *M*-sequence.

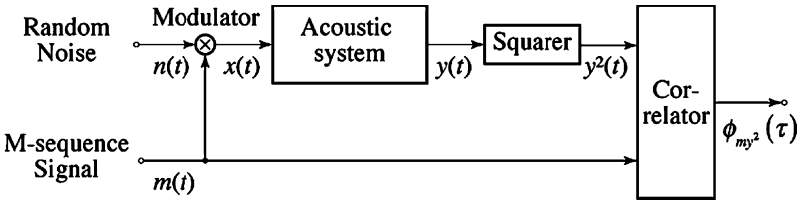


FIGURE 11 Basic system of the *M*-sequence modulation method.



FIGURE 12 (**a**) Test signal $x(t)$ modulated by *M*-sequence and (**b**) the observed waveform $y^2(t)$ under the influence of background noise.

difficulty, a modified method called "*M*-sequence modulation method" was developed [16,17]. Figure 11 shows the principle of the method. Figure 12a shows the input signal $x(t)$, which is a band-limited noise $n(t)$ switched on and off intermittently (or in other words amplitude-modulated) by an *M*-sequence signal $m(t)$:

$$x(t) = \frac{1}{2}\{m(t)+1\} \cdot n(t). \tag{16}$$

If this signal is supplied to a multi-path acoustic system with an impulse response $h(t)$,

$$h(t) = \sum_{i=1}^{N} a_i \delta(t - T_i), \tag{17}$$

where $\delta(t)$ is the Dirac delta function and a_i and T_i are an attenuation factor and a delay time of the ith sound path, respectively, the cross-correlation function $\phi_{my^2}(\tau)$ between the squared output signal $y^2(t)$ and the modulator signal $m(t)$ becomes as follows:

$$\phi_{my^2}(\tau) = \sum_{i=1}^{N} \frac{a_i^2}{2} \phi_{mm}(\tau - T_i)\, \phi_{nn}(0). \tag{18}$$

Since the function $\phi_{mm}(\tau)$ has a triangular waveform as shown in Figure 10b, a_i and T_i can be estimated from the amplitude and delay time of the corresponding peak of $\phi_{my^2}(\tau)$. Figure 12b shows a typical waveform of $y^2(t)$ at an observation point under the influence of background noise.

Practically, it is recommended to make multi-point measurement in which the input signal transmitted to the acoustic system is observed both at a distant observation point and at a reference point near the transmitter simultaneously [17]. The sound signals observed at the two points can be added together and then squared to calculate the cross-correlation function $\phi_{my^2}(\tau)$. A typical illustration is shown in Figure 13, in which there are two triangle correlation peaks; one corresponds to that at the reference point, whereas another at the observation point. The sound-attenuation level ΔL between the two points can be estimated from the following equation:

$$\Delta L = 10 \cdot \lg\left({}^{A_O}/_{A_r}\right), \tag{19}$$

where A_O and A_r are the peak amplitudes of the cross-correlation function at the delay times corresponding to the observation and reference points, respectively. It is reported, e.g., that the method was successfully applied to the measurement of several kilometers sound propagation over sea surface [18].

3.2 Synchronized Integration Method

The *M*-sequence modulation method makes it possible to estimate multi-path sound propagation separately, but it may cause an underestimation of sound attenuation when the path difference is not clear [17]. Thus, a new technique named the synchronized integration method, which is similar to the *M*-sequence modulation method, was developed [19]. Figure 14 shows the principle of the method. A rectangular signal is used as a gate signal instead of the *M*-sequence and an integrator is used instead of the correlator. As shown in Figure 15, at the sound source, a band noise periodically switched on and off by a gate signal $p(t)$ is emitted as a test sound from the loudspeaker intermittently, while at an observation site the sound observed is band-filtered and then squared to be reduced to sound power, which is discretely time-integrated into two digital counters under the

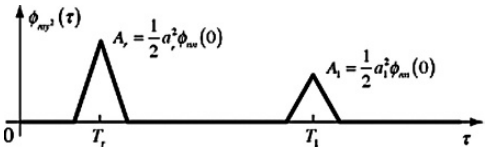


FIGURE 13 Correlation function obtained by 2-point observation.

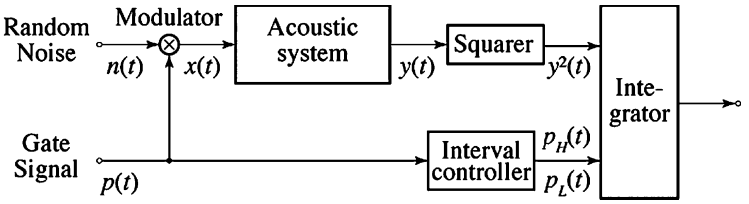


FIGURE 14 Basic system of the synchronized integration method.

control of two gate signals $p_{\mathrm{H}}(t)$ and $p_{\mathrm{L}}(t)$. One counter $E_{\mathrm{T}}^{\mathrm{H}}$ integrates sound power during an interval in which the test sound is surely present together with background noise, while the other $E_{\mathrm{T}}^{\mathrm{L}}$ integrates that of background noise in the absence of the test sound during the following interval of the same length. After the completion of time integration during a predetermined time interval, we can estimate the sound exposure of the test sound except background noise as the difference of the two integrated sound exposures $E_{\mathrm{T}}^{\mathrm{H}}$ and $E_{\mathrm{T}}^{\mathrm{L}}$. If we measure the sound exposure

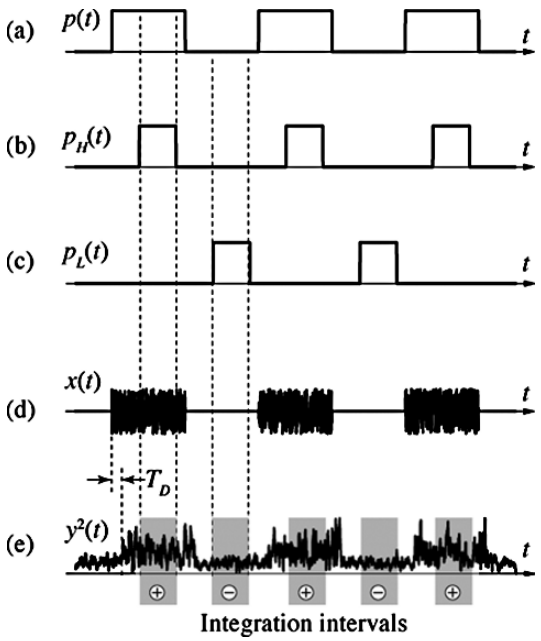


FIGURE 15 Time chart of the synchronized integration method. (**a**) Gate signal $p(t)$. (**b**) Interval signal $p_{\mathrm{H}}(t)$. (**c**) Interval signal $p_{\mathrm{L}}(t)$. (**d**) Test signal $x(t)$. (**e**) Observed sound power $y^2(t)$. T_{D} is the propagation delay. Sound power is integrated upward and downward alternatively according to the signal shown in (**b**) and (**c**).

at an observation site as well as at a reference point near the sound source, we can calculate the sound attenuation ΔL from the source to the observer as follows:

$$\Delta L = 10 \cdot \lg \left(\left(E^H_{T_o} - E^L_{T_o} \right) / \left(E^H_{T_r} - E^L_{T_r} \right) \right), \tag{20}$$

where subscripts o and r mean the observation and reference points, respectively. Note that this method gives us an estimate of the total sound exposure without distinction of multiple sound paths if any.

4 • Noise Sources in Road and Railway Transport – *Yasushi Takano*

Noise from road and rail transport is a serious concern for nearby communities. The only way to reduce noise, without building expensive noise barriers, is to find out where and how the noise is generated and reduce the noise emissions themselves. Researchers and engineers have developed various methods of locating noise sources in moving vehicles. This section discusses some historical methods of locating noise sources in automobiles and railway cars.

If we know the distance and velocity of a moving vehicle, it is possible to use a pass-by measuring apparatus, with super directivity, to identify the noise source distribution. As the vehicle passes by the measuring apparatus, time variations in the measured noise level will be interpreted as noise source distributions along the direction of movement. One of the first attempts with this method was to use a one-dimensional microphone array with fixed directivity as shown in Figure 16 [20,21]. This array consisted of sets of 9-microphone line arrays and was placed parallel to the direction of movement of the target vehicles. The microphones on each line array were separated by a half wavelength of an acoustic wave at the center frequency of a 1/3 or 1/1 octave band. Using a simple analog addition circuit without time delay, this array has a sensitivity perpendicular to the array as shown in the figure. The array system has a constant directivity for the entire frequency range of interest by combining line arrays with suitable microphone spacing. Moritoh et al. placed the array system 25 m away from a railway track, where the environmental noise was evaluated, and found that rail and wheel noise, pantograph noise, and aerodynamic noise can be estimated from the peak values of the measured noise level, as shown in Figure 17. The resolution of the noise source

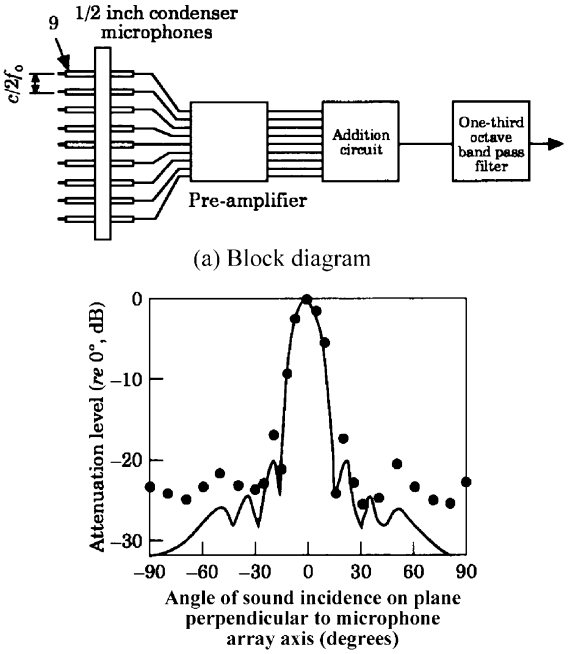


FIGURE 16 Block diagram and directivity of one-dimensional structured array ($f_0 = 1\,\text{kHz}$) [21].

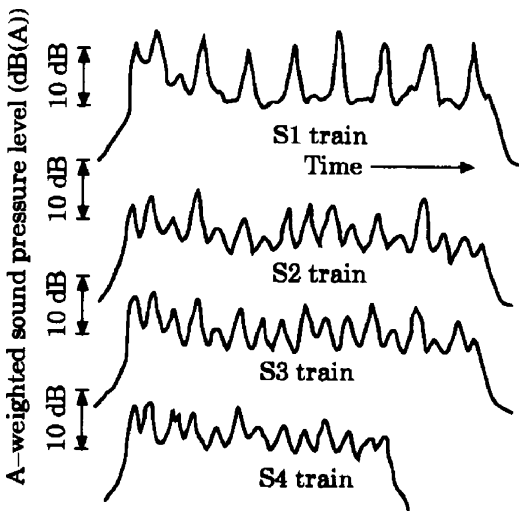
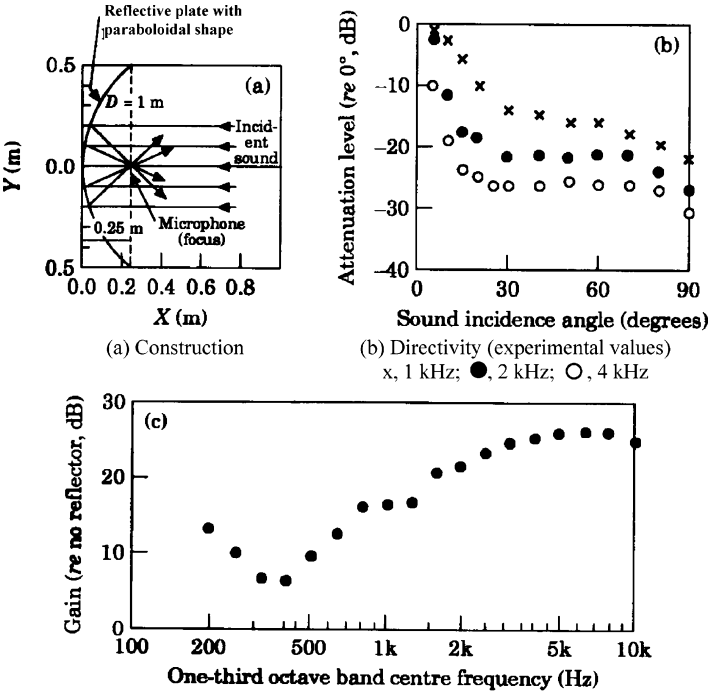


FIGURE 17 Time history of A-weighted sound-pressure level measured with one-dimensional structured array at 25 m distance (trains speed = 220–240 km/h) [21].

depends on the distance and the velocity of the target. This array system is still widely used to locate and analyze high-speed train noise. However, the resolution of noise sources with this line array is several meters. It is therefore difficult to obtain detailed distributions for noise sources.

To localize noise sources with improved resolution, Moritoh et al. [21] developed a parabola-microphone system to obtain more refined information about noise sources in high-speed trains. The device consisted of a microphone and a paraboloidal reflector (parabola apparatus), as shown in Figure 18. Although its directivity and sensitivity depended on the frequency of incident sound, it gave a two-dimensional source distribution and had better definition than the previous microphone array for high-frequency sound. The noise level time history obtained from the parabola apparatus (1/3 octave-band center frequency 8 kHz) is shown in Figure 19. Noise generated from windows, doors, and gaps between adjacent cars can clearly be seen. Detailed observation revealed that this noise was generated from the ridges of the window and doorframes.



(a) Construction

(b) Directivity (experimental values)
x, 1 kHz; ●, 2 kHz; ○, 4 kHz

(c) Gain characteristics (incident angle = 0°, experimental values)

FIGURE 18 Structure and performance of parabola apparatus [21].

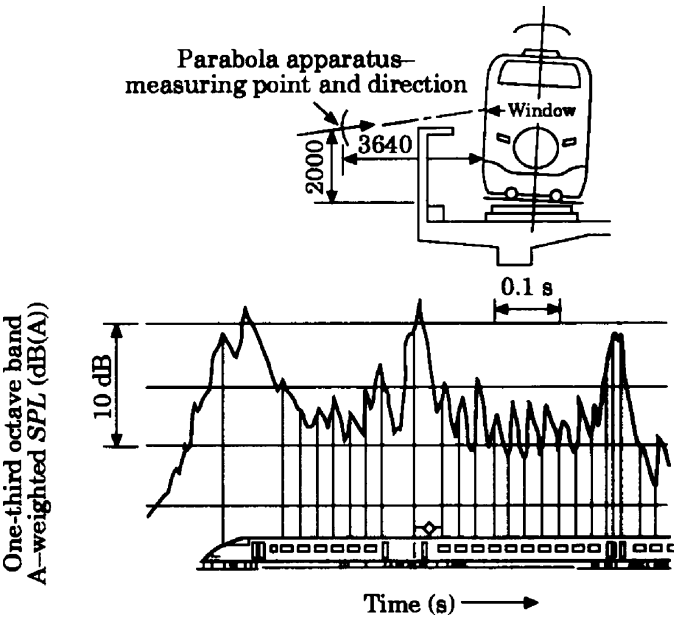


FIGURE 19 Time history of A-weighted one-third octave band ($f_0 = 8\,\text{kHz}$) sound-pressure level measured with parabola apparatus (train speed $= 270\,\text{km/h}$) [21].

It is also possible to place measuring apparatus on vehicles themselves. Oshino and Tachibana [22] developed a special automobile equipped with an automatic microphone scanning system, as is shown in Figure 20 [22]. They measured the sound

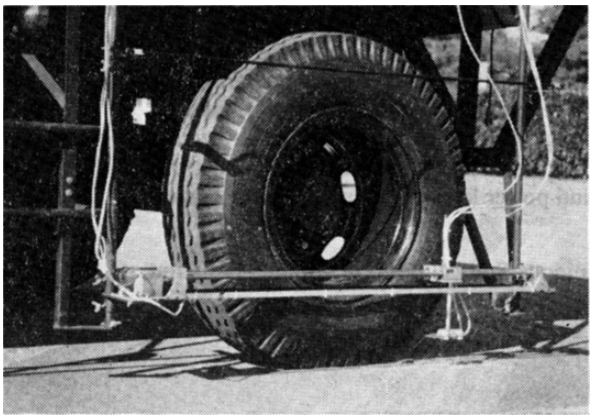
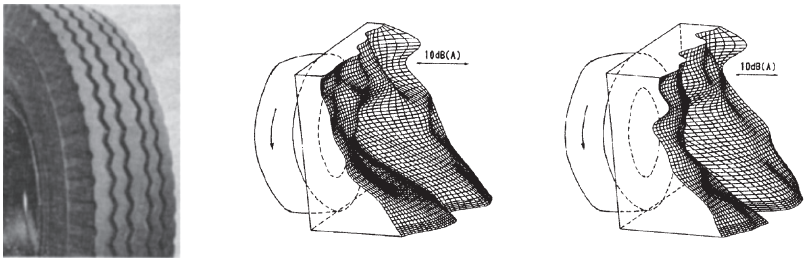


FIGURE 20 Automatic intensity probe scanning system [22].



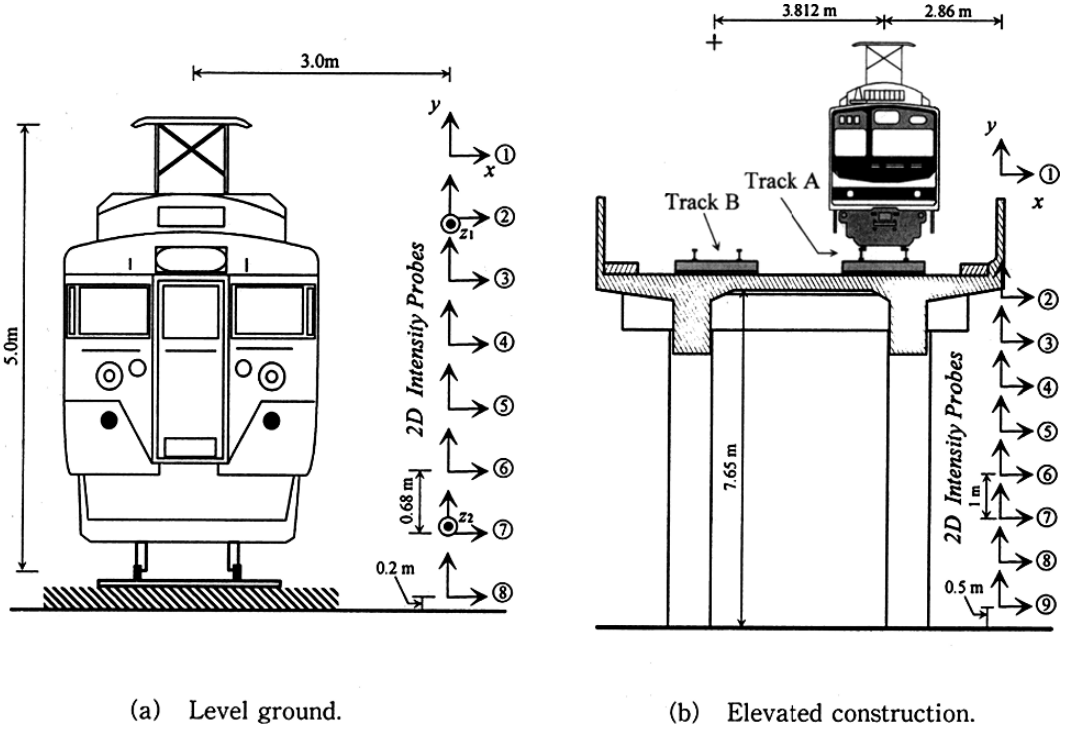
(a) Tread pattern (b) Overall intensity (c) 1 kHz in 1/3 octave band intensity

FIGURE 21 Measured normal sound intensity radiated from tire [22].

radiation characteristics of actual rolling tires with different tread patterns with a 2-microphone sound intensity technique. The normal sound intensity distribution in a plane parallel to the sidewall of the tire was measured during actual running conditions. Their results revealed that it was possible to identify the source positions on a rolling tire on an actual road surface with this method, as we can see in Figure 21. They confirmed that noise radiation patterns vary according to differences in tread patterns. However, this method requires both a large scanning system to cover the entire noise source region and a way of reducing the aerodynamic noise of the measurement apparatus itself to achieve a good signal-to-noise ratio.

Sound intensity probes can also be configured as an array for pass-by measurement [23]. Kawahara et al. [24] developed sets of one-dimensional intensity probes, aligned vertically, to draw a contour map of noise source distributions from a high-speed train. Hidaka et al. [25] developed a system using a two-dimensional intensity probe, also aligned vertically, to measure the intensity vector to establish the direction of the noise source. Joo et al. [26] developed a three-dimensional intensity sensor array system to measure the sound intensity vectors of noise radiated from railways in three dimensions. The measurement setup is outlined in Figure 22. Intensity vectors on the vertical and horizontal planes, which indicate the location of noise sources, are shown in Figures 23 and 24. An intensity contour map can also be drawn to locate probable noise sources as in Figure 25.

Developments in multi-channel data acquisition systems have enabled us to use even a larger number of sensors to measure noise source distributions with the digital beamforming technique. Barsikow et al. [27,28] developed a one-dimensional microphone array system whose directivity was controlled digitally. They demonstrated that Doppler effects can



(a) Level ground. (b) Elevated construction.

FIGURE 22 Arrangement of intensity probes at test site [26].

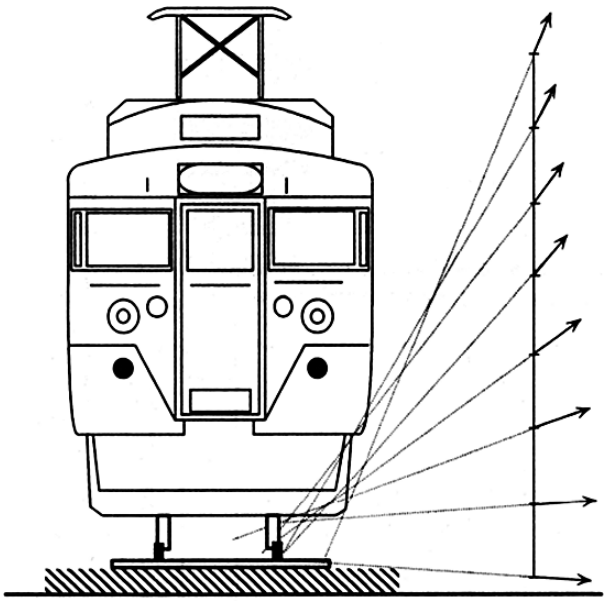


FIGURE 23 Sound intensity vectors for radiated noise from railway car [26] (400 Hz 1/3 octave band in *X*–*Y* plane).

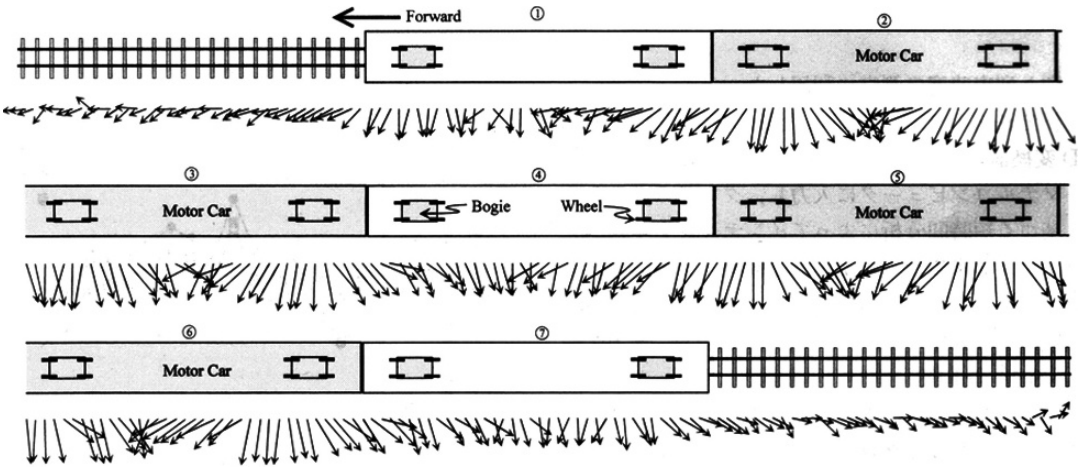


FIGURE 24 Sound intensity vectors in 400 Hz 1/3 octave band in X–Z plane calculated from measurement results at point 7 in Figure 22a [26].

be canceled by correcting time delay, which varies with time, using Eq. (21) [29]:

$$s_i(t) \approx \frac{\sum_{j=1}^{N} W_j p_j(t + r_{ij}(t)/c)}{\sum_{j=1}^{N} W_j}, \tag{21}$$

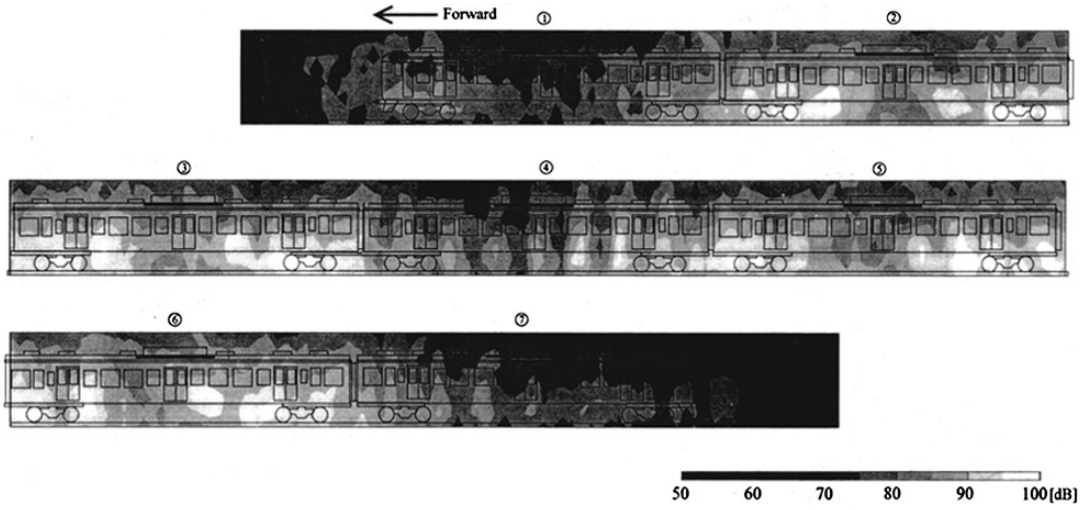
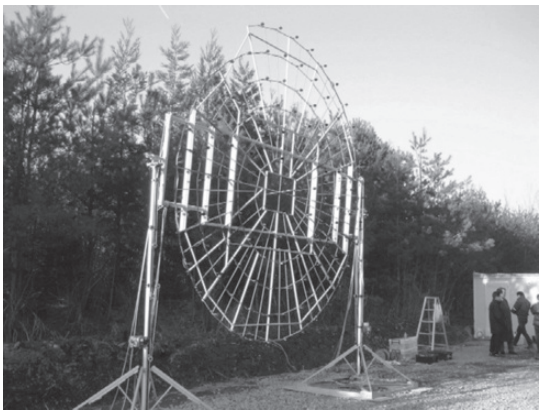


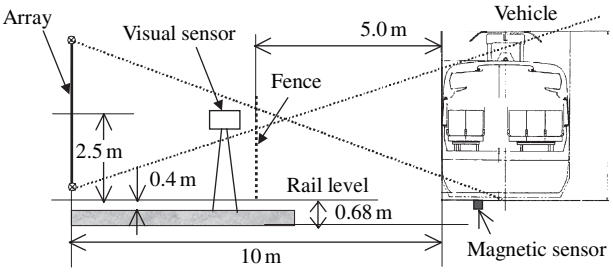
FIGURE 25 Sound intensity contour map in 500 Hz octave band [26].

where $p_j(t)$ = sound pressure at jth microphone (Pa), $s_i(t)$ = sound pressure at unit distance from ith point source (Pa), $r_{ij}(t)$ = distance from ith point source to jth microphone (m), and W_j = window function.

By combining two line arrays, Schmitz et al. and Torii et al. developed cross or X-shaped microphone array systems and succeeded in measuring two-dimensional noise source distributions from high-speed trains [30–32]. However, the cross or X-shaped arrays had large sidelobe levels, which distorted the noise source distributions [33]. Arrays, with an optimized microphone arrangement, have been developed to reduce the sidelobe level [34, 35]. Takano et al. developed a 114-channel spiral-shaped array, synchronized with a visual sensor to acquire a picture of a moving source during acoustical measurement, as can be seen in Figure 26. The efficiency of the system was verified through experimental results using automobiles with speakers as shown in Figure 27.

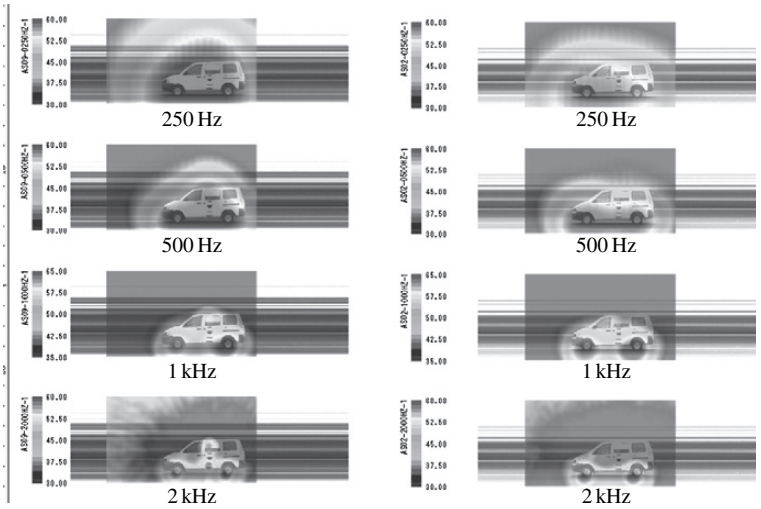


(a) 114 ch spiral array



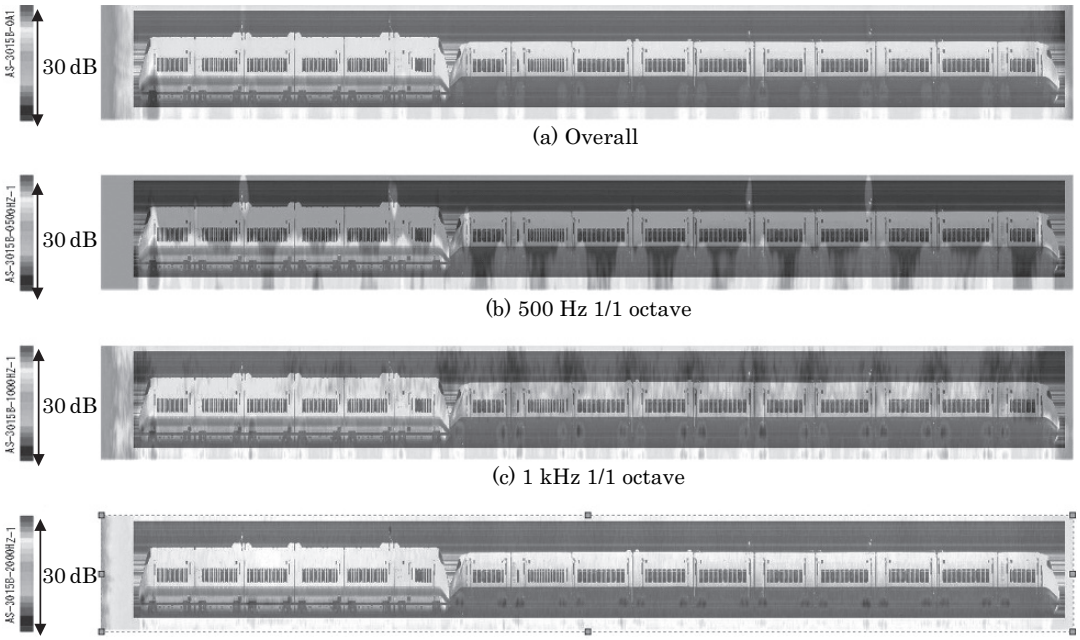
(b) Measurement setup

FIGURE 26 Measurement setup for random array for railways (114-channel spiral array) [37].



(a) Two speakers 1 m apart (b) Speakers off (background noise)

FIGURE 27 Measured results for random array in Figure 26 (two speakers, radiating random noise, are attached to automobile running at 100 km/h) [37].



(a) Overall

(b) 500 Hz 1/1 octave

(c) 1 kHz 1/1 octave

(d) 2 kHz 1/1 octave

FIGURE 28 Measured results for random array in Figure 26 (Shinkansen train running at 275 km/h) [37].

Noise sources were visualized within 50 cm at 1 kHz with a microphone array 10 m away from the source. This system was also applied to high-speed trains to locate noise source distributions as shown in Figure 28 [36,37].

5 • Detection of Aircraft Sound Events – *Ichiro Yamada*

Sound source discrimination is essential to the unmanned monitoring of aircraft sound. Usually, unmanned sound monitors have a simple function to detect single sound events using thresholds on sound level and duration, but it is not always effective because there are many single events similar to aircraft sound. Non-acoustic information such as flight information and radar data has come into increasing use for classifying a pile of sound events. Change in magnitude of secondary radar radio wave from flying aircraft is also applied to the discrimination of aircraft sound [38,39]. Such means, however, neither guarantee that sound events are really caused by aircraft movements even if the time of occurrence coincides with each other nor may be always available. Therefore, it is essential to classify sound events either as aircraft or as non-aircraft by applying distinctive acoustic features of the sound itself.

Sound spectral information is available for the discrimination of sound sources. One work deals with acoustic recognition of aircraft types in flight using a discriminatory analysis based on maximum-held third octave-band spectra of sound events [40]. Others discuss heuristic neural network techniques based on sound spectra for the recognition of aircraft types [41]. These methods may realize high performance in sound source discrimination, but it should be reminded that sound characteristics in the frequency region higher than 1 kHz may greatly vary dependent on climate changes, which means that renewal of reference patterns or re-training of the network is inevitable for keeping the performance.

Information on sound arrival direction brings us another means of discriminating sound sources. Aircraft fly-over sounds usually come down from above, while other sounds come horizontally along the ground or come up from below. This distinction leads us to a simple cross-correlation method of detecting aircraft flyover sound based on elevation angle of sound incidence, which can be estimated from the peak time delay of cross-correlation function between sound signals observed by a pair of microphones arranged vertically [42,43]. Pre-whitening of sound signals using pre-emphasis and inverse filtering showed to bring an improvement in the peak time delay estimation of the

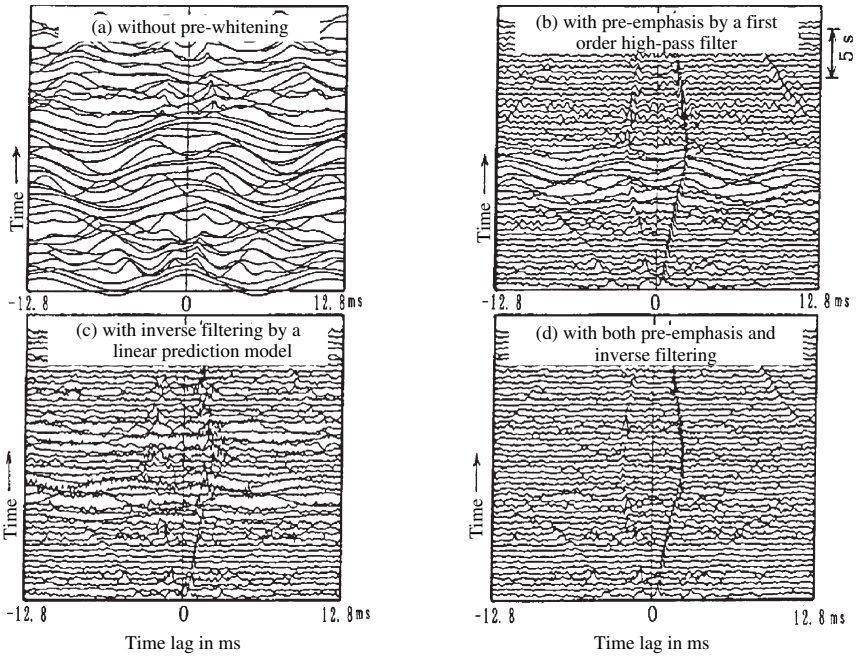


FIGURE 29 Temporal changes of cross-correlation function between propeller aircraft fly-over sound observed by microphones arranged vertically above ground.

cross-correlation function [44]. Figure 29 shows an illustration of the pre-whitening process applied to cross-correlation function of a propeller aircraft fly-over sound.

The cross-correlation method was applied to the measurement of aircraft flight position, being based on the principle of triangulation [45]. Four-microphone array systems were located at two sites facing each other across the flight path and elevation angle of arrival sound was independently estimated at each site from peak time delays of cross-correlation functions.

The cross-correlation method was further extended to a generalized method of classifying sound sources, based on information of three-dimensional sound arrival direction expressed in azimuth ϕ and elevation θ or otherwise in a unit direction vector $\overrightarrow{V}$ on the unit sphere [46–48]. Suppose that an incident plane sound wave is observed by three microphones pairs installed orthogonal to each other at a place, as is shown in Figure 30. The instantaneous three-dimensional sound arrival direction can be estimated using Eq. (22) only from peak time delays (τ_x, τ_y and τ_z) of cross-correlation functions of observed sound signals,

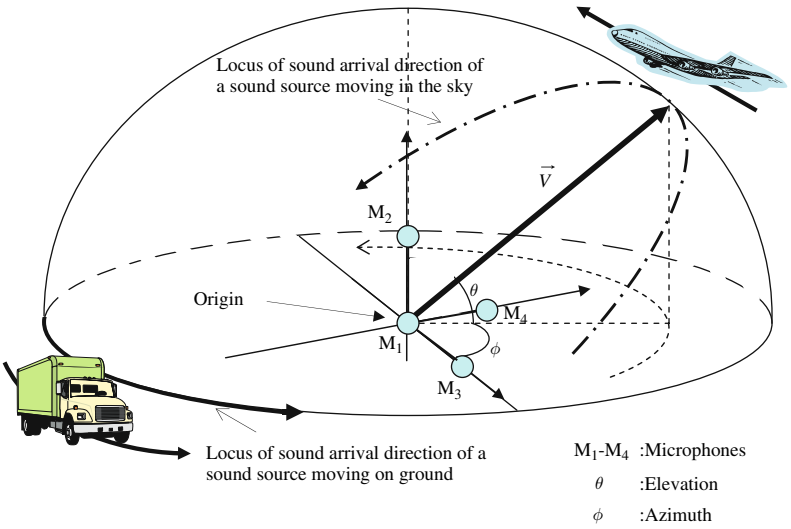


FIGURE 30 Schematic diagram to estimate three-dimensional sound arrival direction.

provided that microphone spacing is the same among the three microphone pairs:

$$\tan\phi = \frac{\tau)y}{\tau_x}, \qquad \tan\theta = \frac{\tau_z}{\sqrt{\tau_x^2 + \tau_y^2}}. \tag{22}$$

Figure 31 shows such a microphone array system actually installed around an airport. Figure 32 illustrates loci of instantaneous sound arrival directions that were repeatedly estimated for five sound events; (a) and (b) show the sound arrival direction gradually moves along a locus on the upper hemisphere of the unit sphere for approaching aircraft, (c) it stays near a fixed point for a fixed sound source, (d) it passes along the horizon for road traffic, and (e) it scatters widely on the unit sphere due to wind noise. Figure 33 shows another illustration, in which sound arrival direction was observed near an engine run-up facility at an airport [49]. The sound arrival direction, expressed by arrows indicating azimuth and a line graph of elevation, points to a fixed direction during the engine run-up operation, while it is gradually turned around following the sound source during aircraft fly-over.

Information on a locus of sound arrival directions, i.e., a series of unit direction vectors for a sound event can be compressed to a single vector if the sound source moves on a straight flight path, as is shown in Figure 34. At that time, the unit direction

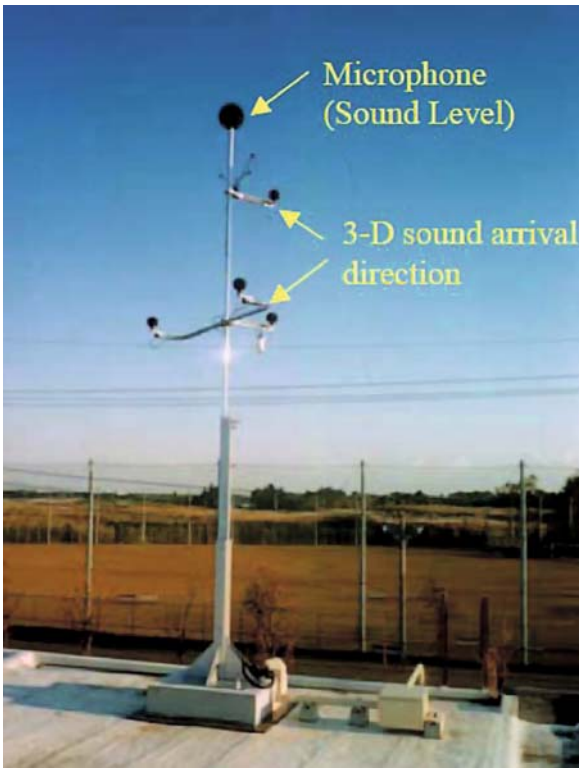


FIGURE 31 Microphone array to estimate three-dimensional sound arrival direction.

vector follows a great-circle route. It can be transformed to a unit normal vector $\vec{N}$ by taking a vector product of two unit direction vectors $\vec{V}_1$ and $\vec{V}_2$ on the route, i.e., $\vec{N} = \vec{V}_1 \times \vec{V}_2$. Practically, we should take an average of unit normal vectors calculated repeatedly during the event, because individual calculations may be contaminated with errors due to background noise or due to discrepancy with the assumption of straight paths. The resultant magnitude of the average normal vector gives us a measure of how fine the great-circle route assumption holds. For example, if the sound is wind noise, both $\vec{V}$ and $\vec{N}$ scatter throughout on the unit sphere, resulting in a small average normal vector. Figure 35 shows an illustration of flight route classification, based on average normal vectors of sound events observed during 1 month at a site 20 km distant from a runway [50]. The figure shows a contour map of frequency distribution of average normal vectors together with target windows of flight route

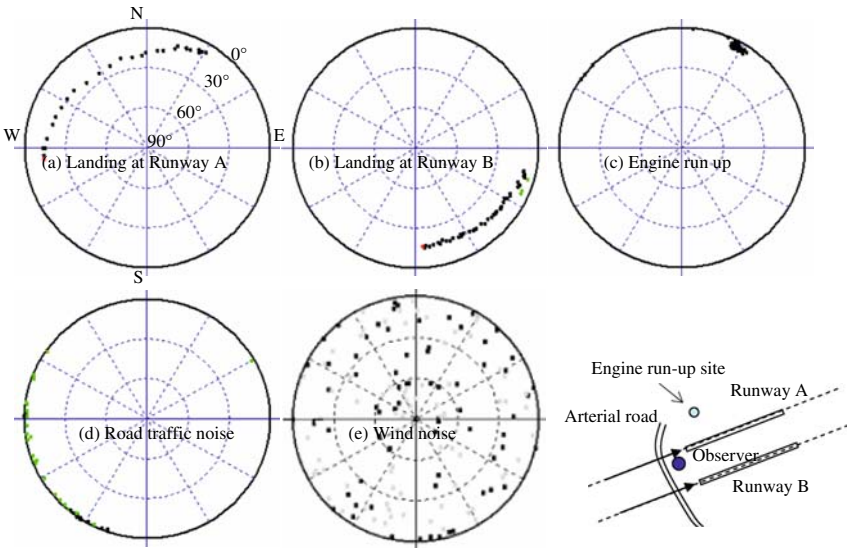


FIGURE 32 Change of three-dimensional sound arrival direction expressed in azimuth and elevation. Elevation is expressed in *dotted concentric circles* at every 30°.

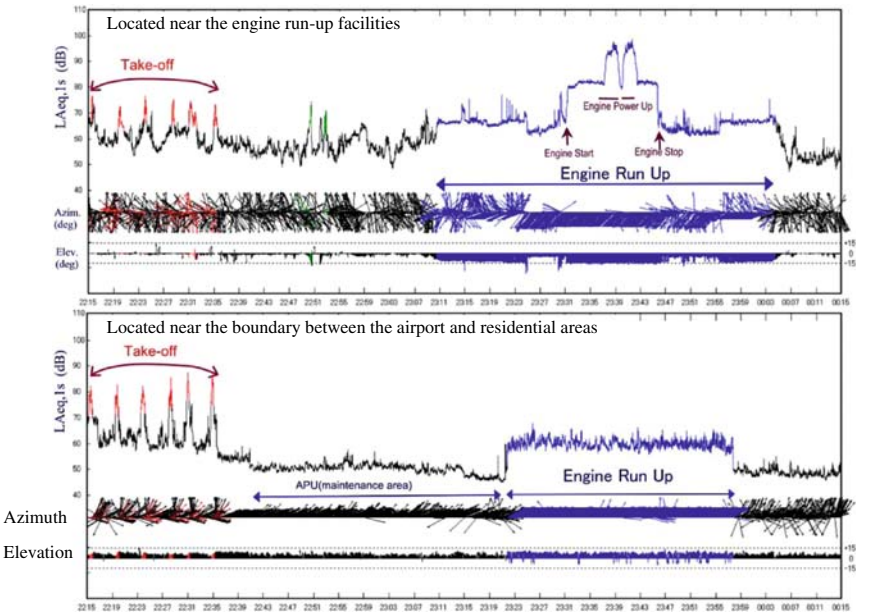


FIGURE 33 Time histories of short-term average A-weighted sound-pressure levels (*L*Aeq,1s) with instantaneous sound arrival direction expressed in azimuth and elevation.

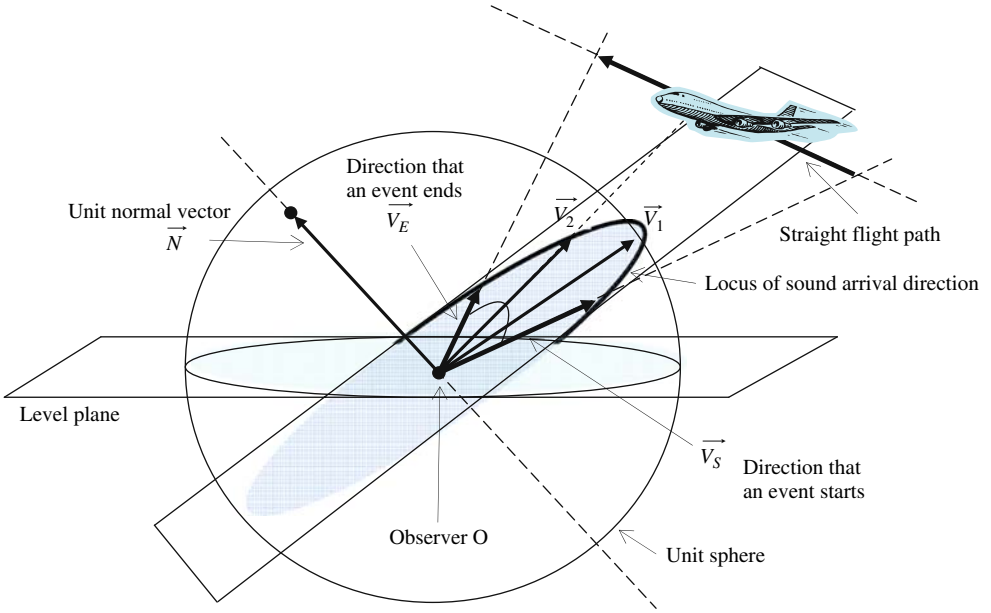


FIGURE 34 Relationship of a locus (a great-circle route) of three-dimensional sound arrival direction (unit direction vectors) with a normal vector.

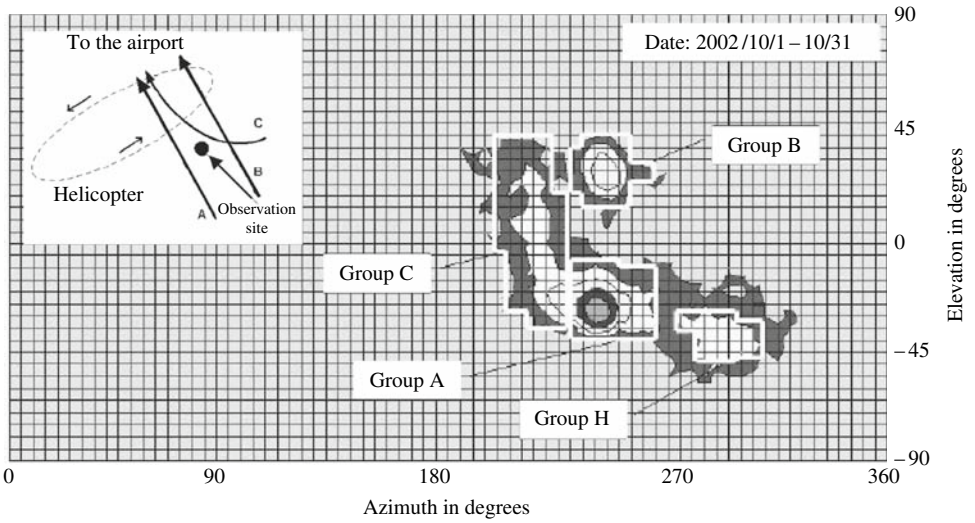


FIGURE 35 Contour map of frequency distribution of average normal vectors of aircraft sound events measured by unmanned monitoring with target windows of flight route categories.

categories A, B, C (large jet aircraft) and H (helicopters), which were empirically determined for individual categories from independent 2-day measurements.

Finally, similar techniques to the cross-correlation method are also reported: time-delay estimation by cross-correlation of beam-formed outputs of two arrayed receivers for the acoustic detection of jet aircraft transits [51], sound source localization based on cross-correlation of a four-microphone array outputs with non-linear least square and/or extended Kalman filter optimization procedure [52], and sound source identification using a spherically arranged microphone array [53,54].

References

1. Y. Miki, "Acoustical properties of porous materials – modifications of Delany-Bazley models," J. Acoust. Soc. Jpn (E), 11, 19–24 (1990).
2. M.E. Delany and E. N. Bazley, "Acoustical properties of fibrous absorbent materials," Appl. Acoust., 3, 105–116 (1970).
3. C.I. Chessell, "Propagation of noise along a finite impedance boundary," J. Acoust. Soc. Am., 62, 825–834 (1977).
4. S.I. Thomasson, "A powerful asymptotic solution for sound propagation above an impedance boundary," Acustica, 45, 122–125 (1980).
5. G.A. Daigle, "Technical assessment of the effectiveness of noise walls," Noise/News Int., 7(3), 137–161 (1999 September).
6. Z. Maekawa, "Noise reduction by screens," Appl. Acoust., 1, 157–173 (1968).
7. N. Aoshima, "Computer-generated pulse signal applied for sound measurement," J. Acoust. Soc. Am., 69(5), 1484–1488 (1981).
8. Y. Suzuki, F. Asano, H. Kim and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," J. Acoust. Soc. Am., 69(5), 1119–1123 (1995).
9. K. Kimura and K. Yamamoto, "A method for measuring oblique incidence absorption coefficient of absorptive panels by stretched pulse technique," Appl. Acoust., 62, 617–632 (2001).
10. W.D.T. Davies, System Identification for Self-Adaptive Control, Chap. 3, Wiley-Interscience, London (1970).
11. N. Aoshima and J. Igarashi, "The M-sequence signal applied to the correlation measurement of the wave propagation," Reports of the 6th International Congress on Acoustics, F-1-5, 17–20 (1968).
12. M. Garai, "Measurement of the sound-absorption coefficient in situ: the reflection method using periodic pseudo-random sequences of maximum length," Appl. Acoust., 39, 119–139 (1993).
13. M. Garai, "The new European standard for qualifying added devices," CD-ROM of 18th ICA Kyoto (2004).
14. ISO 10847:1997, "Acoustics – *in-situ* determination of insertion loss of out door noise barriers of all types."

15. N. Aoshima and J. Igarashi, "M-sequence signal applied to the correlation measurement of acoustical system," J. Acoust. Soc. Jpn, 24(4), 197–206 (1968).
16. N. Aoshima and J. Igarashi, "Field measurement of noise propagation by M-sequence correlation method," Inst. Space and Aeronautical Sci., Research Note, No. 9 (1975).
17. Y. Ishii, J. Igarashi, I. Yamada and Y. Miki, "Noise propagation measurement using techniques of signal processing," Inst. Space and Aeronautical Science Report 14(1)C, 619–642 (1978).
18. K. Konishi, Y. Tanioku and Z. Maekawa, "Long time measurement of long range sound propagation over an ocean surface," Appl. Acoust., 61, 149–172 (2000).
19. Y. Ishii, I. Yamada, Y. Miki and J. Igarashi, "Noise propagation measurement by synchronized integration technique," J. Acoust. Soc. Jpn, 35(5), 226–234 (1979).
20. M. Arai and T. Chino, "Highly directional microphone system for wide-band measurement (in Japanese)," Proc. Acoustical Soc. Japan '78–5, 502 (1978).
21. Y. Moritoh, Y. Zenda and K. Nagakura, "Noise control of high speed Shinkansen," J. Sound Vibr., 193(1), 319 (1996) (Reprinted with permission from Elsevier).
22. Y. Oshino and H. Tachibana, "Noise source identification on rolling tires by sound intensity measurement," J. Acoust. Soc. Jpn (E), 12(2) 87, (1991) (Reprinted with permission from the Acoustical Society of Japan).
23. Y. Matui, R. Kaneko, M. Murase and K. Mizuno, "Research on visualization of noise distribution between the Shinkansen and a noise barrier," Proc. Inter-Noise 94, 135 (1994).
24. M. Kawahara, H. Hotta, M. Hiroe and J. Kaku, "Source identification and prediction of Shinkansen noise by sound intensity," Proc. Inter-Noise '97, 72 (1997).
25. Y. Hidaka, H. Tachibana, Y. Matsui and R. Kaneko, "Measurement of sound radiation from Shinkansen train by sound intensity method," Proc. Inter-Noise '95, 215 (1995).
26. J.S. Joo, Y. Hidaka and H. Tachibana, "A study on the sound radiation characteristics of conventional railway by sound intensity method (in Japanese)," J. Acoust. Soc. Jpn, 56(1), 32 (2000) (Reprinted with permission from the Acoustical Society of Japan).
27. B. Barsikow and B. Mueller, "A detailed study of the vertical distribution of noise sources generated by a high-speed train," Proc. Inter-Noise '85, 299 (1985).
28. B. Barsikow, W. F. King III and E. Pfizenmaier, "Wheel/rail noise generated by a high-speed train investigated with a line array of microphones," J. Sound Vibr., 118(1), 337 (1987).
29. B. Barsikow and W. F. King III, "On removing the doppler frequency shift from array measurements of railway noise," J. Sound Vibr. 120(1), 190 (1988).
30. K.P. Schmitz, "Sound measurements on TRANSRAPID 07 – methods and results," Proc. STECH '93, 2, 33 (1993).

31. A. Torii, Y. Takano, K. Terada, M. Sebata, A. Iida, K. Horihata and H. Fujita, "Shinkansen's sound source measurements using microphone arrays," Proc. Inter-Noise '92, 1171 (1992).
32. Y. Takano, K. Terada, E. Aizawa, A. Iida and H. Fujita, "Development of 2-dimensional microphone array measurement system for noise sources of fast moving vehicles," Proc. Inter-Noise '92, 1175 (1992).
33. Y. Takano, "X-shaped two-dimensional microphone array system for measuring noise-source distribution on moving vehicles," JSME Int. J. Ser. C, 41(1), 46 (1998).
34. A. Nordborg, A. Martens, J. Wedemann and L. Willenbrink, "Wheel/rail noise separation with microphone array measurements," Proc. Inter-Noise 2001, 2083 (2001).
35. J. Hald and J. J. Christensen, "A class of optimal broadband phased array geometries designed for easy construction," Proc. Inter-Noise 2002, N447 (2002).
36. Y. Takano, K. Sasaki, T. Satoh, K. Murata, H. Mae and J. Goto, "Development of noise source visualization system for high-speed vehicles (in Japanese)," Proc. 2003 Symposium on Environmental Engineering, 65 (2003) (Reprinted with permission from the Japan Society of Mechanical Engineers).
37. Y. Takano, K. Sasaki, T. Satoh, K. Murata, H. Mae and J. Goto, "Development of visualization system for high-speed noise sources with a microphone array and a visual sensor," Proc. Inter-Noise 2003, N930 (2003).
38. T. Okuda, H. Tsuru, N. Hayashi, T. Okabe and S. Ohashi, "Noise monitoring by local self-governing bodies around new Tokyo International Airport", Proc. Inter-Noise 96, 1091–1934 (1996).
39. R. C. Muchall, "Experiences with an intelligent measurement system for monitoring aircraft noise without radar assistance," Proc. Inter-Noise 01 (2001).
40. I. Yamada et al., "Acoustic recognition of aircraft types in flight," Internoise 83, 827–830 (1983).
41. K. M. Adams, "Success and failure analysis of neural network identification of aircraft noise," Proc. Inter-Noise 2001.
42. G. Nishinomiya et al., "Aircraft noise identification system by correlation technique," IEEE BC-24, No. 24 (1978).
43. T. Ono et al., "Aircraft noise monitoring system with identification by correlation technique," Internoise 79, 721–724 (1979).
44. I. Yamada et al., "Improvement of the performance of cross correlation method for identifying aircraft noise with pre-whitening of signals," J. Acoust. Soc. Jpn (E), 13(4), 241–252 (1992).
45. H. Kobatake et al., "Acoustical identification of aircraft position," Report of the Institute of Space and Aeronautical Science, the University of Tokyo, 9(4), 854–869 (1973).
46. I. Yamada et al., "Unattended noise monitoring and source identification of aircraft noise," Proc. Inter-Noise 89, 1133–1136 (1989).

47. I. Yamada et al., "Identification of outdoor sound sources with the aid of automatic tracking of sound arrival direction," Proc. Inter-Noise 92, 903–906 (1992).
48. I. Yamada et al., "Recent works on acoustical source identification for unattended monitoring of aircraft noise," Proc. 15th ICA, 339–342 (1995).
49. T. Ohshima and I. Yamada, "Source classification of environmental noise using short-term L_{Aeq} sound arrival direction," Proc. Inter-Noise 1999.
50. K. Fukushima, "Discrimination of aircraft noise using sound arrival direction for unattended noise monitoring system," Proc. Inter-Noise 2003, N814 (2003).
51. B.G. Ferguson, "Time-delay estimation techniques applied to the acoustic detection of jet aircraft transits," J. Acoust. Soc. Am., 106, 255–264 (1999).
52. N. Roosnek, "Passive ranging with four microphones in a spatial configuration," Proc. Inter-Noise 2001, 2159–2162 (2001).
53. H. Tsuru, O. Kohashi, M. Sakai, T. Sugaya and S. Sugaya, "Monitoring of flight path by aircraft noise," Proc. Acoust. Soc. Jpn, 653–654 (1997).
54. N. Tanaka, Y. Tadahira, K. Yamashita and S. Ohashi, "A study on measurement technique of the noise with spherical microphone array," Proc. Acoust. Soc. Jpn, 861–864 (September 2004).

53

Measurements of Physical Characteristics Using Sound at Audible or Lower Frequencies

Ippei Torigoe

Department of Mechanical Engineering Kumamoto University, Kumamoto, Japan

Sound is a mechanical disturbance from a state of equilibrium that propagates through an elastic material medium [1]. The particle vibrations are accompanied by fluctuations in medium properties, and sound therefore carries information about the medium along the path of travel of the acoustic wave. Ultrasound measurement is a substantial area of research and industry [2,3], however, measurement using audible or even lower frequencies has had significantly less attention, although various measurements using lower frequencies have been reported. Examples include: photoacoustic spectroscopy [4,5], fault diagnosis [6–8], measurement of flow rate or flow velocity [9–20], density of fluid [21–24], viscosity of fluid [25–27], pressure [28,29], level [30–32], angular velocity [33–35], temperature [36,37], thermal diffusivity [38], length and gap [39–42], capacity or volume [43–52], surface area [53–56], Young's modulus [57], and sugar content [58]. In this chapter, we will show several examples

of such measurements using audible or lower frequency sound: measurement of gap (Sect. 1), volume (Sect. 2), surface area (Sect. 3), angular velocity (Sect. 4), flow velocity (Sect. 5), differential pressure (Sect. 6), and kinematic viscosity (Sect. 7). Most of these methods boil down to the question how to measure the acoustic impedance accurately, eliminating the effect of other influences. Since any change in the acoustic impedance due to the change in the target quantity is a small fraction, as is typically the case with the measurement of surface area, in order to eliminate noises it is necessary to apply signal processing on every level; signal processing with analog electric circuits, digital signal processing, and noise elimination with the structure and configuration of the measuring system itself. For example, a differential structure is employed in the measurements of acoustic micrometer, volume meter, surface area meter and acoustic rate gyro, comparative measurement in acoustic volume meter and surface area meter, a bridge configuration in acoustic surface area meter and rate gyro, a use of information contained in the phase of sound pressure in acoustic rate gyro and kinematic viscosity sensor, modulation in chopper microphone, lock-in detection in acoustic volume meter, surface area meter, impedance flow meter, chopper microphone, and kinematic viscosity sensor, and so on.

1 • Acoustic Micrometer [59]

Figure 1 shows a schematic diagram of the acoustic micrometer. A short sensing tube is connected to the front side of a speaker box, with the opposite, flanged end placed close to the surface of an object, leaving approximately a 100 μm gap. The backside of the speaker box is confined but connected to the atmosphere through an equalizing capillary vent. The pressure in the two sections of the speaker box is measured by electret microphones. The output signals of the microphones are band-pass filtered and the phase difference between them is measured.

The acoustic impedance Z_a of the annular gap, which is formed by the flange of the tube and the object surface, is approximately given by

$$Z_a = \left\{ \frac{6\mu}{\pi d^3} + i\frac{3\rho\omega}{5\pi d} \right\} \cdot \lg\left(\frac{b}{a}\right) \tag{1}$$

where d is the gap distance, μ is the coefficient of viscosity of air, ρ is the density of air, ω is the angular frequency of sound, a and b are the inner and the outer radius of the flange [38]. For small

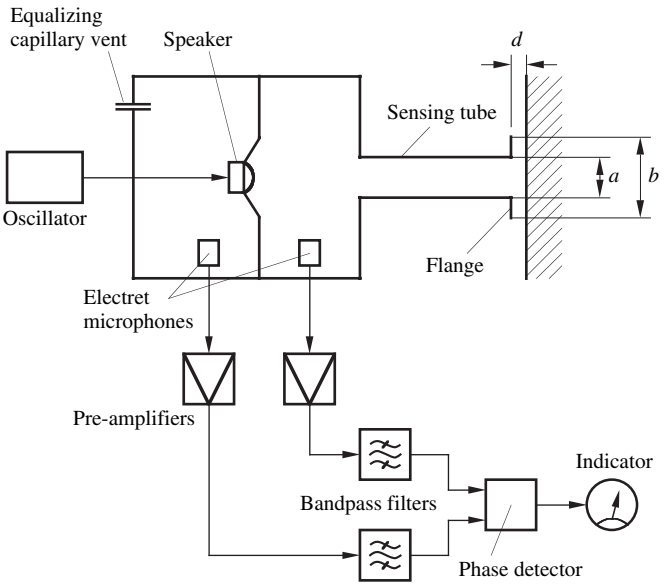


FIGURE 1 Schematic diagram of the acoustic micrometer.

gaps, the first term is dominant and the acoustic impedance Z_a changes in inverse proportion to the cube of the gap d. The amplitude and phase of the sound pressure within the front cavity is therefore very sensitive to changes in the gap d.

The gap can be determined from the amplitude of the sound pressure in the front cavity. However, the amplitude of the detected sound pressure is proportional to the sensitivity of microphone, which in turn, is susceptible to any change in operating environment, for example ambient temperature. On the other hand, it is a much simpler process to control the phase characteristics of microphones in the range from 100 to 1000 Hz. In the acoustic micrometer depicted in Figure 1, the sound pressure is also picked up in the back cavity of the speaker, and the difference in the phase between the sound pressures of the front and the back cavities is detected as a measure of the gap d. The effect of the change in the amplitude of speaker vibration is also eliminated in this configuration.

Figure 2 shows the experimental setup. The acoustic micrometer is held in a C-shaped brass clamp with the sensing flange facing the spindle face of a mechanical micrometer head. The gap d can be changed by feeding the spindle forward or backward. Figure 3 shows the phase difference recorded against time while changing the gap d. In this experiment, the sound frequency is 125 Hz, a and b are 4 and 6 mm, and the sound

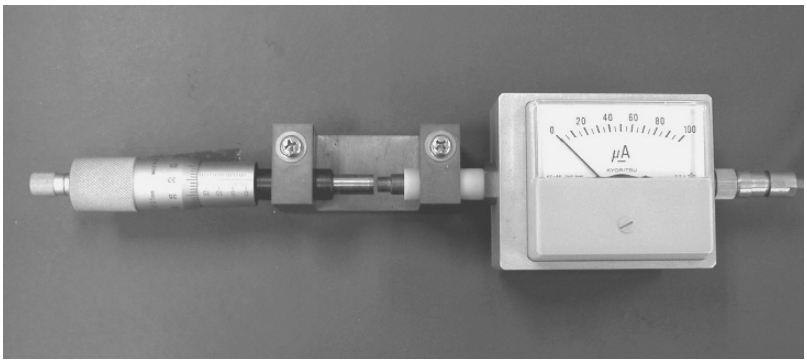


FIGURE 2 Experimental setup for the gap measurement.

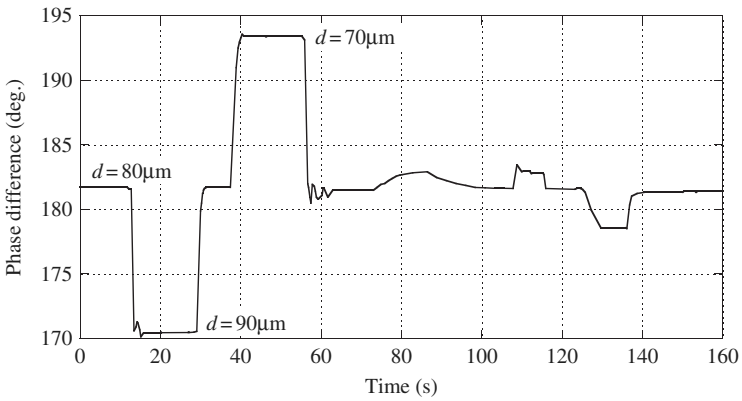


FIGURE 3 A plot of measured gap.

pressure level within the cavities is about 120 dB. A gradual increase in the phase difference around $t = 80$ s was due to the decrease of the gap, which was caused by the thermal expansion of the sensing tube brought about by touching it with fingers. A rapid increase at about $t = 110$ s was caused by pinching the clamp with fingers. At $t = 130$ s the whole experimental setup was rotated to the position in which the sensing tube was vertical, a slight deformation in the clamp was caused by the weight of the acoustic manometer. The gap was measured to a resolution of 0.1 μm with this apparatus.

2 • Acoustic Volume Meter [45,51]

For sufficiently low frequencies, the acoustic impedance of a container is capacitive; its amplitude is nearly independent of shape and in inverse proportion with the capacity of the

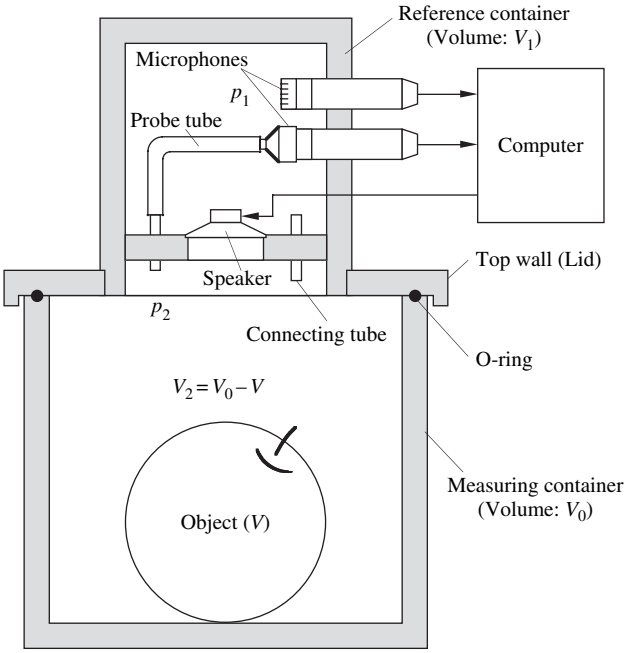


FIGURE 4 Configuration of the acoustic volume meter.

container. The capacity of a container, or the volume of an object which is placed in a known container, can thus be determined from the acoustic impedance measurement.

Figure 4 shows the configuration of the acoustic volume meter. A reference container with a known volume (V_1) is installed on the top wall of a measuring container whose empty volume (V_0) is known. A speaker is also mounted in this top wall. Two condenser microphones are fixed to the reference container; one of them picks up the pressure fluctuation in the reference container and another picks up that in the measuring container led through a probe tube. The two containers are connected via a tube which acts to equalize the pressure and the composition of both containers. The top wall, together with the reference container, the speaker and the microphones, is lifted when an object to be measured is placed in the measuring container.

The speaker is driven sinusoidally at a sufficiently low frequency (usually 39 Hz) to generate sinusoidal fluctuations in the volume of each container which are equal in magnitude while opposite in phase. Although air moves through the connecting tube, the length of the tube is short compared with the wavelength of the driving frequency and the air within the tube can be considered to move in phase. Therefore, the volume change in each of the two containers caused by air movement has the

same magnitude and opposite phase. So, at a given moment, the volume change in each container is equal in magnitude but opposite in sign.

When a change of the volume $+\Delta V$ is applied to the measuring container, the resulting sound pressure p_1 and $-p_2$ within each container can be given as follows, under the condition that the change in state of gas is adiabatic:

$$\begin{aligned} \frac{p_1}{p_s} &= \gamma \frac{\Delta' V}{V_1} \\ \frac{p_2}{p_s} &= \gamma \frac{\Delta V}{V_2} = \gamma \frac{\Delta V}{V_0 - V}, \end{aligned} \tag{2}$$

where γ is the ratio of specific heats, p_s is the average static pressure of the air in the volume, V is the volume of the object, and $V_2 = V_0 - V$ is the remnant volume in the measuring container. Using these equations, the volume of the object placed in the measuring container is calculated from the ratio of the magnitudes of sound pressure within the two containers:

$$V = V_0 - V_1 \frac{p_1}{p_2}. \tag{3}$$

The outputs of the condenser microphones e_1 and e_2 are converted into digital signals, and their Fourier component magnitudes E_1 and E_2 at the speaker driving frequency are calculated. The ratio of the sound pressure magnitudes is estimated at E_1/E_2. Employing this configuration and a regular calibration procedure, this volume meter eliminates the errors which would be caused by the change in the gas composition (humidity), the atmospheric pressure, the amplitude of the speaker vibration, and the sensitivity of the microphones.

The acoustic volume meter has been applied to the measurement of the volume of weights (Figure 5) [60]. In order to satisfy the uncertainty required for the class E_1 weights of OIML [61], the buoyancy correction of air for the volume of a weight is essential. Measurement of the volume of a weight must have a relative uncertainty of less than 3×10^{-3}. Although the conventional hydrostatic weighing method is capable of measuring the volume itself with a relative uncertainty of the order of 10^{-6}, it requires a complicated procedure involving wetting in a reference liquid and then drying the weight. In addition, there are problems associated with moisture content and contamination after removing the weight from the liquid. In practical applications, it has been proposed that acoustic volume meters be used for the measurement of the volume of



FIGURE 5 Acoustic volume meter applied for the measurement of weight volume.

weights [62]. The acoustic volume meter is suitable for containers and objects to which the hydrostatic method is not applicable. One successful example is the measurement of the volume of combustion chamber of reciprocating engines (Figure 6) [60]. Applications of the acoustic volume meter to the measurement of various quantities such as the sugar content of fruit [58], the density of ice in a glacier, fuel level in fuel tanks [32] have also been reported.

3 • Acoustic Surface Area Meter [56]

The heat flow between an object and the surrounding medium is proportional to the surface area of the object. This heat conservation law can be applied to the measurement of the surface area of an object. If the wall of a cavity is precisely adiabatic, the acoustic impedance of the cavity is capacitive. For actual cavities, however, the wall is nearly isothermal since the wall, in comparison with air, is highly conductive and has a far larger heat capacity. Heat conduction occurs between the air and the isothermal wall, and a real component arises in the acoustic impedance of a cavity, which has a magnitude that is some function of the cavity wall area. Therefore, the surface area of a cavity, or of an object placed in a known chamber, can be obtained from the measurement of the acoustic impedance. However, the real component of the acoustic impedance of a chamber is a small fraction of the whole content. It is not

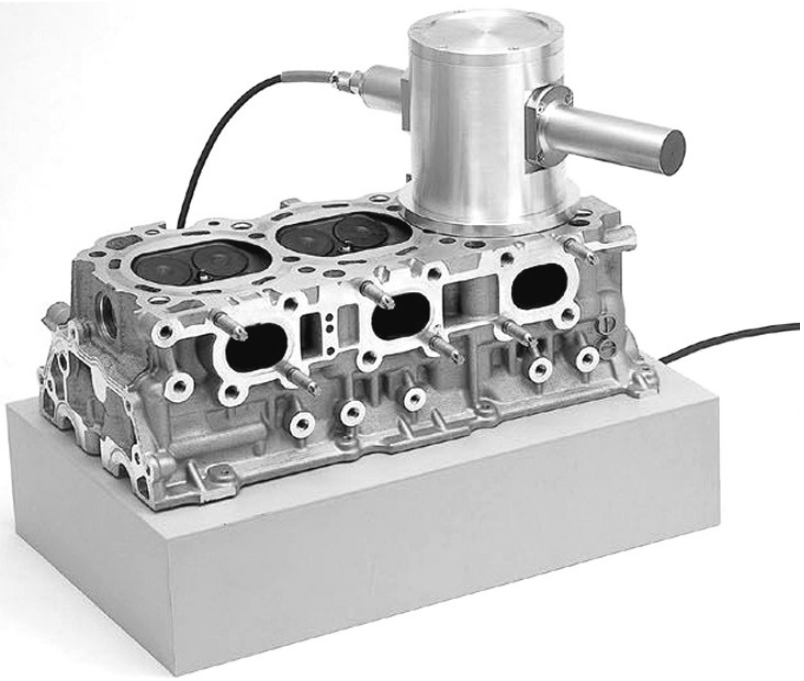


FIGURE 6 Acoustic volume meter applied for the measurement of combustion chamber capacity.

easy to measure this component accurately, while eliminating all additional effects. Therefore, in order to attain this goal, a bridge technique is employed.

Figures 7 and 8 show the acoustic surface area meter. The device consists of two 80 mm × 80 mm × 70 mm chambers; a measuring chamber in which an object is placed and a reference chamber. On the top of each chamber, bellows and a micrometer head are installed. The volume of the measuring chamber can be adjusted by feeding the micrometer head forward and backward, whereas the inner surface area remains constant. The reference chamber is filled with silicone oil (dimethylpolysiloxane) to the level of the bottom of the measuring chamber. On the side wall of the reference chamber a rotary stage is attached, and to this stage a rotor shaped like the rotating electrode of a variable capacitor (Figure 7b) is fixed concentrically. The inner surface area of the reference chamber can be changed by rotating the rotor forward and backward without the volume being varied. Between the two chambers, a speaker is mounted and driven sinusoidally at a frequency of 40 Hz to create sinusoidal volume perturbations

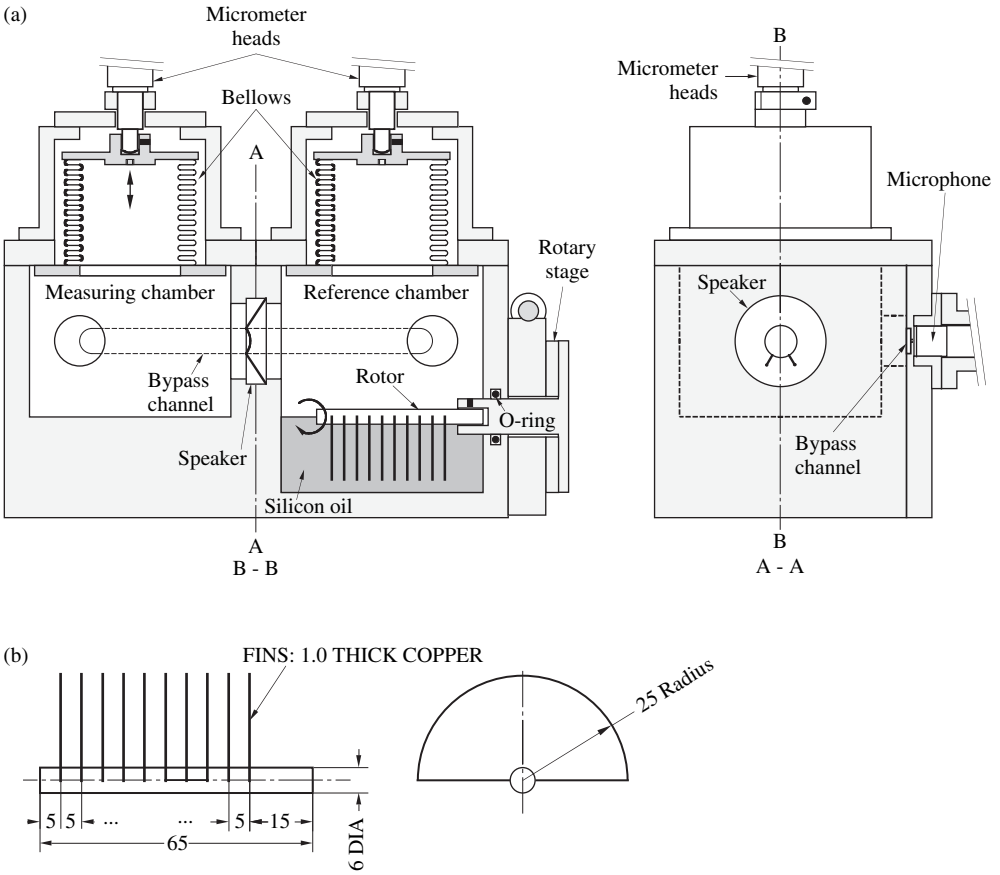


FIGURE 7 (**a**) Configuration of the acoustic surface area meter. (**b**) The rotor (units:mm).

in the two chambers that are equal in magnitude but opposite in sign. A rectangular groove is cut on the backside wall to provide a bypass channel which connects the two chambers. At the midpoint of this channel a condenser microphone is installed to pickup the sound pressure there. Using a lock-in amplifier, the output of the microphone is converted to two d.c. voltages.

The acoustic impedance Z_I of a cavity of an isothermal wall is given by

$$Z_I = \frac{\gamma p_s}{i\omega V_0}\left\{1+(1-i)(\gamma-1)\frac{S_0\delta}{2V_0}\right\}^{-1} \approx \frac{\gamma p_s}{\omega V_0}(\gamma-1)\frac{S_0\delta}{2V_0}+\frac{\gamma p_s}{i\omega V_0}\left\{1-(\gamma-1)\frac{S_0\delta}{2V_0}\right\}, \quad (4)$$

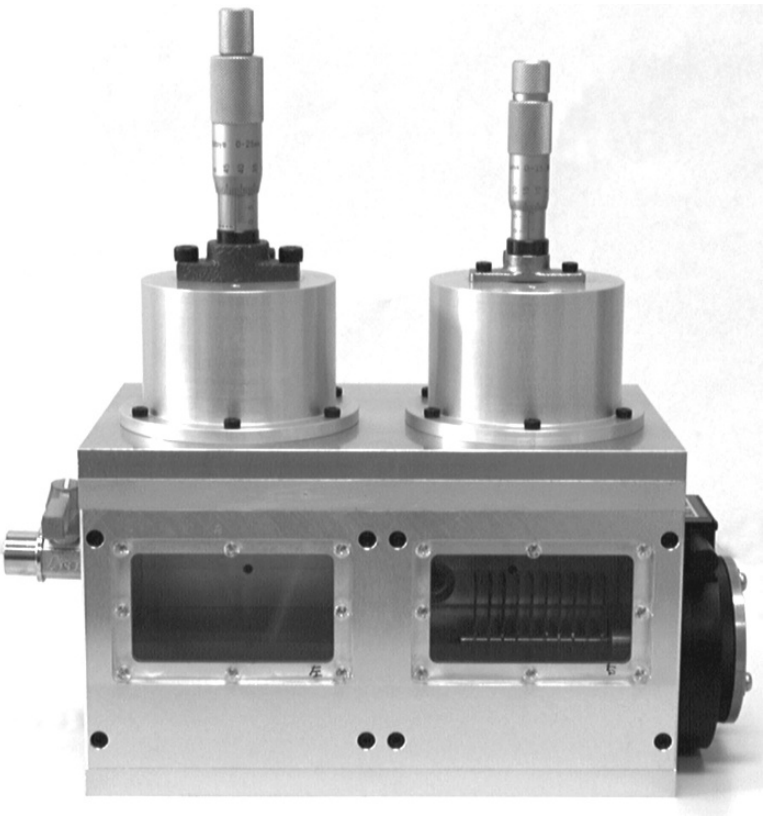


FIGURE 8 Photograph of the experimental device.

where γ is the ratio of specific heats, V_0 is the volume of the cavity, p_s is the atmospheric pressure, ω is the angular frequency, S_0 is the inner surface area of the cavity, and δ is the thickness of the thermal boundary layer. Using the thermal conductivity κ, the density ρ, and the specific heat at constant pressure c_p of air, δ is given by

$$\delta = \sqrt{\frac{2\kappa}{\rho \omega c_p}}. \tag{5}$$

When compared with the acoustic impedance of an adiabatic cavity

$$Z_A = \frac{\gamma p_s}{i\omega V_0}, \tag{6}$$

we can see that there is a real component in the acoustic impedance and additional terms in the imaginary component.

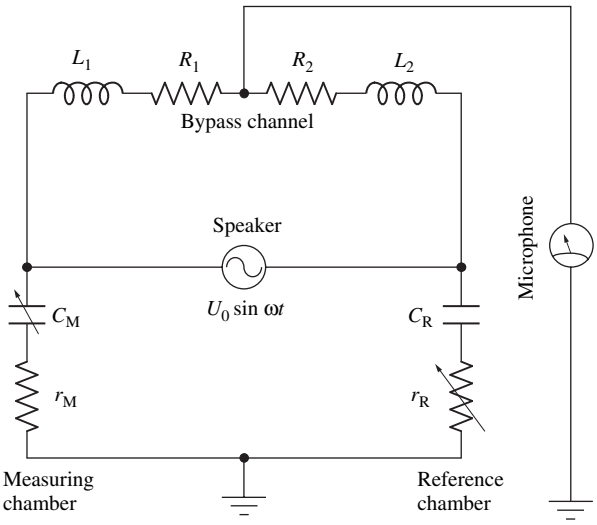


FIGURE 9 Electrical equivalent for the acoustic surface area meter.

Figure 9 shows the electrical equivalent for the acoustic surface area meter. The measuring chamber is represented as a variable capacitor with a small serial resistance, and the reference chamber is represented by a capacitor and a variable resistance. The speaker is represented as a current source whose intensity is proportional to the volume velocity generated by it. The electrical equivalent for the bypass channel is a series of inductors and resistors. Since the microphone, the null detector, is installed at the midpoint of the bypass channel, the balance equation for this bridge circuit is

$$r_R = r_M, \quad C_R = C_M \tag{7}$$

which means that the acoustic bridge is balanced when the volume and the inner surface area of each chamber become equal. The surface area of an object can be determined from the angular deflection of the rotor between the balance points before and after the object is placed.

Figure 10 shows the result of the experiment measuring surface areas of objects of several shapes. Copper discs, steel balls, brass cylinders, and aluminum foils were measured. The abscissa is the surface area which is calculated from the object dimensions and, for objects containing edges, corrected considering the edge effect [51]. The ordinate represents the measured surface area into which the angular deflection reading is converted; the scale was determined from a single-point calibration using a standard object (disc).

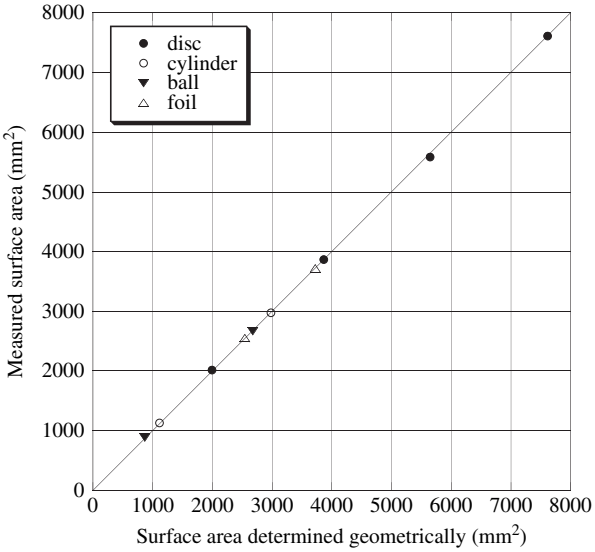


FIGURE 10 Measured surface area of objects of various shapes ($\omega/2\pi = 40\,\text{Hz}$).

In the determination of the surface area on the basis of Eq. (4), the error in each of the variables is propagated to that in the measured surface area. More basic quantities such as temperature also affect the measurement of these variables. In order to eliminate these influences the bridge technique has been employed in the experimental device. If a perfect symmetry of the acoustic bridge is attained, all the influences will cancel and the surface area of an object can be determined only from the inner surface area of the reference chamber. In addition, there are uncertainty contributions from the medium compressibility, the calibration error, the error due to the curvature of object surface [50], and random error. Uncertainty is also caused by unpredictable fluctuations of the measurement conditions: noises in the electronic circuit, the ambient pressure fluctuation, mechanical vibrations, the temperature fluctuation in the measuring chamber introduced by the fluctuation of room air temperature, the fluctuation of gas content (the humidity in particular) in the measuring chamber, etc. The measurement uncertainty due to random effects is estimated from the experimental standard deviation. The combined overall standard uncertainty [63] in the measurement of the experiment of Figure 10, apart from the influence of the curved surface, ranged from $74\,\text{mm}^2$ (8.5 %) to $148\,\text{mm}^2$ (2 %).

4 • Acoustic Rate Gyro [35]

Figure 11 shows the configuration of the acoustic rate gyro. The device is partitioned into two containers (volumes V_1 and V_2) by a T-shaped inner wall. A sensing channel of a low-profile rectangular cross-section and a length of l is formed between the top wall of the device and the T-shaped wall. This channel connects the two containers. A speaker is installed in the inner wall and driven sinusoidally at an angular frequency ω to yield a sinusoidal vibration of the air within the sensing channel in the x-direction. Two small taps are drilled a distance d apart along the y-axis in the top wall, and two electret microphones are installed to detect the sound pressure in the sensing channel.

The electrical equivalent for this device is also a bridge circuit. In the acoustic rate gyro, however, the volumes V_1 and V_2 are adjusted using the adjust plug so that the bridge is slightly off balance at the midsection of the sensing channel and small sinusoidal signals are picked up by the microphones. In this condition, if the device is rotated about the z-axis at an angular

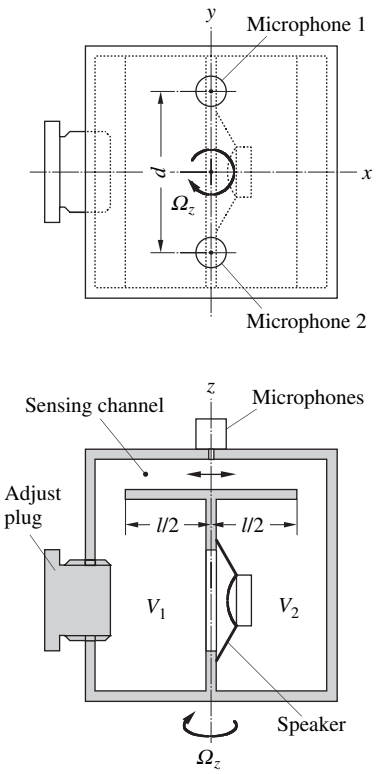


FIGURE 11 Configuration of the acoustic rate gyro.

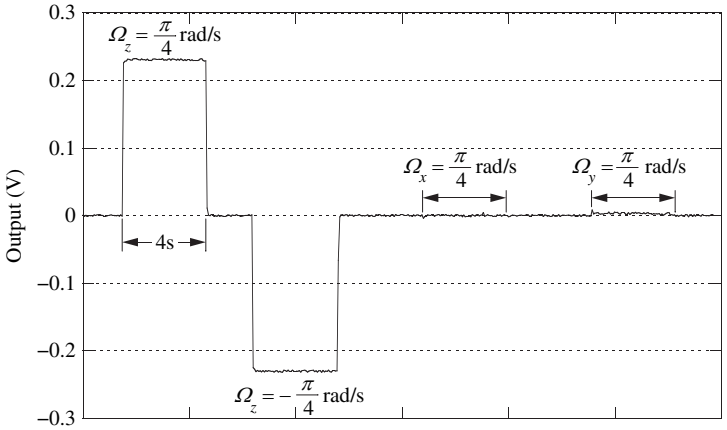


FIGURE 12 A plot of measured angular velocity.

velocity Ω_z, a phase difference

$$2\theta = 2\tan^{-1}\left(2\frac{V_1+V_2}{V_1-V_2}\cdot\frac{d\Omega_z}{\omega l}\right) \tag{8}$$

arises between the outputs of the two microphones because of Coriolis force exerted on the vibrating air in the sensing channel.

Figure 12 shows the experimental results. The phase difference between the two microphone signals was converted into a voltage by a phase detector and plotted against time (the frequency $\omega/2\pi = 213\,\text{Hz}$, $V_1 = 40\,\text{ml}$, $V_1 - V_2 = 0.9\,\text{ml}$, $l = 30\,\text{mm}$, $d = 39\,\text{mm}$). The device detects rotation about the z-axis but is insensitive to that about the x-axis and y-axis.

5 • Acoustic Impedance Flowmeter [18]

At low Strouhal numbers, the acoustic resistance of an aperture increases in proportion to the mass velocity of flow which grazes the aperture. This is, intuitively speaking, because the acoustical energy stored in the vicinity of the aperture is blown away by the flow. Interaction between flow and sound can cause a deterioration of the performance of acoustic devices such as silencers. However, we can exploit this phenomenon for measuring the flow rate.

Figure 13 shows the experimental setup of the acoustic impedance flowmeter. Two apertures are drilled in the wall of a probe shaped like a Pitot-static tube. This probe is placed in such a position that the apertures are exposed to grazing flow from a wind tunnel, and is driven by a speaker at a constant velocity

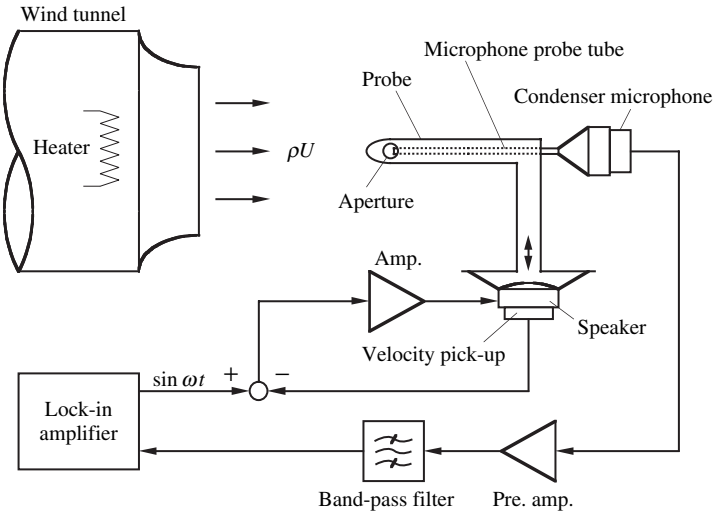


FIGURE 13 Schematic diagram of the acoustic impedance flowmeter.

amplitude. Since the wavelength at the driving frequency (10 Hz in this experiment) is much larger than the probe dimension, we can neglect the phase shift between oscillating velocities at any two points in the probe. The sound pressure within the probe is picked up by a condenser microphone at the aperture position. The lock-in amplifier detects the in-phase component in the output of the microphone, which is proportional to the acoustic resistance of the aperture.

Figure 14 shows measured mass flow velocities plotted against time. During the measurement the fan was switched between high and low in order to change the flow velocity, and the heater was switched on or off to change the density of air. The values presented in the graph are the mass flow rate independently calculated from the velocity, temperature, and the pressure of air.

6 • Chopper Microphone [29]

Figure 15 shows a schematic diagram of the chopper microphone – an ultra sensitive differential pressure manometer. The device consists of a chopper (switching valve), a microphone, two buffers, and a synchronous detector. The chopper switches two input pressures and introduces them alternately to the microphone. The pressure applied to the microphone is switched between p_1 and p_2 alternately (the buffers improve the transient response of the switched pressure). The output of the microphone is then a square wave whose amplitude is modulated by

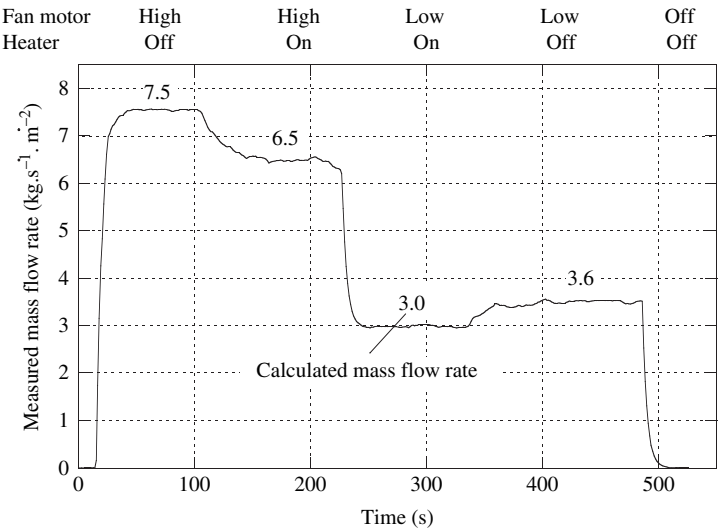


FIGURE 14 Measured mass flow velocity.

the difference of the two input pressures. By amplifying and then demodulating the microphone output signal with a reference signal as shown in Figure 16, we can obtain an output voltage proportional to the pressure difference $p_1 - p_2$. Since the demodulated output is not affected by an offset in the microphone and the a.c. amplifier, we have a differential manometer which is as sensitive as a microphone and drift free. A trial device which employed a rotating chopper (the chopping frequency = 10 Hz)

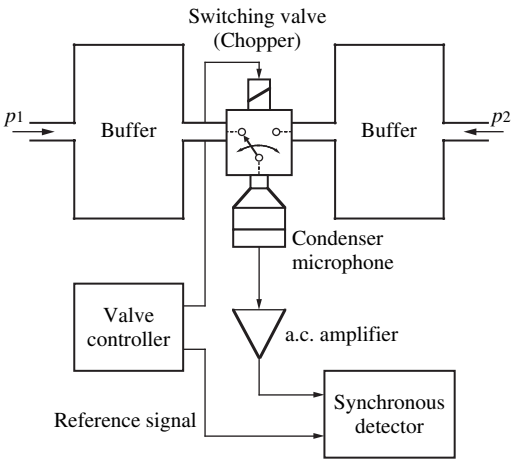


FIGURE 15 Schematic diagram of the chopper microphone.

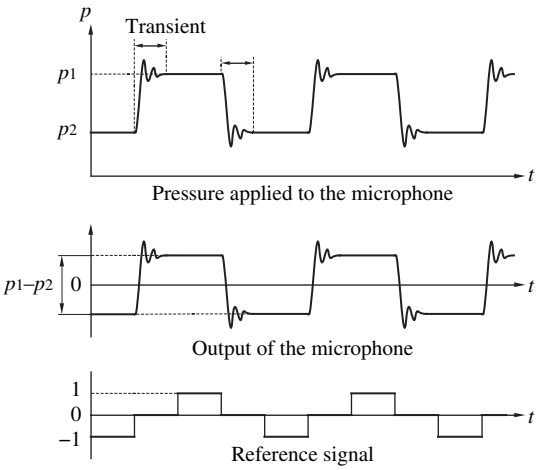


FIGURE 16 Schematic diagram of the signal waveforms.

was capable of detecting differential pressure to a resolution higher than 10^{-3} Pa.

There is a small cavity between the chopper and the microphone diaphragm. When the microphone is connected to the higher pressure port, the cavity is compressed to this pressure, then the microphone is switched to another pressure port and the cavity is rarefied to the lower pressure. This means that a small volume of gas is conveyed via the cavity in each switching cycle; the rate is proportional to the pressure difference, the cavity volume, and the switching frequency. The input resistance of the chopper microphone is thus finite, as is not the case with ordinary diaphragm-type manometers. Utilizing this property from a different point of view, we can exploit it for measuring and/or controlling micro mass flow of gases [20].

7 • Kinematic Viscosity Sensor [27]

Flow field is characterized by the kinematic viscosity ν – the ratio of coefficient of viscosity μ to density ρ of the fluid – together with the flow velocity and the characteristic length. In most applications, therefore, it is important to have knowledge of the kinematic viscosity of fluid. Figure 17 shows the principle of the kinematic viscosity sensor. The sensor consists of a sensing channel, a driving speaker, microphones, and phase-detecting electronics. The speaker is driven in a sinusoidal manner so as to vibrate the fluid contained within the sensing channel. The sinusoidal acceleration of the fluid causes a sinusoidal variation of the axial pressure gradient in the channel. Two pressure taps are drilled a small distance apart along the major axis of the

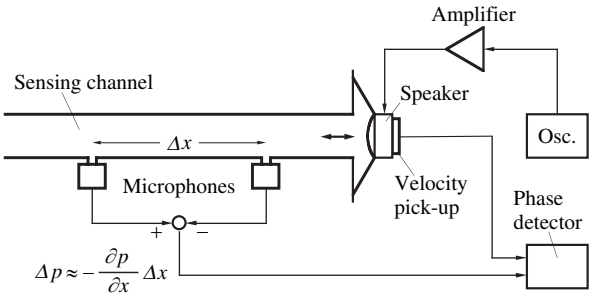


FIGURE 17 Schematic diagram of the kinematic viscosity sensor.

channel. The pressure difference, the finite difference approximation for the axial pressure gradient in the channel, between these taps is detected using the microphones.

This pressure gradient variation includes two components: (i) a component proportional to the velocity and viscosity of fluid in the sensing channel, and (ii) a component proportional to the acceleration and density of the fluid. The ratio of these two components – the phase angle between the fluid velocity and the pressure gradient – is a function only of the vibration frequency ω and kinematic viscosity ν of the fluid. For example, for a channel with a low-profile rectangular cross-section, the relationship between the pressure gradient and the flow velocity is written as

$$-\frac{\partial p}{\partial x} = \mathrm{i}\omega\rho u_{\mathrm{a}} \left\{ 1 - \frac{1}{1+\mathrm{i}} \frac{\delta_{\mathrm{v}}}{h} \tanh\left(1+\mathrm{i}\right) \frac{h}{\delta_{\mathrm{v}}} \right\}^{-1}, \tag{9}$$

where p is the pressure, x is the distance along the channel axis, u_{a} is the flow velocity averaged over the channel cross-section, h is the height of the channel, and the thickness of the vibration boundary layer δ_{v} is given as

$$\delta_{\mathrm{v}} = \sqrt{\frac{2\nu}{\omega}}. \tag{10}$$

The kinematic viscosity ν of the fluid can be determined from the measurement of the phase angle at a given vibration frequency or from the measurement of the frequency that results in a given phase shift.

Figure 18 shows the experimental results for the measurement of kinematic viscosity of air, with the temperature of air varied in order to change the kinematic viscosity. The phase shift between the pressure gradient and the flow velocity was maintained at a constant value ($\pi/4$ rad), and the frequency that produced this

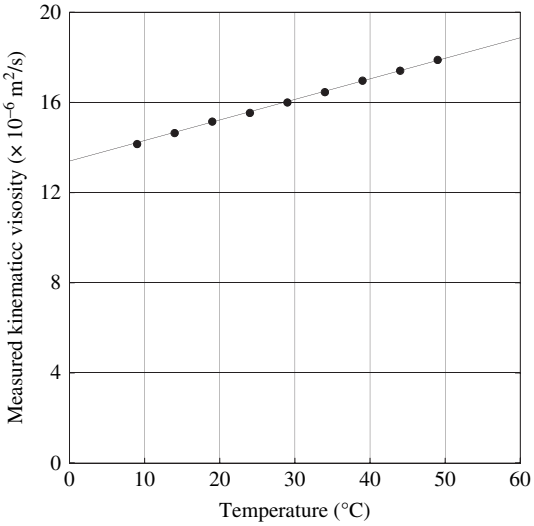


FIGURE 18 Measured kinematic viscosity of air.

phase shift was measured (for the channel in this experiment, the frequency was in a range around 40 Hz). The solid line shows the published values.

It may be worth pointing out in passing that if we control the vibration velocity amplitude at a constant value and use a precise manometer, we can measure the density ρ and the viscosity coefficient μ of the fluid separately using the same configuration and a phase-sensitive detector [23].

References

1. The entry "sound", The New Encyclopedia Britannica, Volume 11, p.26 (1979)
2. A.P. Cracknell, Ultrasonics, Wykeham Publication (1980)
3. D. Ensminger, Ultrasonics: Fundamentals, Technology Applications, 2nd Rv& Exp ed., CRC Press (1988)
4. A. Rosencwaig, Photoacoustics and Photoacoustic Spectroscopy, Wiley (1980)
5. C. Suzuka, New Infrared Gas Analyzer and its Application, Keisoku Gijutsu (Measurement Technology)1, 10(5), 71–75 (1982)
6. R.A. Collacott, Mechanical Fault Diagnosis and Condition Monitoring, Chapman and Hall (1977)
7. K. Mori, I. Torigoe and N. Momosaki, Nondestructive inspection of concrete structures by using sound wave, Mater. Sci. Forum, 465–466, 355–360, on-line http://www.scientific.net (2004)
8. I. Torigoe, K. Mori, Y. Sato, Y. Ohshima and T. Iwamoto, Non-destructive and con-contacting test method utilizing sound intensity, J. Jpn. Soc. Non-destruct. Inspect., 55(2), 84–88 (2006)

9. H.E. Hartig and H.B. Wilcox, Fluid meter, US Patent 1881543 (1932)
10. T.J. Scarpa, Acoustic flowmeter, US Patent 3580092 (1971)
11. T. Horiuchi, Measurement of flow rate of gas and powder using noise, J. Soc. Instrum. Cont. Eng., 14(7), 547–552 (1975)
12. K.M. Forman and R.A. Oman, Acoustic emission flow measurement system, US Patent 3958458 (1976)
13. E.J. Dewath, Acoustical wave flowmeter, US Patent 4003252 (1977)
14. Y. Ishii, Vibrating Pitot tube, Trans. Soc. Instrum. Cont. Eng., 16(4), 556–561 (1980)
15. J.E. Potzick and B. Robertson, Long-wave acoustic flowmeter, ISA Trans., 22(3), 9–15 (1983)
16. I. Torigoe, Mass flowmeter utilizing a pulsating source, Trans. Soc. Instrum. Cont. Eng., E1(1), 77–82, on-line at http://srv01.sice.or.jp/~e-trans/(1987)
17. I. Torigoe, A new method for measuring flow acceleration and its application to pulsating flow measurement, Trans. Soc. Instrum. Cont. Eng., 23(8), 794–799 (1987)
18. I. Torigoe and Y. Ishii, Acoustic impedance flowmeter, Trans. Soc. Instrum. Cont. Eng., 26(2), 145–150 (1990)
19. P. Shakkottai, E.Y. Kwack and L.H. Back, Aeroacoustic flow sensor, Rev. Sci. Instrum., 62(9), 2205–2212 (1991)
20. I. Torigoe, Switched cavity: an approach to micromass-flow sensors, Rev. Sci. Instrum., 73(6), 2475–2478 (2002)
21. J.O. Kim and H.H. Bau, On line, real-time densimeter – theory and optimization, J. Acoust. Soc. Am., 85(1), 432–439 (1989)
22. B.C.B. Coleman and P. Fryer, Fluid density measurement using transducers based upon loudspeaker technology, Rev. Sci. Instrum. 62(12), 3068–3074 (1991)
23. I. Torigoe and Y. Ishii, Density meter utilizing pressure gradient variation, Trans. Soc. Instrum. Cont. Eng., 31(3), 308–314 (1995)
24. I. Torigoe, Simulated acoustic inductor, Trans. Soc. Instrum. Cont. Eng., 38(5), 514–516 (2002)
25. M. Hoummady and F. Bastien, Acoustic wave viscometer, Rev. Sci. Instrum., 62(8), 1999–2003 (1991)
26. K.A. Gillis, J.B. Mehl and M.R. Moldover, Greenspan acoustic viscometer for gases, Rev. Sci. Instrum., 67(5), 1850–1857 (1996)
27. I. Torigoe, Kinematic viscosity meter utilizing the phase of pressure gradient variation, Rev. Sci. Instrum., 74(1 Part I), 164–167 (2003)
28. J.Ohga, Mechanical resonance type pressure transducer using stiffness of the cavity, J. Acoust. Soc. Jpn, 30(10), 602–607 (1975)
29. I. Torigoe and Y. Ishii, Chopper microphone, Membership Journal of Jpn. Soc. Appl. Phys., 52(5), 425–428 (1983)
30. I. Shimbo, Liquid level measuring system using sound tube, J. Soc. Instrum. Cont. Eng., 5(12), 593–598 (1955)
31. I.H. Cohn and W.E. Dunn, Precision techniques for measuring liquid quantity, Cont. Eng, 15(1), 51–55 (1968)
32. K. Watanabe, Y. Takebayashi and D.M. Himmelblau, Volume measurements of deformed materials in tanks, ISA Trans, 27(4), 9–19, (1988)

33. C.E. Granqvist, Gyroscope device with vibrating gas particles or particles of another sound transferring medium, US patent 2999389 (1961)
34. M. Bruneau, C. Garing and H. Leblond, A rate gyro based on acoustic mode coupling, J. Acoust. Soc. Am., 80(2), 672–680 (1986)
35. Y. Ishii, A rate gyro using an acoustic bridge, Trans. Soc. Instrum. Cont. Eng., 26(12), 1356–1361 (1990)
36. Y. Kagawa and H. Tsubota, Vibrator sensors for measurement of atmospheric pressure, temperature and humidity, J. Acoust. Soc. Jpn, 29(3), 144–150 (1973)
37. M. Komori, Method and apparatus for leak test of gas piping, Jpn patent H9-163584 (1997)
38. I. Torigoe, Method and apparatus for measuring thermal properties, Jpn patent 2005-22452 (2005)
39. H. Itokawa, A new method of measuring a small displacement by acoustic impedance change, Annu. Rep. Inst. Ind. Sci. Univ. Tokyo, 2, 255–275 (1952)
40. R. Ichimiya, K. Miyashita and M. Goto, Accurate measurement of displacement by means of sound frequency in a resonant pipe, Trans. Jpn. Soc. Mech. Eng. (C), 51(471), 3077–3081 (1985)
41. F. Sakao, Sensing displacement by measuring acoustic pressure, Trans. Soc. Instrum. Cont. Eng., 21(12), 1336–1341 (1985)
42. Y. Ishii, Acoustic micrometer, J. Acoust. Soc. Jpn, 43(9), 648–654 (1987)
43. R.W. Leonard, Simplified porosity measurements, J. Acoust. Soc. Am., 20, 39–41 (1948)
44. A.B. Dubois et al., A rapid plethysmographic method for measuring thoratic gas volume: a comparison with a nitrogen washout method for measuring functional residual capacity in normal subjects, J. Clin. Invest., 35, 322–326 (1956)
45. I. Torigoe and Y. Ishii, Volume measurement using Boyle–Charles' law, Proc. Acoust. Soc. Jpn Annu. Conf., 437–438 (Oct. 1982)
46. W.G. Deskins, D.C. Winter, H. Sheng and C. Garza, Use of a resonating cavity to measure body volume, J. Acoust. Soc. Am., 77(2), 756–758, (1985)
47. Y. Ishii, Volume measurement using an acoustic resonator, Proc. 27th Soc. Instrum. Cont. Eng. Annu. Conf., 131–132 (1988)
48. Y. Yanagida, Y. Ishii and I. Torigoe, Volume measurement using the antiresonant frequency of an acoustic system, Proc. Acoust. Soc. Jpn Annu. Conf., 501–502 (Oct. 1989)
49. R. Ichimiya, T. Endo and A. Ishikawa, Study of volume measurement using sound signal, Trans. Jpn. Soc. Mech. Eng. (C), 57(541), 2888–2893 (1991)
50. I. Torigoe and Y. Ishii, Acoustic bridge volumeter, Trans. Soc. Instrum. Cont. Eng., E1(1), 164–170, on-line at http://srv01.sice.or.jp/~e-trans/(1994)
51. Y. Ishii, Acoustic volume meter, J. Soc. Instrum. Cont. Eng., 36(4), 288–291 (1997)

52. I. Torigoe, Measurement of human body volume employing a static puff, Trans. Soc. Instrum. Cont. Eng., 35(9), 1129–1134 (1999)
53. I. Torigoe, Surface area measurement utilizing an acoustic bridge, Proc. 35th Soc. Instrum. Cont. Eng. Annu. Conf., 975–976 (1996)
54. I. Torigoe and Y. Ishii, Surface area measurement utilizing sound, Trans. Soc. Instrum. Cont. Eng., 34(3), 182–187 (1998)
55. I. Torigoe, Edge correction of the acoustical surface area, Trans. Soc. Instrum. Cont. Eng., 34(11), 1739–1741 (1998)
56. I. Torigoe and Y. Ishii, Surface area measurement utilizing an acoustic bridge, J. Acoust. Soc. Am., 107(5 Part 1), 2548–2553 (2000)
57. L. Kiesewetter et al., Determination of Young's moduli of micromechanical thin films using the resonance method, Sens Actuat A, 35, 153–159 (1992)
58. T. Sugiura et al., Non-destructive estimation of sugar content of peach by using of specific gravity, J. Jpn. Soc. Hortic Sci., 65 (Suppl. 1), 154–155 (1996)
59. Y. Ishii and I. Torigoe, Measurement using low frequency sound, J. Acoust. Soc. Jpn, 53(5), 399–406 (1997)
60. http://www.rion.co.jp/dbcon/pdf/VM-230_120_300.pdf
61. 1994 OIML regulation R111, edition 1994(E) Weights of classes E1, E2, F1, F2, M1, M2, M3, Bureau International de Metrologie Legale
62. T. Kobata, M. Ueki, A. Oiwa and Y. Ishii, Measurement of the volume of weights using an acoustic volumeter and the reliability of such measurement, Metrologia, 41, S75–S83 (2004)
63. BIMP, IEC, IFCC, ISO, IUPAP, OIML, Guide to the expression of uncertainty in measurement, ISO (1993)

54

Community Response to Noise

Sandy Fidell

Fidell Associates, Inc., Woodland Hills, CA, USA

The primary effects of community noise on residential populations are speech interference, sleep disturbance, and annoyance. This chapter focuses on transportation noise in general and on aircraft noise in particular because aircraft noise is one of the most prominent community noise sources, because airport/community controversies are often the most contentious and widespread, and because industrial and other specialized forms of community noise generally pose more localized problems.

Great variability and seeming unpredictability are among the most obvious features of transportation noise exposure and its adverse effects on residential populations. Short- and long-term outdoor noise levels in communities can vary by 40 dB or more with population density and time of day. Some populations tolerate with apparent equanimity 24-h, time-weighted noise exposures that are 30 dB greater than those which others find unacceptable. Likewise, individuals who have adapted to sleeping in noisy, high-density urban surroundings may have difficulty sleeping in the stillness of rural settings. Despite their great ranges, certain regularities underlie the variability of community noise and its effects.

1 • Statistical Character of Environmental Noise in Populated Areas

Time-weighted average outdoor sound levels in motorized societies are strongly affected by population density and time of day. Galloway, Eldred, and Simpson (1974) have shown that outdoor cumulative noise exposure levels in populated areas can be approximated as shown in Equation 1. Sound levels are expressed in Equation 1 in units of Day–Night Average Sound Level (abbreviated as DNL and represented in mathematical expressions as L_{dn}), a 24-h time-weighted average sound level, with a 10-dB "penalty" applied to the hours between 2200 and 0700.

The so-called nighttime "penalty" built into some measures of cumulative noise exposure might equally well be described as a daytime license. From a regulatory perspective, there is little to be gained from encouraging operators of noise sources to refrain from noisy activities at times of day at which they would be unlikely to engage in them anyway. For example, if the goal of regulation is to avoid speech interference (for example, in schools, offices, or residences), why penalize noisy activities at times of day when people are unlikely to benefit much from such restraint?

It is sometimes assumed that nighttime noise penalties minimize sleep disturbance. Since sleep disturbance is much more closely linked to single event levels inside sleeping quarters than to 24-h average outdoor levels (Fidell, Pearsons, Tabachnick, and Howe, 2000), however, a weighting applied over a 9-h interval is at best an awkward regulatory tool. A regulatory goal of discouraging specific activities which generate high episodic noise levels, such as nighttime aircraft operations, would be more effectively implemented by means other than a time weighting for long-term average sound levels.

Neither the magnitude nor the time of imposition of the penalty is based on systematic evidence of greater sensitivity to noise exposure during the hours of application. Different nations have adopted variant policies for the number, duration, and timing of temporal weightings, based on little more than intuition. In U.S. urban areas away from airports and highways, the average difference at 100 measurement sites between an unweighted 24-h equivalent level and DNL is about 3.4 dB (Galloway et al., 1974).

Population density is expressed in people per square mile.

$$L_{dn}(\text{dB}) = 10\lg(\text{population density}) + 22 \qquad (1)$$

TABLE 1 Examples of typical DNL values in areas of varying population density

Setting	Nominal population density (people per square mile)	Expected day–night average sound level (db)
Wilderness (unpopulated)	< 1	22
Rural	500	49
Low density (suburban)	5, 000	59
Moderate density (urban)	25, 000	67
High density (urban)	100, 000	72

Table 1 contains examples of typical DNL values estimated from Equation 1. Population density is closely related not only to noise exposure but also to lifestyle. High population density living implies more high-rise and party wall residential construction, greater use of public transportation, more street traffic, more indoor living, and less exposure to natural outdoor ambient sound levels than low population density and rural living.

The median population density of urban areas of the United States is about 5,000 people per square mile. In a typical suburban setting, this figure implies an average household size of about four people per detached, single-family dwelling built on a half-acre lot. Equation 1 implies that half of the U.S. urban population lives in areas with outdoor noise levels in excess of $L_{\mathrm{dn}} = 59\,\mathrm{dB}$, while the other half lives in areas with lower outdoor noise levels.

Community noise exposure fluctuates seasonally, diurnally, and in the shorter term. These fluctuations can be usefully characterized as the sum of two Gaussian processes (Fidell, Horonjeff, and Green, 1981). At any given measurement point in a populated area, a "distant" process, composed of the sum of a great many remote noise sources, has a relatively low mean and variance. A "local" process, composed of the sum of a relatively small number of noise sources in the immediate vicinity of the measurement point, has a high mean and variance. When plotted in cumulative form, the distribution of short-term outdoor noise levels in populated areas often has a distinct breakpoint separating segments with steep and shallow slopes. The lower centiles of the distribution are created by the distant

process, while the upper centiles of the distribution are created by the local process.

The breakpoint between the steep slope (low variance, low level) and shallow slope (high variance, high level) segments varies with time of day, as the distant and local processes contribute differentially to the joint distribution of measured noise levels. In the late night and early morning hours, for example, the breakpoint shifts toward higher centile values as the bulk of the urban noise measurable at any given point is generated by the distant process. At busier times of day, the breakpoint shifts toward lower centile values as the local process contributes more greatly to the joint distribution.

2 • Predicting the Annoyance of Environmental Noise

According to the U.S. Federal Interagency Committee on Noise (FICON 1992), "…the percent of the [noise] exposed population expected to be Highly Annoyed (%HA) [is] the most useful metric for characterizing or assessing noise impact on people." Noise-induced annoyance is rather loosely defined as a general adverse reaction to community noise exposure. The prevalence of a consequential degree of annoyance in a community ("percentage highly annoyed") is simply 100 times the number of survey respondents who describe themselves as highly annoyed when their opinions are solicited in a social survey, divided by the number of respondents to whom the question was posed. Identification of the prevalence of a consequential degree of annoyance as a summary measure of transportation noise impacts on exposed populations is a relatively recent development that followed from the work of Schultz (1978). Prior to Schultz's work, activity (speech and sleep) interference and complaint behavior served as the more common bases for evaluating both individual and population-level aircraft noise impacts (Fidell, 2003).

FICON (1992) not only identifies annoyance as its preferred "summary measure of the general adverse reaction of people to noise" for assessment of community-level environmental noise effects but also endorses a specific fitting function (given in Equation 2 and seen in Figure 1) as its preferred dosage-effect relationship. The slope of this curvilinear relationship varies continuously over its range, from about 1% highly annoyed per 1 db increase in DNL at lower levels to as much as 3% highly annoyed per 1 db increase in DNL at higher levels.

$$\%\text{highly annoyed} = 100/(1+\exp(11.13–0.141\ L_{\text{dn}})). \qquad (2)$$

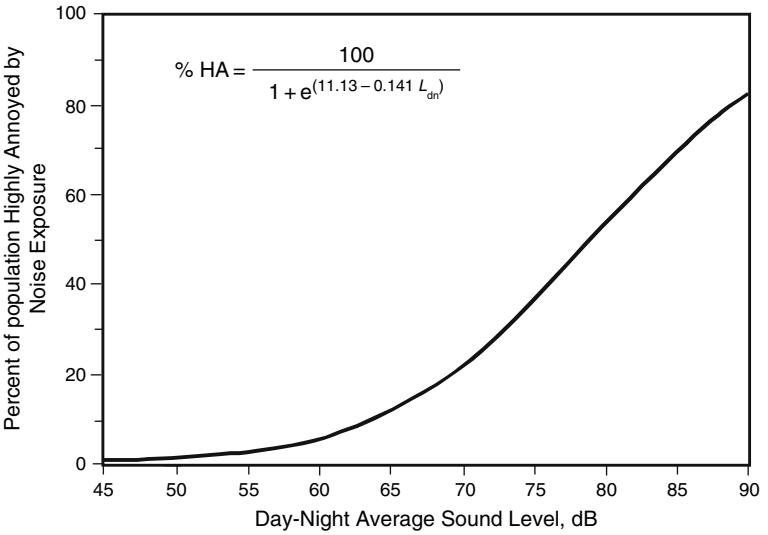


FIGURE 1 Dosage–effect relationship favored by the U.S. Federal Interagency Committee on Noise (FICON, 1992).

Figure 2, adapted from Fidell and Silvati (2004), shows not only the great variability of the data on the prevalence of aircraft noise-induced annoyance but also the underestimation bias of the FICON curve with respect to the empirical means of the data (shown as dark triangles) within ±2.5 dB exposure intervals.

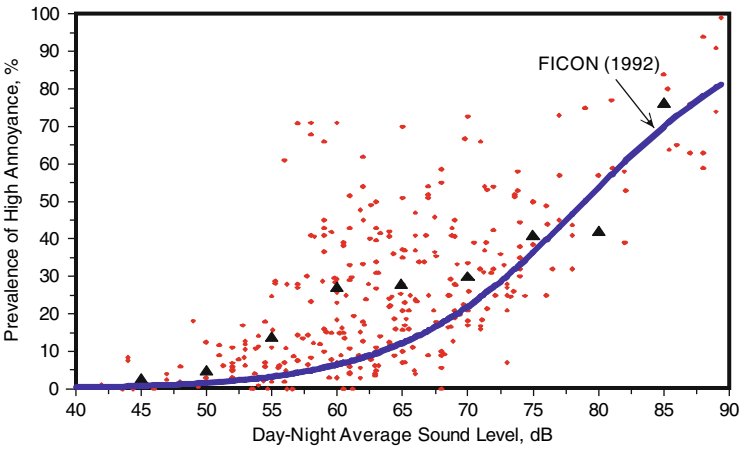


FIGURE 2 Underestimation of FICON (1992) dosage–effect relationship with respect to prediction of the prevalence of aircraft noise-induced annoyance (*individual points*). *Filled triangles* are weighted mean annoyance prevalence rates within ±2.5 dB intervals.

Perversely, the error of underestimation is the largest in the region of greatest regulatory interest ($55\,\text{dB} \leq L_{\text{dn}} \leq 65\,\text{dB}$).

Neither the functional relationship illustrated in Figure 1 nor the distribution of field observations seen in Figure 2 is self-interpreting for purposes of regulatory policy. No obvious inflection point is apparent in the data that would strongly compel identification of a regulatory threshold or a definition of an environmentally "significant" level of community response. To the extent that such guidance can be inferred from the data or summary relationships at all, an argument might be made for an inflection point in the vicinity of $L_{\text{dn}} = 55\,\text{dB}$ (Schomer, 2005). Selection of such a regulatory policy point is an expressly nontechnical judgment, however, that must balance conflicting societal interests, costs, and benefits.

Prediction of a prevalence rate for aircraft noise-induced annoyance based on exposure alone is an unreliable expedient because the descriptive approach that Schultz (1978) pioneered does not explain the great variability of community reaction. Predictions of aircraft noise impacts made by applying FICON's purely descriptive fitting function account for less than one-fifth of the variance in aircraft noise annoyance data (Job, 1988; Miedema and Vos, 1998; Fidell, 2003) and are prone to sizable error (Silvati and Fidel, 2004). Noise exposure per se is not a particularly effective predictor of the prevalence of annoyance, and a given level of noise exposure is neither a necessary nor a sufficient antecedent condition for either annoyance or complaints with transportation noise.

Functional relationships between community response to noise and acoustic predictor variables are thus susceptible to oversimplifications and overinterpretations that ignore the problems of underlying imprecision and multiple causation. For example, predictions of absolute numbers of individual awakenings, or minutes of speech interference associated with projected numbers of aircraft operations, are not uncommon. Given the limited precision of the data from which such functions have been inferred, however, overly precise predictions of annoyance, sleep disturbance, and speech interference are unjustifiable.

In part, the seeming unpredictability of annoyance prevalence rates may reflect the fact that a 24-h, time-weighted average sound level such as DNL may not be an optimal acoustic predictor variable. Reliance on DNL as a predictor is predicated upon the acceptance of the "equal energy hypothesis" — the notion that the number, sound level, and duration of noise events are fully fungible determinants of annoyance. Because DNL is an exposure metric that reflects the total acoustic energy of

a noise environment, it is a convenient measure of the entire noise burden imposed by all noise sources on a community. However, the evidence in favor of the equal energy hypothesis is equivocal. The findings of Fidell and Pearsons (2003), for example, suggest that numbers and levels of prospective aircraft operations (a critical matter in many transportation noise impact assessments and environmental disclosure documents) may not be interchangeable in the manner posited by the equal energy hypothesis.

As a total exposure metric, DNL of necessity correlates highly with numbers, levels, and durations of noise events. Given the high correlation between DNL and most other environmental noise metrics computed from numbers, levels, and/or durations of noise events, it is unlikely that any other exclusively acoustic index can serve appreciably better than DNL as a predictor of community reaction to environmental noise.

It is more likely that DNL is an imperfect predictor of annoyance because community reaction to aircraft noise is strongly influenced by situational and non-acoustic factors (Rosenblith and Stevens, 1953; Green and Fidell, 1991; Schomer, 2002; Fidell, 2003; Miedema, Fields, and Vos, 2005). For example, by soliciting annoyance judgments under controlled listening conditions of pairs of sounds of identical power spectra but different phase spectra, Fidell, Sneddon, Pearsons, and Howe (2002) demonstrated that the meaning of a sound can affect its annoyance. Pairs of sounds were created by scrambling the phase spectra of meaningful sounds to create meaningless twin signals of the same equivalent level. Although these signal pairs were indistinguishable to a sound level meter, their judged annoyance varied considerably, by orders of magnitude in some cases.

Fidell, Schultz, and Green (1988) and Green and Fidell (1991) have modeled self-reports of noise-induced annoyance as though they were the product of a decision-making process, in which social survey respondents decide in which response category (commonly, "Not at all annoyed", "Slightly annoyed", "Moderately annoyed", "Very annoyed", or "Extremely annoyed") to report themselves. Like all other perceptual decisions, a respondent's self-report of degree of annoyance inextricably confounds true sensitivity to physical stimuli with response bias.

Green and Fidell (1991) extended the psychophysical theory of signal detectability (Swets, 1964) to the analysis of environmental noise impacts. Noise exposure is viewed as a treatment administered to a population of individuals in a community. The reactions of individuals to their exposure may be represented by a random variable x that is assumed to be exponentially

distributed, with a mean population value of m. This mean value is assumed to be a power law transformation of noise exposure, with an exponent of 0.3 (the rate of growth of loudness with sound level). Individuals describe themselves as highly annoyed by noise exposure if their annoyance, x, exceeds some response criterion, A. The proportion of the population describing itself as highly annoyed can be predicted from the area of the distribution of x that is greater than the value of A chosen for reporting annoyance. The prevalence of a consequential degree of noise-induced annoyance in a community is therefore given by $e^{-A/m}$. As illustrated in Figure 3, this function yields a dosage-response curve with a fixed (ogival) form and slope, which response bias translates horizontally along the cumulative noise exposure axis.

The only free parameter of the model, A, is estimated from social survey data. Deviations from the assumed exponential rate of growth in the prevalence of annoyance with cumulative noise exposure are attributed to response bias. Green and Fidell (1991) showed that this model can generate a dosage-response relationship derived from first principles that retrospectively accounts for more than 70% of the variance in a worldwide set of observations of noise-induced annoyance. They also showed that much of the remaining variance may be due to sampling and measurement error.

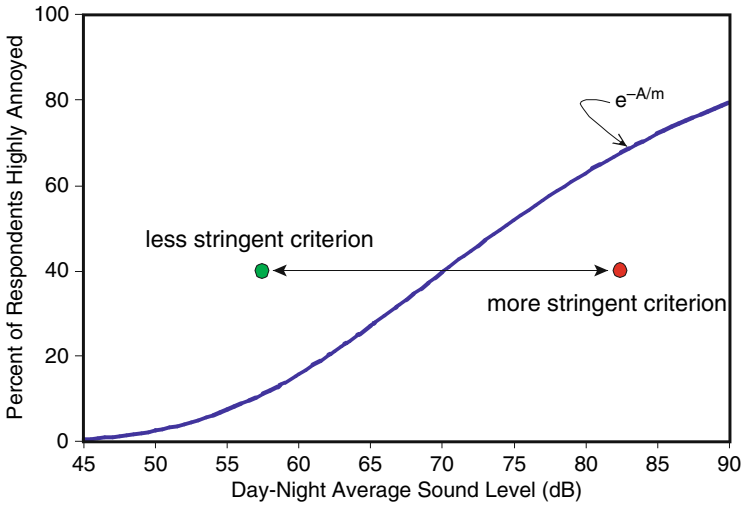


FIGURE 3 Translation of dosage–effect relationship derived from first principles along the abscissa by response bias (non-acoustic) factors.

Predictive application of this theory-based model requires prospective assessment of the degree of response bias in a given community: that is, the ability to reliably quantify the degree to which a community is favorably or unfavorably disposed toward noise exposure as a function of non-acoustic factors. In the context of airport/community controversies, these factors may include the economic dependence of a community on airport operation; demographic, political, land use, and community development factors; the extent of recent and historical media coverage of aircraft noise issues; the degree to which community interests are respected by airport administrations; and so forth.

It is convenient for interpretive purposes to express the effect of response bias in units that can be compared directly to noise exposure values. Algebraic manipulation of A yields a D^* parameter that may be thought of as an average noise exposure level at which people describe themselves as highly annoyed. As calculated by Fidell and Silvati (2004), this value is 69.3 dB for community reaction to aircraft noise exposure. Figure 4 shows that a predictive relationship based on first principles yields a more defensible fit to the empirical data on aircraft noise annoyance than the FICON (1992) relationship.

The common assumption that annoyance prevalence rates inferred from measurements made under steady-state exposure conditions may be generalized to settings in which noise exposure is increasing or decreasing is problematic. Fidell, Silvati, and Pearsons (1998) have shown that a genuine decrease in aircraft

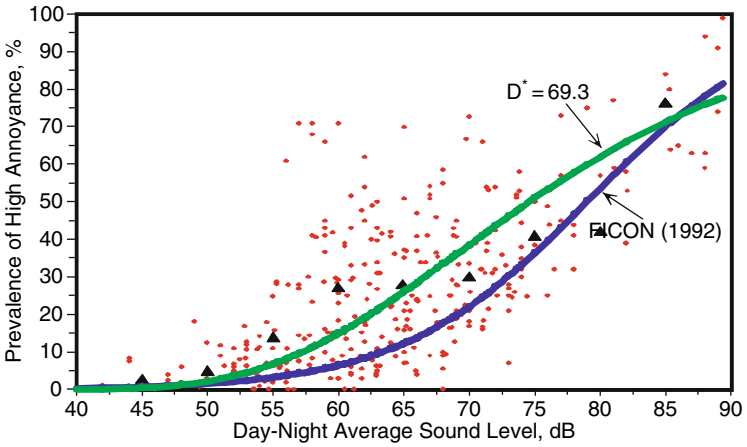


FIGURE 4 Comparison of theory-based prediction of aircraft noise-induced annoyance with FICON (1992) relationship, all data points, and mean annoyance prevalence rates.

noise exposure at Seattle–Tacoma International Airport that occurred gradually (over the course of several years) went essentially unnoticed in affected neighborhoods near the airport. On the other hand, Fidell, Silvati, and Haboly (2000) showed that a sudden increase at relatively low noise exposure levels at Vancouver International Airport was associated with a disproportionately large increase in the prevalence of annoyance in a neighborhood newly exposed to aircraft landing noise.

An earlier study (Fidell, Horonjeff, Mills, Baldwin, Teffeteller, and Pearsons, 1985) had shown that short-term ("last week") annoyance prevalence rates tracked changes in noise exposure associated with intermittent runway maintenance closures, but that longer term annoyance prevalence rates showed considerable hysteresis. This study modeled the temporal changes in attitude associated with changes in noise exposure as an exponential process analogous to charging or discharging a capacitor. Although the study was unable to produce precise estimates of the time constants of the process, it appears that at least weeks to months are required to fully charge or discharge noise-induced annoyance following a notable change in exposure.

The large variance in the level of community annoyance reported for any given level of DNL noise exposure in multiple studies has important policy and planning implications. Communities with much higher percentages of households that are highly annoyed by aircraft noise than predicted by the FICON-endorsed relationship are likely to oppose airport expansion plans, even when planning paradigms based on FICON's relationship suggest that flight activity should be acceptable to a large proportion of the population.

Many of the limitations of the current paradigm for predicting community reaction to transportation noise follow from the noncausal nature of Schultz's (1978) relationship. The most obvious shortcoming of the many curve-fitting exercises that have followed Schultz's is that none accounts for the better part of the variance in what is now a very large body of social survey data on the prevalence of annoyance associated with environmental noise exposure. Efforts to explain the variance in annoyance rates through differences in the personal or situational characteristics of the respondents have been largely unsuccessful (Fields, 1993), and are not, in any event, of much relevance for public policy purposes. All neighborhood residents are entitled to equal protection of the law from exposure to the noise of street traffic and aircraft, which insonifies the homes of the more- and less-sensitive alike.

Only ad hoc explanations are available for large differences in annoyance prevalence rates in different communities with the

same noise exposure. The lack of systematic explanations means that accurate predictions of the prevalence of annoyance in communities exposed to change in noise levels (for example, from increases in air traffic) remain elusive. The lack of systematic understanding further means that prediction of the benefits of costly measures intended to mitigate noise exposure cannot be made with confidence, and that regulatory policies remain poorly supported by technical analysis (Fidell, 1999).

Two other issues that the relationships seen in Figure 4 do not address are (1) annoyance due to secondary emissions (indoor rattling noises produced by household paraphernalia due to either ground-borne or airborne noise and/or vibration) and (2) the role of ambient noise levels on the noticeability and annoyance of low-level noise intrusions. Figure 5 shows a summary relationship for annoyance due to rattle, constructed from the results of two studies of the annoyance of runway sideline noise (Fidell, Pearsons, Silvati, and Sneddon, 2002; Fidell, Silvati, Pearsons, Lind, and Howe, 1999). The predictor variable, Low-Frequency Sound Level, is a single-event metric calculated from airborne sound levels in the six one-third octave bands between 25 and 80 Hz, inclusive.

Sounds that are either inaudible or unnoticeable in settings with higher ambient noise environments may be of greater concern in unusually low-ambient noise environments (Fidell, Silvati, Tabachnick, Howe, Pearsons, Knopf, Gramann, and Buchanan, 1996). In such settings, the proportion of time that sounds are audible (Horonjeff and Anderson, 2004) or noticeable (Sneddon, Pearsons, and Fidell, 2004) may be of greater concern.

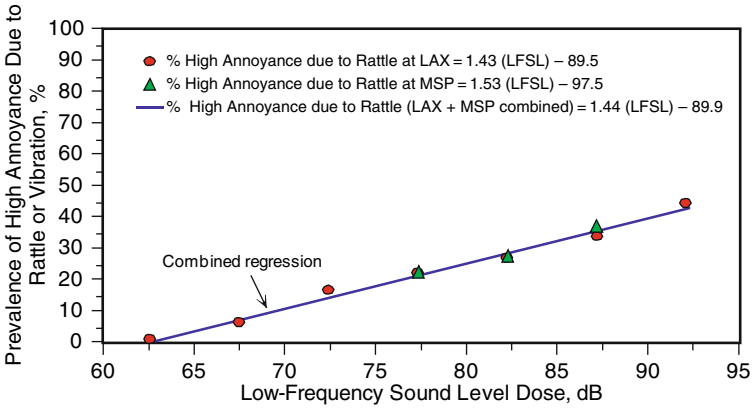


FIGURE 5 Dosage–effect relationship for annoyance due to interior rattle induced by low-frequency noise from aircraft ground operations.

3 • Predicting Sleep Disturbance

Unlike annoyance, noise-induced sleep disturbance is not plausibly predictable from 24-h time-weighted average outdoor sound levels. Instead, indoor sound exposure levels (short-term, single-event equivalent levels, normalized to a 1 s interval) have been found to be more closely associated with in-home measurements of behavioral awakening (Fidell, Pearsons, Tabachnick, and Howe, 2000). ANSI (2000) recommends the relationship seen in Equation 3 for prediction of the prevalence of behaviorally confirmed awakening in residential populations familiar with the local nighttime residential noise environment (Figure 6):

$$\text{Prevalence of awakening due to individual aircraft noise intrusions} = -7.02 + 0.14\ (\text{SEL}). \tag{3}$$

It should be noted that this relationship, derived by regression analysis, accounts for only about one-fifth of the variance in field measurements; and that the slope of the relationship is so shallow that it provides little useful guidance for defining regulatory thresholds. Note also that the predictor variable in Equation 2 is an indoor sound exposure level at the sleeper's ear, which can be only roughly estimated from outdoor noise measurements. It is further unclear whether prevalence of awakening can be reliably extrapolated from the circumstances in which observations were

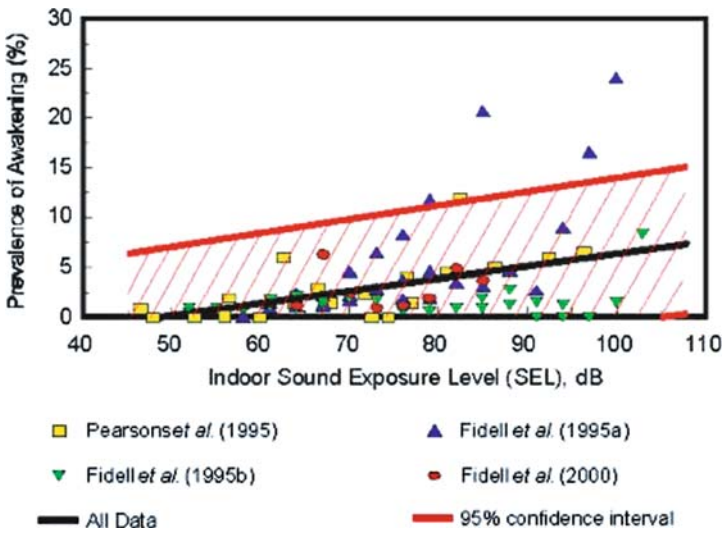


FIGURE 6 Dosage–effect relationship for prevalence of awakening due to indoor sound exposure levels in sleeping quarters.

made (i.e., modest numbers of individual awakenings associated with relatively small numbers of aircraft overflights in a few airport settings.) The effects of novel exposure conditions and unfamiliar or changed exposure regimes on sleep disturbance remain poorly understood.

4 • Predicting the Likelihood of Noise-Induced Complaints

Annoyance, the officially preferred measure of community reaction to transportation noise, is an attitude, while a complaint is a behavior. While annoyance prevalence rates are at least partially predictable from cumulative noise exposure levels, complaints appear to be more closely linked to exceptional noise events and relatively short-term periods of noise exposure than to long-term average conditions. The great bulk of aircraft noise complaints received by airports, for example, comes from residential areas beyond the 65-dB DNL contours of airports (GAO, 2000). These complaints are inexplicable from the perspective of standard interpretive criteria for the compatibility of residential land uses with aircraft noise, which hold that such land uses are fully "compatible" with aircraft noise at exposure levels below $L_{dn} = 65$ dB.

Note that the term "compatibility", as used in land use compatibility recommendations issued by a regulatory agency originally chartered to foster the growth of civil aeronautics, does not correspond to the commonly understood definition of mutual and reciprocal compatibility between conflicting interests. "Land use compatibility" as used by the U.S. Federal Aviation Administration, for example, implies land uses which do not threaten the continued use and expansion of airport infrastructure, not the compatibility of aircraft noise exposure with the unfettered enjoyment of residential property.

Recent studies suggest that aircraft noise complaints may be more closely linked with proximity to flight tracks than with the integrated energy of exposure to multiple overflights (Fidell and Howe, 1998; Fidell, 2003). For example, Figure 7 shows two "mountains" in complaint density (rendered as shaded false elevation) along the extended centerline of the primary departure runway of a small municipal airport. These complaint mountains lie under the primary departure routing for noisy business jets from the airport. Fewer complaints are observed along the extended centerlines of the airport's other runway ends primarily because they support quieter operations (by lighter aircraft and for approaches) and because land use planning has minimized

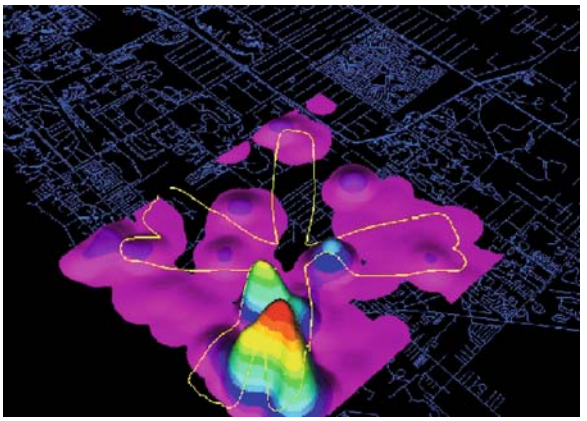


FIGURE 7 Complaint density pseudo-terrain as shaded false elevation near Naples Municipal Airport, with 95 dB maximum A-weighted aircraft noise contour-line superimposed.

residential densities in these areas. The contour draped over the complaint density pseudo-terrain that encompasses the bulk of the high ground is the 95 dB maximum A-level contour. (The airport's 65 dB DNL contour closes much nearer to the end of the runway.)

Figure 7 was constructed by geo-coding (assigning latitude/longitude coordinates) to each of the complainants' home addresses, and then representing complaints per unit time per unit area as elevation. The geographic correspondence between the 95 dB maximum A-weighted aircraft sound level contour and the complaint concentrations suggests that at this airport, it is the maximum level of individual aircraft noise levels, rather than the integrated exposure created by the entire fleet of aircraft operating at the airport, that is most closely associated with complaints.

5 • Summary

Although community reaction to environmental noise has been intensively studied for half a century, much remains to be learned about the origins and mechanisms by which environmental noise generates annoyance. Interpretive criteria for tolerable levels of community noise remain ad hoc in nature and are unsupported by systematic, theory-based understanding. They therefore tend to reflect the charters and political goals of regulatory agencies as much as objective information about the acceptability of noise exposure to communities.

Acknowledgment Portions of the text of this chapter paraphrase the text of Fidell (2003) and Fidell and Silvati (2004).

References

American National Standards Institute (ANSI). (2000) "Quantities and Procedures for Description and Measurement of Environmental Sound – Part 6: Methods for Estimation of Awakenings Associated with Aircraft Noise Events Heard in Homes", ANSI S12.9-2000/Part 6.

Federal Interagency Committee on Noise (FICON). (1992). Final Report: Airport Noise Assessment Methodologies and Metrics, Washington, DC.

Fields, J., (1993), "Effects of personal and situational variables on noise annoyance in residential areas," J. Acoust. Soc. Am. 93(5), 2753–2763.

Fidell, S., (1999). "Assessment of the effectiveness of aircraft noise regulation," Noise & Health 3, 17–25, ISSN 1463–1741.

Fidell, S. (2003). "The Schultz curve 25 years later: a research perspective", J. Acoust. Soc. Am. 114(6), 3007–3015.

Fidell, S., Horonjeff, R., and Green, D. M. (1981). "Statistical analyses of urban noise", Noise Control Eng., 16(2), 75–80.

Fidell, S., Horonjeff, R., Mills, J., Baldwin, E., Teffeteller, S., and Pearsons, K., (1985). "Aircraft noise annoyance at three joint air carrier and general aviation airports," J. Acoust. Soc. Am. 77, 1054–1068.

Fidell, S., and Howe, R. (1998). "Use of Airport Noise Complaint Files to Improve Understanding of Community Response to Aircraft Noise", NASA Contractor Report 207650, NASA Langley Research Center, Hampton, Virginia.

Fidell, S., Howe, R., Tabachnick, B. Pearsons, K., and Sneddon, M. (1995b). "Noise-Induced Sleep Disturbance in Residences Near Two Civil Airports", NASA Langley Research Center Contract No. NAS1-20101, NASA Contractor Report 198252, Hampton, VA.

Fidell, S., and Pearsons, K. (2003). "Sensitivity to prospective transportation noise exposure", Noise control. Eng. J. 51(2), 106–113.

Fidell, S., Pearsons, K., Howe, R., Tabachnick, B., Silvati, L., and Barber, D. (1995). "Field study of noise-induced sleep disturbance", J. Acoust. Soc. Am. 98(2), 1025–1033.

Fidell, S., Pearsons, K., Silvati, L., and Sneddon, M. (2002). "Relationship between low-frequency aircraft noise and annoyance due to rattle and vibration", J. Acoust. Soc. Am. 111(4), 1743–1750.

Fidell, S., Pearsons, K., Tabachnick, B. G., and Howe, R. (2000). "Effects on sleep disturbance of changes in aircraft noise near three airports", J. Acoust. Soc. Am. 107(5), Pt. 1, 2535–2547.

Fidell, S., Pearsons, K., Tabachnick, B., Howe, R., Silvati, L., and Barber, D. (1995a). "Field study of noise-induced sleep disturbance", J. Acoust. Soc. Am. 98, 1025—1033.

Fidell, S., Schultz, T. J., and Green, D. (1988). "A theoretical interpretation of the prevalence rate of noise-induced annoyance in residential populations", J. Acoust. Soc. Am. 84(6), 2109–2113.

Fidell, S., and Silvati, L. (2004). "Parsimonious alternatives to regression analysis for characterizing prevalence rates of aircraft noise annoyance", Noise Control Eng. J. 52(2), March–April, 2004, 56–68.

Fidell, S., Silvati, L., and Haboly, E., (2002). "Social survey of community response to a step change in aircraft noise exposure," J. Acoust. Soc. Am. 111(1), Pt. 1, 200–209.

Fidell, S., Silvati, L., and Pearsons, K., (1998). "Noticeability of a decrease in aircraft noise," Noise Control Eng. J. 46(2), 49–56.

Fidell, S., Silvati, L., Pearsons, K., Lind, S., and Howe, R. (1999). "Field study of the annoyance of low-frequency runway sideline noise", J. Acoust. Soc. Am. 106(3) Pt. 1, 1408–1415.

Fidell, S., Silvati, L., Tabachnick, B., Howe, R., Pearsons, K. S., Knopf, R. C., Gramann, J., and Buchanan, T. (1996). "Effects of aircraft overflights on wilderness recreationists", J. Acoust. Soc. Am. 100(5), 2909–2918.

Galloway, W. J., Eldred, K. McK., and Simpson, M. A. (1974). "Population Distribution of the United States as a Function of Outdoor Noise Level," U.S. Environmental Protection Agency, Report EPA-550/9-74-009-A, Washington, DC.

Government Accounting Office (GAO). (2000). "Aviation and the Environment: Results from a Survey of the Nation's 50 Busiest Commercial Service Airports," GAO/RCED-00-222, Washington, DC, pp. 35–36.

Green, D., and S. Fidell (1991). "Variability in the criterion for reporting annoyance in community noise surveys", J. Acoust. Soc. Am. 89(1), 234–243.

Horonjeff, R., and Anderson, G. (2004). "Queuing For Quiet – The Natural Soundscape Microstructure from a Visitor Perspective", 148th Meeting of the Acoustical Society of America, San Diego, CA.

Job, R. F. S. (1988). "Community response to noise: A review of factors influencing the relationship between noise exposure and reaction", J. Acoust. Soc. Am. 83(3), 991–1001.

Miedema, H., Fields, J., and Vos, J. (2005). "Effect of season and meteorological conditions on community noise annoyance", J. Acoust. Soc. Am. 117(5), 2853–2865.

Miedema, H., and Vos, J. (1998). "Exposure-response relationships for transportation noise", J. Acoust. Soc. Am. 104(6), 3432–3445.

Pearsons, K., Barber, D., Tabachnick, B., and Fidell, S. (1995). "Predicting noise-induced sleep disturbance", J. Acoust. Soc. Am. 97(1), 331–338.

Rosenblith, W., Stevens, K. N., and the staff of Bolt Beranek and Newman (1953). Handbook of Acoustic Noise Control, Volume II, "Noise and Man," WADC Technical Report 52-204, Wright Air Development Center, Ohio.

Schomer, P. (2002). "On normalizing DNL to provide better correlation with response," Sound and Vibration, 14–23.

Schomer, P. (2005) "Annoyance criteria", Noise Control Eng. J. 53(4), 125–135.

Schultz, T. J. (1978). "Synthesis of social surveys on noise annoyance", J. Acoust. Soc. Am. 64, 377–405.

Sneddon, M., Pearsons, K., and Fidell, S. (2004). "Laboratory study of the noticeability and annoyance of sounds of low signal-to-noise ratio", Noise Control Eng. J. 51(5), 300–305.

Swets, J. (1964), "Signal Detection and Recognition by Human Observers", John Wiley and Sons, Inc., New York, NY.

PART IX
ACTIVE NOISE CONTROL

Stephen J. Elliott

Institute of Sound and Vibration Research, Southampton, UK

Introduction
Stephen J. Elliott 1023

55 Forward Control in Single-Channel Systems
Stephen J. Elliott 1029
Introduction • Control of Deterministic Disturbances • Adaptive FIR Controllers • Equivalent Feedback Controller

56 Forward Control in Multichannel Systems
Stephen J. Elliott 1057
Introduction • Optimal Control of Tonal Disturbances • Steepest Descent Algorithms • Trade-Off Between Error Attenuation and Control Effort: The Leaky Algorithm • Robustness to Plant Uncertainties and Plant Model Errors • Applications: Controlling Propeller Noise in Aircraft

57 Feedback Control Systems
Stephen J. Elliott 1081
Introduction • Analogue Controllers • Application: Active Headsets • Digital Controllers • Hybrid Feedback Controllers

Introduction

Stephen J. Elliott

Institute of Sound and Vibration Research, Southampton, UK

Active sound control involves the manipulation of the sound field due to an original "primary" source, using interference with another sound field generated by a number of controllable "secondary" sources. In most applications, the level of the sound field needs to be reduced, although in order to improve the sound quality by manipulating the spectral balance in a sound field, some recent systems have also included the ability to amplify the sound at some frequencies.

Active sound control requires the spatial matching of the primary and secondary sound fields, and the temporal matching of their waveforms. It is the spatial matching that sets the fundamental physical limits to active control and generally restricts its use to low-frequency applications, for which the acoustic wavelengths are relatively long. In the active control of sound in ducts and the control of interior noise in cars or propeller aircraft (some current industrial applications), active control is only useful in attenuating sound below a few hundred hertz. This is because the physical dimensions of the duct or enclosures become large compared with the acoustic wavelength above these frequencies. If, however, active control is only required over a small spatial region, such as in an active headset, the working range can be extended to about 1 kHz. The physical principles that determine these frequency limitations are discussed in several books (Nelson and Elliott, 1992; Hansen and Snyder, 1997).

Chapters 55–57 concentrate on the problem of temporal matching between the primary and secondary sound fields, which must be achieved by processing the electrical signals used to drive the secondary sources (Kuo and Morgan, 1996; Clark, Saunders and Gibbs, 1997; Elliott, 2001). The general problem addressed by active control is illustrated on the left-hand side of Figure 1,

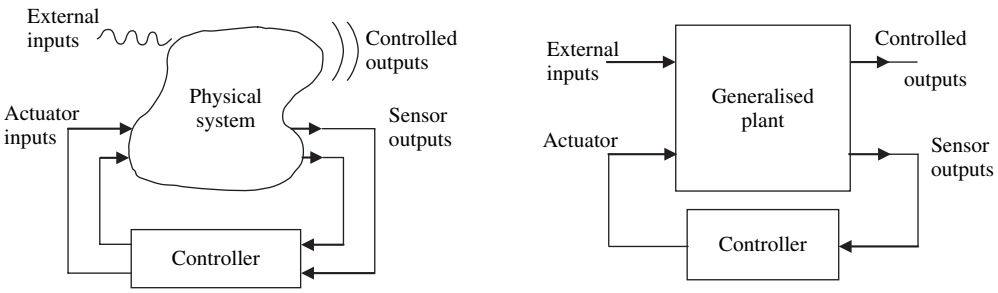


FIGURE 1 The general active sound control problem is shown on the left-hand side, in which an external input acts on a physical system to produce an output, which can also be controlled by the action of a number of actuators acting on the system, driven by the outputs of an electrical controller fed from a number of sensor signals. The general block diagram of such a control system is shown on the right-hand side in which the various responses of the physical system under control are represented as elements of the multi-input, multi-output generalised plant.

in which a physical system is subject to an external excitation and so generates some response, which in this case would be sound. In order to control this output, the electrical signals from a number of sensors are processed by a controller to produce the output signals fed to a number of actuators. The general block diagram of such a system is shown on the right-hand side of Figure 1, in which the response of the physical system, i.e. the controlled and sensor outputs, to the external and actuator inputs are represented as elements of a multi-input generalised "plant".

There are two main strategies for active control, which can both be expressed in the form of this general block diagram. If an external "reference" signal is available, which gives time-advance information about the sound to be cancelled, then a feedforward control strategy can be used. An example of such an application is the control of sound in a duct as illustrated on the upper part of Figure 2, where the sound wave, propagating from left to right, is detected using a microphone well before it reaches the secondary actuator, which is a loudspeaker in this case. This type of active sound control system was proposed by Paul Lueg in 1936. The objective is to reduce the residual sound propagating in the duct and this can be quantified by the output of an "error" sensor placed further down the duct. In general, the actuator output may also affect the reference sensor signal, but this can be internally accounted for in the controller,

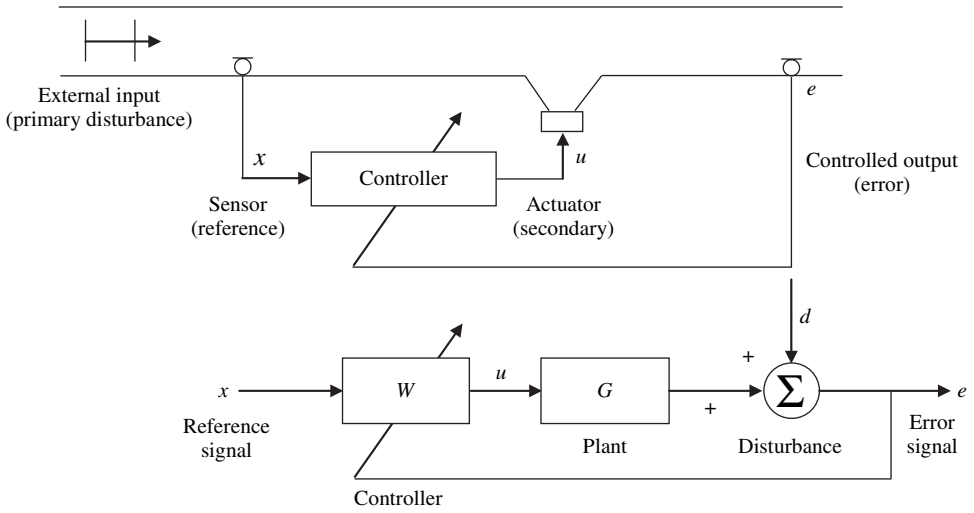


FIGURE 2 Physical arrangement of a feedforward control system used to reduce noise propagating from left to right in a duct (*upper*), and the equivalent electrical block diagram (*lower*).

so that the general block diagram reduces to that shown on the lower part of Figure 2, in which the single-input, single-output plant response now denotes the behaviour of the electroacoustic system that includes the loudspeaker, the acoustic response of the duct and the microphone. If the feedforward controller is fixed, it cannot compensate for changes in the disturbance or the plant and is thus "open loop". In most active control systems, the output of the error sensor is used to adapt the controller in order to maintain performance in the face of disturbance and plant changes, and then becomes "closed loop".

The development of adaptive feedforward controllers has been strongly influenced by the similar problem of electrical noise cancellation, in which adaptive digital filters are used to maintain good performance in the face of a changing environment (Widrow and Stearns, 1985). The signal processing roots of active sound control originate in the algorithms, particularly the LMS algorithm, developed for electrical noise cancellation. Adaptive feedforward control in single-channel systems is discussed in Chapter 55 for both tonal and random disturbances. To achieve active control over a larger region of space, such as in a car or aircraft cabin, however, many actuators and sensors are required. The feedforward control of tonal disturbances in such multi-channel systems is discussed in Chapter 56.

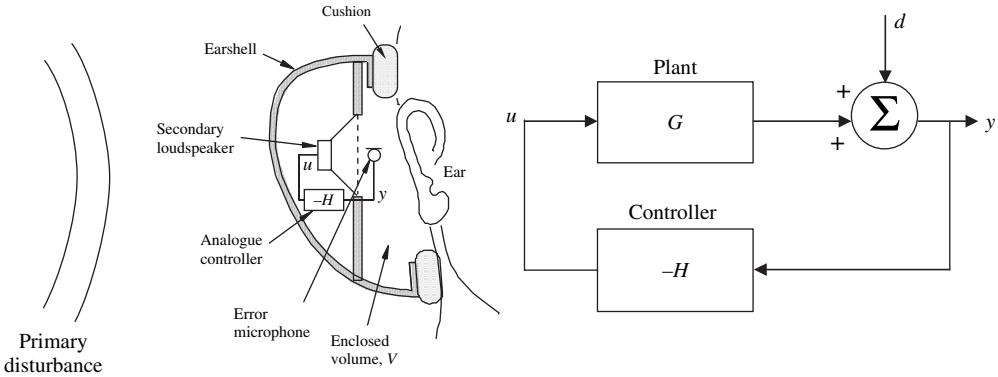


FIGURE 3 Physical arrangement of a feedback control system used to reduce sound in an active headset (*left-hand side*) and the equivalent electrical block diagram (*right-hand side*).

Active sound systems can also use direct feedback controllers, and this principle is employed in the active headset illustrated on the left-hand side of Figure 3. In this application, there are no external reference signals to provide a priori information about the disturbance. The single microphone thus acts as both reference and error sensor. Its output is fed directly back to the loudspeaker through the electrical controller, which is generally analogue and is denoted by $-H$ in Figure 3, to emphasise the fact that negative feedback is being implemented. The right-hand side of Figure 3 shows the equivalent block diagram of such a feedback system, which is analysed in Chapter 57 of this section.

Although the algorithms used for adaptive feedback controllers have their origins in the signal processing community, there has recently been a good deal of activity in using rigorous analysis methods developed by the control communities to understand their behaviour. The control community has also developed a number of ways of analysing feedback systems over the last 50 years, and in some cases the active sound control community has been slow in recognising these insights. It has not, however, always been easy to develop a common understanding between the active sound control community, with its background in signal processing, and the conventional control community. This stems partly from differences in the implicit assumptions made by the different communities and partly from their difference in emphasis. There has recently been more of a convergence of understanding between these two academic communities, however (Elliott and Veres, 2005), and a more mature understanding of the interaction between signal processing and control is now emerging.

References

Clark R.L., Saunders W.R. and Gibbs G.P. (1997) *Adaptive Structures*, John Wiley & Sons, New York.

Elliott S.J. (2001) *Signal Processing for Active Control*, Academic Press, New York.

Elliott S.J. and Veres S. (2005) Guest Editors of a Special Issue of the International Journal of Adaptive Signal Processing and Control on Adaptive Control of Sound and Vibration, vol. 19, no. 2–3, pp. 63–65.

Hansen C.H. and Snyder S.D. (1997) *Active Control of Noise and Vibration*, E&FN Spon, an imprint of Chapman Hall, London.

Kuo S.M. and Morgan D.R. (1996) *Active Noise Control Systems*, John Wiley & Sons, New York.

Nelson P.A. and Elliott S.J. (1992) *Active Control of Sound*, Academic Press, New York.

Widrow B. and Stearns S.D. (1985) *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ.

55

Forward Control in Single-Channel Systems

Stephen J. Elliott

Institute of Sound and Vibration Research, Southampton, UK

1 • Introduction

Feedforward methods have a long and important history in active control. Dating back to Lueg's original duct controller, as discussed in Introduction, in which an upstream reference microphone was used to give advanced information about the waveform of the pressure disturbance propagating down the duct. In this application, there is no restriction on the type of disturbance that can be cancelled, provided the reference sensor is placed far enough upstream to overcome any delays in the processor, so that the optimal controller is causal. Thus, random or transient signals can be completely cancelled with a feedforward controller under noise-free conditions, provided that enough time-advance information is available from the detected signal.

Another class of signal for which feedforward controllers are widely used is that for which the signals are periodic. In this case, the future values of a signal's waveform can, in principle, be perfectly predicted from its waveform in the past. In practice, however, a measured waveform is never perfectly predictable but may be closely synchronized to the process that produces it, for example a rotating or reciprocating machine. In this case, a *reference signal* can generally be obtained from the machine, which can be used instead of that from the reference sensor in a

feedforward controller. This still gives 'advanced' warning of the waveform, although the fact that the waveform is substantially unaltered over several cycles is also being used in the controller.

In this chapter, the control of periodic disturbances is considered first, in both the time and the frequency domains, because this type of disturbance is found in many active control problems. The more complicated case is then described in which a random disturbance is controlled by an FIR controller adapted using the filtered-x LMS algorithm. Finally, the use of adaptive filters for the practical implementation of controllers for tonal disturbances is discussed, and an exact analysis of its response is presented in terms of an equivalent feedback loop.

All the control systems discussed in this chapter will be assumed to be 'digital', i.e. they operate on discrete-time signals. Clearly, the physical signal under control, whether it be the acoustic pressure variation in a duct or the velocity of a structure, is an 'analogue' or continuous-time signal. It is therefore important to establish the relationship between the continuous-time signals present in the physical system under control, and the discrete time signals used by the controller. Although the theory of sampled-data systems is very well developed (Kuo, 1980; Franklin et al., 1990; Åström and Wittenmark, 1997), we will endeavour here to give a physical description of the processes involved and the conditions under which the continuous-time signals can be readily inferred from their discrete-time counterparts.

The block diagram of a linear continuous-time *plant*, subject to a *disturbance* $d_c(t)$, and driven by a discrete-time feedforward controller is shown in Figure 1a. The name 'plant' is widely used in the control literature to denote the physical system under control, which, in the context of active control, is the response between the input to the secondary actuator and the output of the sensor used to measure the residual error signal. The output of this *error sensor* in the absence of active control is termed the disturbance signal and is caused by the influence of the original, primary, source on this sensor. Note that the disturbance signal plays a similar role to the desired signal in the electrical noise-cancelling problem. It is, however, assumed that the controlled output is added to the disturbance in Figure 1, because of the superposition of physical signals which occur in a linear control system, rather than the controlled output being subtracted from the desired signal to form the error in the electrical cancellation problem. The complete discrete-time control system includes not just the digital controller, $W(z)$, but also the data converters and analogue anti-aliasing and reconstruction filters. The transfer function of the continuous-time plant in the Laplace domain is equal to $G_c(s)$, and it is assumed that an analogue

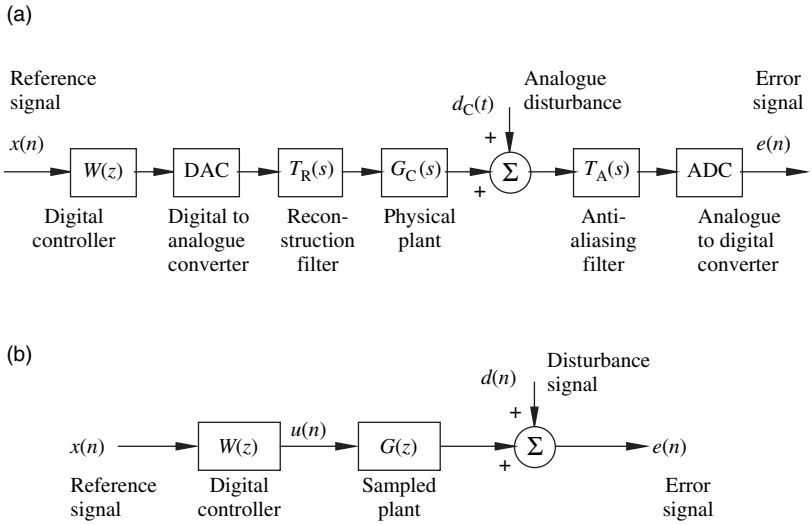


FIGURE 1 Block diagram of a continuous-time 'analogue' plant driven by a sampled-time 'digital' feedforward controller (**a**) and its equivalent sampled-time representation (**b**).

reconstruction filter, with transfer function $T_R(s)$, is used after the digital to analogue converter (DAC). The DAC itself generally has an inherent zero-order hold (see Franklin et al. (1990), for example), whose Laplace-domain transfer function is equal to $T_s(z)$.

We will assume that the combined response of the reconstruction filter and the plant provides sufficient attenuation of the frequency components produced by the DAC above half the sample frequency for the output of the plant to be effectively band-limited to below this frequency. Similarly, we will assume that the anti-aliasing filter, with transfer function $T_A(s)$, removes any components of the continuous time disturbance above half the sample frequency. Under these conditions, the discrete-time error signal is a faithful representation of the continuous-time error signal, and, more importantly, the filtered continuous time error signal behaves smoothly between the discrete sampling time.

Assuming no disturbance, the z domain transfer function from the sampled input $u(n)$ to the sampled output $e(n)$ is the effective plant response seen by the digital controller, which is also known as the pulse-transfer function (Franklin et al., 1990). This transfer function, $G(z)$, corresponds to the plant response 'seen' by the digital controller and includes the response of the continuous-time system under control, the data converters, and the analogue anti-aliasing and reconstruction filters. It is also convenient to

include any processing delay in the controller within this sampled plant response. This delay is typically one full sample for active control systems that operate synchronously and at relatively high sampling rates, but could also be a fraction of a sample delay for systems that operate at a slower sampling rate, and in which the input and output of the controller are not sampled synchronously. The effective plant response seen by the digital controller can be computed by taking the z transform of the sampled impulse response of the continuous-time signals, which is itself given by the inverse Laplace transform of the s-domain transfer function. Thus

$$G(z) = \mathcal{Z}\mathcal{L}^{-1}\left[T_{\mathrm{Z}}(s)T_{\mathrm{R}}(s)G_{\mathrm{c}}(s)T_{\mathrm{A}}(s)\right], \tag{1}$$

where $\mathcal{Z}$ denotes taking the z transform and $\mathcal{L}^{-1}$ denotes taking the inverse Laplace transform. The block diagram of the equivalent discrete-time feedforward controller is shown in Figure 1b, where the discrete-time plant response $G(z)$ includes the responses of all the data converters and analogue filters.

1.1 Feedforward Control

In order to describe most clearly the fundamental approaches to feedforward control, we will restrict ourselves in this chapter to a discussion of digital control systems that have a single reference signal, $x(n)$, whether derived from a reference sensor or a synchronizing device, a single secondary actuator, with input $u(n)$, and a single error sensor, with output $e(n)$. From the block diagram of such a system shown in Figure 1b, the z transform of the signal driving the secondary source, $U(z)$, can be written in terms of the reference signal $X(z)$, and the transfer function of the digital controller, $W(z)$, so that

$$U(z) = W(z)\,X(z). \tag{2}$$

If we assume that the digital controller is of FIR form, with coefficients $w_0 \ldots w_{I-1}$, the discrete-time domain version of the control signal that drives the secondary actuator can be written as

$$u(n) = \sum_{i=0}^{I-1} w_i x(n-i). \tag{3}$$

Note that the current sample of the reference signal $x(n)$ is assumed to be used in calculating the current sample of the control signal and thus the controller is assumed to act instantaneously. Although any practical controller will inevitably have a processing delay, we have assumed that this is contained within the response of the discrete time plant, $G(z)$ in Figure 1b, as

described above, which allows equation (3) to be written in a way that is consistent with the signal-processing literature.

The z transform of the final error signal in Figure 1b, $E(z)$, is the linear superposition of the disturbance, $D(z)$, and the control input $U(z)$, modified by the response of the physical system under control, the discrete-time plant $G(z)$, so that

$$E(z) = D(z) + G(z)\ U(z). \tag{4}$$

The difference between a feedforward control system and an electrical noise cancellation system is the presence of the plant response between the output of the control filter and the observation point for the error signal. The plant response would represent little problem if it were known with perfect precision and was minimum phase, because an inverse plant response could then be included at the output of the control filter, and the net response between the output of the control filter and the error observation point would be unity, as is the case for electrical noise cancellation. In practice, however, the plant response is rarely minimum phase, nor is it known with perfect precision, or even completely constant over time. The non-minimum phase characteristics of the plant arise in active control largely because of the distributed nature of the system under control, which allows the disturbances to propagate as waves. In the simplest case, this gives rise to a pure *delay* in the plant response. Such a delay between the filter output and the error observation point would destabilize the normal LMS algorithm used for electrical noise cancellation. The adaptive filtering problem thus has to be re-examined to take the plant response into account. The algorithms used to adjust the control filter in an adaptive feedforward system are based on the same principles as those used for adaptive digital filters (Widrow and Stearns, 1985), but to account for the presence of the plant they generally require an *internal model* of the plant's response. For the adaptive control algorithm to be stable, this model generally has to represent the plant's response with a certain precision. The response of the plant generally changes with time and/or operating conditions, and so the accuracy with which the internal model must represent the plant's response is an important issue.

2 • Control of Deterministic Disturbances

Periodic disturbances are generated by rotating or reciprocating machines. If the disturbance were perfectly periodic and the plant response did not change, optimal control could be achieved with

a fixed, time-invariant, controller. In practice, most disturbance signals encountered in active control are slowly changing, and thus not perfectly periodic, and adaptive controllers are required to maintain good performance. It is possible to implement adaptive control systems that control such periodic disturbances either in the time domain or in the frequency domain. It will be assumed below that the fundamental frequency of the disturbance is known. This information is normally derived from an external periodic reference signal, generated by a tachometer on a rotating or reciprocating engine for example. Methods of estimating this fundamental frequency from the disturbance signal itself, and thus implementing a control system with no external reference signal, have been discussed by Bodson and Douglas (1997) and Meurers and Veres (1999), for example.

2.1 Waveform Synthesis

In the time domain, the control signal can be generated by passing a periodic reference signal through an FIR digital controller whose impulse response is as long as the fundamental period of the disturbance. Adjusting the coefficients of this filter can then generate any periodic waveform within the sampling constraints of the digital system. A particularly efficient implementation of a periodic controller is achieved if the sampling rate can be arranged to be an exactly integer multiple of the period of the disturbance. The FIR control filter now only has to have as many coefficients as the period of the disturbance divided by the sampling period. Furthermore, if the reference signal is assumed to be a periodic impulse train at the fundamental frequency of the disturbance, the implementation of the controller also becomes extremely efficient (Elliott and Darlington, 1985).

Referring to the block diagram of such a control system shown in Figure 2, the sampled output of the control filter, $u(n)$, is given by the reference signal $x(n)$, filtered by an I-th order FIR controller, so that in general

$$u(n) = \sum_{i=0}^{I-1} w_i x(n-i). \tag{5}$$

In the particular case in which the reference signal is a periodic impulse train, then

$$x(n) = \sum_{k=-\infty}^{\infty} \delta(n-kN), \tag{6}$$

where the signal has N samples per period and $\delta(n)$ is the Kronecker impulse function, which is equal to 1 if $n = 0$ and is

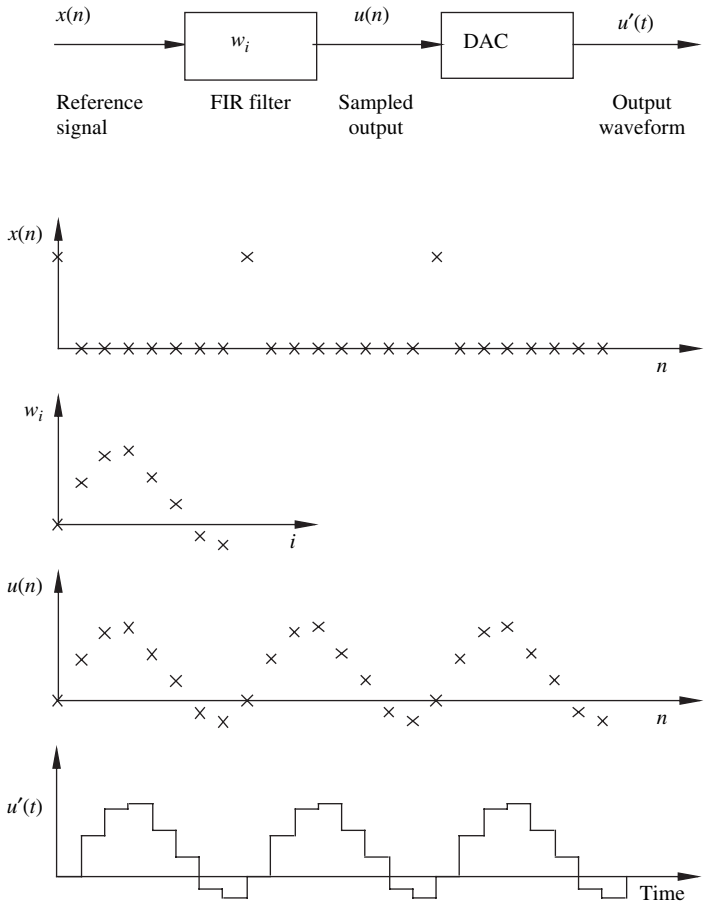


FIGURE 2 Waveform syntheses as FIR digital filtering of a reference signal consisting of a periodic impulse train.

otherwise equal to zero. If the control filter also has N coefficients, its output can be written as

$$u(n) = \sum_{k=-\infty}^{\infty} \sum_{i=0}^{N-1} w_i \delta(n - kN) = w_p, \tag{7}$$

where p denotes the smallest value of $(n - kN)$ for any k and may be interpreted as the 'phase' of $u(n)$ relative to $x(n)$. The output signal is thus a periodic reproduction of the impulse response of the control filter, as illustrated in Figure 4, and the waveform stored in the control filter is effectively 'played out' every cycle after triggering by the impulsive reference signal. The

implementation of the control filter thus requires no computation except the sequential retrieval of the N coefficients w_i. This form of controller was originally investigated by Chaplin (1983) who described it as using *waveform synthesis*. Chaplin originally suggested a 'trial and error' or *power sensing* approach to the adaptation of the coefficients of the control filter in which the mean-square value of the error signal is monitored and the coefficients are individually adjusted in turn until this is minimized. Later, more sophisticated algorithms were suggested for adapting all of the filter coefficients simultaneously using the individual samples of the error sequence, which were described as *waveform sensing* algorithms (Smith and Chaplin, 1983). Some of these algorithms can be shown to be variants of the LMS algorithm, which takes on a particularly simple form when the reference signal is a periodic impulse train (Elliott and Darlington, 1985), and may be viewed as an early form of *iterative learning control* (Owens and Munde, 2000).

2.2 Harmonic Control

A more common method of controlling periodic disturbances is to implement a controller that works in the frequency domain, i.e. operates on each harmonic of the disturbance separately. A sampled periodic disturbance can be represented as the finite summation of harmonics,

$$d(n) = \sum_{k=-K}^{K} D_k e^{\mathrm{j}k\omega_0 Tn}, \tag{8}$$

where k is the harmonic number, D_k is the complex amplitude of the k-th harmonic, the complex amplitude of the harmonics with negative values of k are the conjugates of the positive-numbered ones, i.e., $D_{-k} = D_k^*$, and there are assumed to be a total of K harmonics. The quantity $\omega_0 T$ is 2π times the normalized fundamental frequency, i.e. the true fundamental frequency, $\omega_0/2\pi$ in hertz, divided by the sampling rate, 1/T. If the system under control is linear, the amplitude of the error signal at one harmonic will only depend on the disturbance and the control input at that same harmonic. The control algorithm can then be divided into K *independent* loops, each operating on an individual harmonic, which do not interact in the steady state.

The number of harmonics one chooses to control will depend upon the application. The simplest controller, for a disturbance with a relatively small number of harmonics, may have a separate, tonal, reference signal for each harmonic. An individual control filter may then be used for each harmonic, which only has to have its amplitude and phase adjusted in response to the error

signal at one frequency. This form of controller was proposed by Conover (1956) and in his original control arrangement, analogue circuits are used to generate the reference signals and to change the amplitude and phase of these reference signals. Conover was concerned with the active control of transformer noise, which is most prevalent at the even harmonics of the line frequency, because it is predominantly generated by magnetostriction in the transformer core. He used a full-wave rectifier to double the fundamental frequency and generate a signal rich in harmonics. These harmonics were then individually selected with band-pass filters, to produce tonal reference signals, whose amplitudes and phases could be adjusted before being added together and used to drive the secondary loudspeaker.

Conover originally proposed a manual adjustment of the amplitude and phases of the individual harmonics, but also discusses the need for automatic adaptation and describes how this was difficult to implement using the analogue technology at his disposal. Using digital techniques, it is far easier to automatically adjust the amplitude and phase of a sinusoidal reference signal, and this has led to the rapid development of frequency-domain controllers in recent years. In practice, the real and imaginary parts of the complex control signal, which correspond to the in-phase and quadrature parts of the physical signal, are generally adjusted in a modern system, rather than the amplitude and phase. The complex component of the steady state error signal at the k-th harmonic can be written as

$$E_k = D_k + G_k U_k \tag{9}$$

where D_k was introduced in equation (8), G_k is equal to $G(e^{\mathrm{j}k\omega_0 T})$, which is the complex frequency response of the plant at a normalised angular frequency of $k\omega_0 T$ and U_k is the complex control signal at this frequency. Clearly, the error signal at this harmonic can be driven to zero in this single-channel case if the control input is equal to the optimum value given by

$$U_{k,opt} = -\frac{D_k}{G_k}. \tag{10}$$

Note that because this equation only has to be satisfied at a single frequency, there are no problems with the causality of the controller, which can be implemented using any circuit that gives the correct amplitude and phase values at $k\omega_0 T$, even if the plant response has considerable delay. In practice, however, the plant response and disturbance can never be measured with perfect precision, and an iterative method of adjusting the control

signal is generally used. This has the additional advantage of being able to track non-stationarities in the disturbance, although these must occur on a timescale that is long compared with the period of the disturbance, because otherwise the harmonic series representation of equation (8) would no longer hold.

2.3 Adaptive Harmonic Control

One of the simplest iterative adaptation algorithms is the method of steepest descent. Here we consider changing the real and imaginary parts of the complex control signal in proportion to the negative gradient of the cost function with respect to these control variables. Writing the complex control variable in terms of its real and imaginary parts as

$$U_k = U_{k\mathrm{R}} + jU_{k\mathrm{I}}, \tag{11}$$

the steepest-descent algorithm can be written as the pair of equations

$$U_{k\mathrm{R}}(n+1) = U_{k\mathrm{R}}(n) - \frac{\alpha}{2}\frac{\partial J}{\partial U_{k\mathrm{R}}(n)}, \tag{12a}$$

$$U_{k\mathrm{I}}(n+1) = U_{k\mathrm{I}}(n) - \frac{\alpha}{2}\frac{\partial J}{\partial U_{k\mathrm{I}}(n)}, \tag{12b}$$

where J is the cost function being minimized, n is now the iteration index and α is the convergence coefficient, also known as the adaptation constant, or step size. Alternatively, we can write equations (12) and (13) as a single equation if we define the *complex gradient* to be

$$g_k(n) = \frac{\partial J}{\partial U_{k\mathrm{R}}(n)} + j\frac{\partial J}{\partial U_{k\mathrm{I}(n)}}, \tag{13}$$

in which case the steepest-descent algorithm, equation (12), can be written as

$$U_k(n+1) = U_k(n) - \frac{\alpha}{2}\, g_k(n). \tag{14}$$

Because the cost functions we will consider in this section are restricted to being real variables, we do not need to be concerned about the need to define an analytic gradient of the form discussed by Jolly (1995) or Haykin (1996), for example.

If we now assume a cost function equal to the mean-square value of the error signal, then this is equal to the sum of squared moduli of the error signal at each of the K harmonics,

$$J = \sum_{k=1}^{K} |E_k|^2 . \tag{15}$$

The control input U_k will affect only the k-th harmonic of the error signal in a linear system, and so the complex gradient can be written as

$$g_k(n) = \frac{\partial |E_k|^2}{\partial U_{k\mathrm{R}}(n)} + j\frac{\partial |E_k|^2}{\partial U_{k\mathrm{I}}(n)}. \tag{16}$$

Differentiating the expression for $|E_k|^2$, expressed in terms of the real and imaginary parts of equation (9), with respect to $U_{k\mathrm{R}}$ and $U_{k\mathrm{I}}$, equation (16) can be shown to be equal to

$$g_k(n) = 2G^*{}_k E_k(n), \tag{17}$$

where * denotes conjugation. The steepest-descent algorithm can thus be written as

$$U_k(n+1) = U_k(n) - \alpha G^*{}_k E_k(n). \tag{18}$$

If the adaptation proceeds slowly, the error signal will still be in the steady state at each iteration, so that

$$E_k(n) = D_k + G_k U_k(n), \tag{19}$$

where the disturbance is assumed to be stationary. Using equations (19) and (10), equation (18) can be written as

$$\left[U_k(n+1) - U_{k,\mathrm{opt}}\right] = \left[1 - \alpha |G_k|^2\right]\left[U_k(n) - U_{k,\mathrm{opt}}\right] \tag{20}$$

If $\alpha |G_k|^2$ is considerably smaller than unity, to ensure that the adaptation is slow, then

$$\left(1 - \alpha |G_k|^2\right) \approx e^{-\alpha |G|^2} \tag{21}$$

and so if $U_k(0) = 0$ then

$$U_k(n) \approx \left(1 - e^{-\alpha |G_k|^2 n}\right) U_{k,\mathrm{opt}}. \tag{22}$$

Substituting this into equation (19) yields an expression for the convergence of the error signal as a result of implementing the steepest-descent algorithm, which is given by

$$E_k(n) = D_k e^{-\alpha |G_k|^2 n}, \tag{23}$$

so that the real and imaginary components of the error signal decay to zero exponentially with the same time constant, which is equal to $1/\left(\alpha |G_k|^2\right)$ samples.

In a practical implementation, a perfect model of the plant response is not available and the steepest-descent algorithm becomes

$$U_k(n+1) = U_k(n) - \alpha \hat{G}^*_k E_k(n), \tag{24}$$

where $\hat{G}_k$ is the practical estimate of the plant response, or *plant model* which is used in the control algorithm. Figure 3 shows the block diagram of a practical implementation of this frequency domain adaptive algorithm, in which two quadrature reference signals are used, which correspond to the real and imaginary parts of $e^{jk\omega_o T}$. These are weighted by the real and imaginary parts of the complex control signal U_k, and added to the outputs of any other harmonics being controlled, to produce the control sequence $U(n)$. This is passed through a digital to analogue converter and analogue low pass reconstruction filter to produce the analogue control signal used to drive the loudspeaker. The error signal from the microphone is sampled after passing it through an analogue low pass antialiasing filter and multiplied by the two reference signals to demodulate the complex error signal at this harmonic, E_k. Finally, this complex error signal is

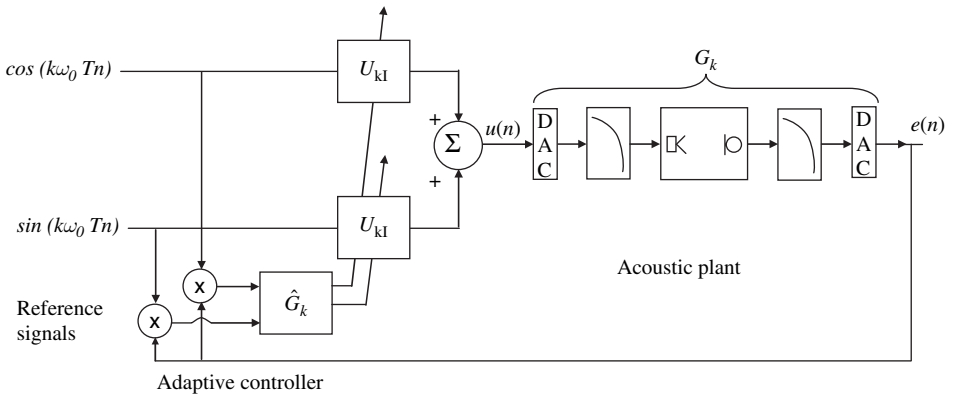


FIGURE 3 Block diagram of the implementation of the k-th harmonic of a frequency-domain adaptive controller for periodic disturbances. The outputs of two quadrature reference signals are weighted by the real and imaginary parts of the complex control signal u_k and fed to the loudspeaker via digital to analogue converted and reconstruction filter. The output from the microphone is passed through an anti-aliasing filter and analogue to digital converted before being multiplied by the two reference signals and averaged to generate the real and imaginary part of the complex error signal, E_k. This is multiplied by the internal model of the plant response, G_k, before being used to adapt the real and imaginary parts of the control signal.

multiplied by the conjugate of the plant model, G_k, and multiplied by α to adapt the real and imaginary parts of U_k according to equation (24). Note that the physical plant response, G_k, comprises the frequency response of all the data converters and analogue filters, as well as that of the loudspeaker and microphone and the acoustic response. This is generally identified before control, to obtain the plant model, either directly in the frequency domain or by identifying an FIR model and taking its Fourier transform. If the fundamental disturbance frequency changes, this must be measured so that an appropriate choice of $\hat{G}_k$ can be made from the frequency response of the plant model.

A more ambitious adaptation algorithm than the method of steepest descent would involve an iterative estimate of the optimum solution, equation (10), which could be written as

$$U_k(n+1) = U_k(n) - \tilde{\alpha}\hat{G}_k^{-1}E_k(n), \tag{25}$$

where $\tilde{\alpha}$ is another convergence coefficient. Notice, however, that the phase of $\hat{G}_k^{-1}$ is the same as that of $\hat{G}^*_k$. The control signal will thus be adjusted in the same direction by both algorithms, and the only difference between them is the amount by which they are adjusted. This will depend on the relative value of the convergence coefficients in the two cases, but it is interesting to note that if

$$\alpha = \frac{\tilde{\alpha}}{\left|\hat{G}_k\right|^2}, \tag{26}$$

then the steepest-descent algorithm, equation (24), is the same as the iterative least squares algorithm (25). In the frequency domain, the method of steepest-descent is thus exactly equivalent to an iterative solution to the least squares problem for single-channel systems.

2.4 Stability Conditions

The conditions for the stability of the adaptive algorithm described by equation (24) can be readily established. Using equation (9) for E_k (at the n-th iteration), we can write equation (24) as

$$U_k(n+1) = U_k(n) - \alpha\left[\hat{G}^*_k D_k + \hat{G}^*_k G_k U_k(n)\right] \tag{27}$$

Equation (10) can now be used to express D_k as $-G_k U_{k,\text{opt}}$, in which case subtracting a factor of $U_{k,\text{opt}}$ from both sides of (27) allows this equation to be written as

$$\left[U_k(n+1) - U_{k,\text{opt}}\right] = \left[1 - \alpha\hat{G}^*_k G_k\right]\left[U_k(n) - U_{k,\text{opt}}\right]. \tag{28}$$

The adaptive algorithm described by equation (24) is thus guaranteed to converge towards the optimum solution provided

$$\left|1-\alpha\hat{G}^{*}_{k}G_{k})\right| < 1. \tag{29}$$

If the real and imaginary parts of $\hat{G}^{*}_{k}G_{k}$ are written as $\mathrm{R}+\mathrm{j}X$, then equation (29) can also be written as

$$|1-\alpha\mathrm{R}-\mathrm{j}\alpha X|^{2} < 1, \tag{30}$$

or

$$(1-\alpha\mathrm{R})^{2}+(\alpha X)^{2} < 1, \tag{31}$$

so that the condition for stability is given by

$$0 < \alpha < \frac{2\mathrm{R}}{\mathrm{R}^{2}+X^{2}}. \tag{32}$$

This is an important result, which will be used on several occasions below. In the single-channel case, considerable physical insight into the stability condition can be derived by assuming that the estimate of the plant's response at $k\omega_{0}T$ can be written as

$$\hat{G}_{k} = MG_{k}e^{\mathrm{j}\phi} \tag{33}$$

where M is a positive real number representing the error in the magnitude, and ϕ represents the error in the phase of the estimated plant response. We can thus write

$$\hat{G}^{*}_{k}G_{k} = M\,|G_{k}|^{2}\,e^{\mathrm{j}\phi} \tag{34}$$

and so the stability condition given by equation (32) can be written as

$$0 < \alpha < \frac{2\cos\phi}{M\,|G_{k}|^{2}}. \tag{35}$$

There are two important points to note about equation (35). First, the stability condition can never be fulfilled if $\cos\phi$ is negative, so the adaptive algorithm will be unstable unless the phase error in the estimated plant response is less than $\pm 90°$, i.e.

$$|\phi| < 90° \text{ for stability.} \tag{36}$$

This important result will arise repeatedly in the following chapters. Second, provided the phase condition on $\hat{G}^{*}_{k}$ is satisfied,

any magnitude error can be compensated for by adjusting the magnitude of the convergence coefficient α. Clearly, the most rapid convergence occurs if there is no phase or magnitude error in $\hat{G}_k^*$ and the convergence coefficient is equal to

$$\alpha_{\text{opt}} = \frac{1}{|G_k|^2}, \tag{37}$$

which corresponds to setting the convergence coefficient in the iterative least squares algorithm of equation (25) to unity. Under these conditions $\left|1 - \alpha \hat{G}_k^* G_k\right|$ is zero and the adaptive algorithm given by equation (24) would converge to the optimal solution in one step. Potential errors in $\hat{G}_k$, however, mean that the convergence coefficient is generally adjusted so that it is a fraction of that indicated in equation (37). With no modelling error, the maximum value of the convergence coefficient for stability is equal to $2\alpha_{\text{opt}}$ according to equation (35). It is interesting to note from equation (28) that, provided the adaptive algorithm is stable, it is guaranteed to converge to the optimum result, regardless of the magnitude and phase errors in the estimate of the plant response in this single-channel, single-frequency case. The same is not true for broadband controllers or for multichannel single-frequency controllers.

3 • Adaptive FIR Controllers

For the control of random signals, controllers are required whose response can be adjusted over the whole range of frequencies in the disturbance, rather than just at a finite number of discrete frequencies as is the case of a harmonic controller. Digital FIR filters are often used for such broadband controllers, and these may be adapted using a modification of the LMS algorithm, as described by Widrow and Stearns (1985), for electrical noise cancellation. The number of coefficients required depends on the extent of the impulse response of the optimal filter or the required frequency selectivity.

This new form of the LMS algorithm can be motivated by assuming that the coefficients of the controller in Figure 4a are changing only slowly with respect to the timescale of the plant dynamics. In this case, we can notionally reverse the order of the controller and plant to produce the equivalent block diagram shown in Figure 4b, in which $r(n)$ is the *filtered reference signal*. Figure 4b is now of exactly the same form as the electrical noise-cancelling case, and the LMS algorithm can be used to adapt its coefficients, provided the reference signal in the equation for

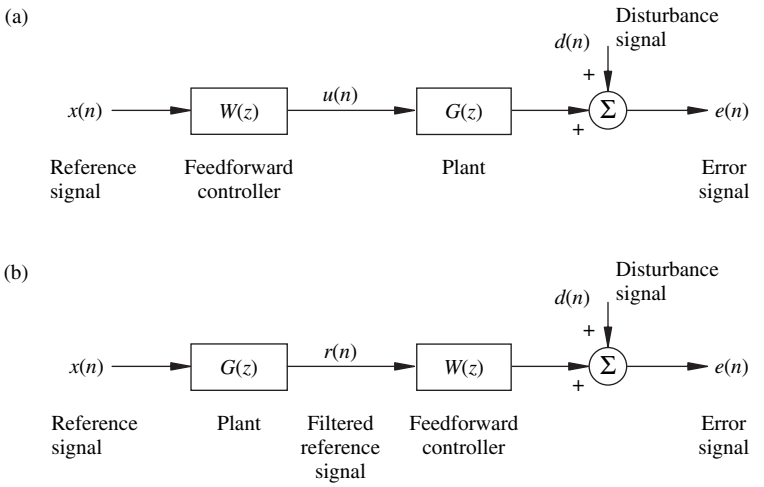


FIGURE 4 Simplified version of the general block diagram for a single-channel feedforward control system assuming perfect feedback cancellation (**a**), and a rearranged version of this for linear and time invariant systems (**b**).

the LMS adaptation is now replaced by the filtered reference signal, whose values can be arranged into a vector $r(n)$. The new algorithm can then be written as

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \alpha \boldsymbol{r}(n)\, e(n), \tag{38}$$

which is known as the *filtered-reference LMS*, or the *filtered-x LMS* (FXLMS), because the reference signal is generally denoted as x. The algorithm was originally proposed by Morgan (1980) and independently for feedforward control by Widrow et al. (1981) and for the active control of sound in ducts by Burgess (1981).

In practice, the filtered reference signal will be generated using an *estimated* version of the true plant response represented as a *plant model*, which can be implemented as a separate real-time filter, $\hat{G}(z)$, to generate the filtered reference signal, $\hat{\boldsymbol{r}}(n)$, as illustrated in Figure 5. The practical version of the filtered-reference LMS algorithm can thus be written as

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \alpha \hat{\boldsymbol{r}}(n) e(n). \tag{39}$$

The need for a separate filter to generate the filtered reference signal in equation (39) clearly increases the computational burden of the filtered-reference LMS algorithm, and so efficient representations of the plant response are very desirable. Fortunately, the impulse response of well-damped acoustic enclosures,

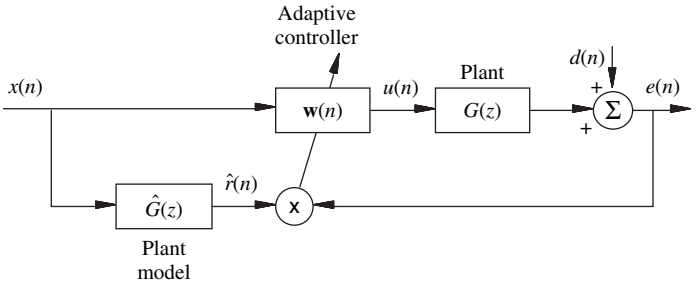


FIGURE 5 Block diagram of the practical implementation of the filtered-reference LMS algorithm, in which a model $\hat{G}(z)$ of the true plant response G(z), is used to generate the filtered reference signal $\hat{r}(n)$.

such as those found in cars and aircraft, do not have a long duration and may be efficiently modelled using FIR digital filters. The LMS algorithm may be used to identify this estimate of the plant response, typically prior to control.

3.1 Conditions for Stability

It is important to establish how well the estimated plant response has to model the true plant in order for the algorithm to behave in a stable and reliable manner. The behaviour of the practical filtered-reference LMS algorithm can be analysed in a similar way to the LMS algorithm in Chapter 55, provided the filter is again assumed to be only slowly adapting. In this case, we can substitute equation (38), for $e(n)$, into equation (39) to obtain

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \alpha\left[\hat{\boldsymbol{r}}(n)d(n) + \hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\boldsymbol{w}(n)\right]. \qquad (40)$$

If the algorithm is stable, it will converge to a solution for which the expectation of the update term in square brackets in equation (40) is equal to zero. Denoting this converged solution as $\boldsymbol{w}_\infty$, we can write

$$\boldsymbol{w}_\infty = -\left[E\left[\hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\right]\right]^{-1} E\left[\hat{\boldsymbol{r}}(n)d(n)\right]. \qquad (41)$$

Thus, even if the practically implemented filtered reference algorithm is stable, it does not generally converge to the truly optimal set of filter coefficients, unless the plant model is perfect, so that $\hat{\boldsymbol{r}}(n) = \boldsymbol{r}(n)$.

Because the controller is only changing slowly we can make the usual independence assumption, as discussed by Widrow and Stearns (1985) for the LMS algorithm. Subtracting a factor of $\boldsymbol{w}_\infty$ from both sides of equation (40) and taking the expectation

value, the average behaviour of the adaptive algorithm can then be written in the form

$$E[\boldsymbol{w}(n+1)-\boldsymbol{w}_\infty]=\left[\mathbf{I}-\alpha E\left[\hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\right]\right]E[\boldsymbol{w}(n)-\boldsymbol{w}_\infty]. \tag{42}$$

The generalized correlation matrix in equation (42) can be decomposed as

$$E\left[\hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\right]=\mathbf{Q}\underline{\Lambda}\mathbf{Q}^{-1} \tag{43}$$

so that equation (42) can be written as

$$\boldsymbol{\nu}(n+1)=[I-\alpha\underline{\Lambda}]\,\boldsymbol{\nu}(n), \tag{44}$$

where the vector of averaged, normalize, and rotated coefficients is equal to

$$\boldsymbol{\nu}(n)=\mathbf{Q}^{-1}E[\boldsymbol{w}(n)-\boldsymbol{w}_\infty]. \tag{45}$$

The eigenvalues that make up the diagonal components of $\underline{\boldsymbol{\Lambda}}$ are not necessarily real, however, and so the magnitude of $1-\alpha\lambda$, which must be less than unity for the algorithm to be stable, is analogous to equation (30) in the single frequency analysis above. The conditions for stability of the filtered reference LMS algorithm are thus, Morgan (1980),

$$0<\alpha<\frac{2Re[\lambda_i]}{\lambda_i^2}\quad\text{for all } i, \tag{46}$$

The stability of the algorithm thus depends on the sign of the real parts of the eigenvalues of the matrix $E\left[\hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\right]$. If the plant model is perfect then this matrix is positive definite and all the eigenvalues are guaranteed to be real and positive. The convergence modes of the algorithm under these conditions, however, are determined by the eigenvalues of the autocorrelation matrix of the filtered reference signals. For the LMS algorithm, the ratio of the maximum to minimum eigenvalue is bounded by the ratios of the maximum and minimum values of the power spectral density of the reference signal (Haykin, 1996). In this case, this condition becomes

$$\frac{\lambda_{\max}}{\lambda_{\min}}\leqslant\frac{\left[\left|G(e^{\mathrm{j}\omega T})\right|^2 S_{xx}(e^{\mathrm{j}\omega T})\right]_{\max}}{\left[\left|G(e^{\mathrm{j}\omega T})\right|^2 S_{xx}(e^{\mathrm{j}\omega T})\right]_{\min}}, \tag{47}$$

so that even if $S_{xx}\left(e^{j\omega T}\right)$ is constant with ωT, a large eigenvalue spread can be generated by differences between the peaks and troughs in the frequency response of the plant.

If the plant model has a different response to that of the physical plant, then the eigenvalues of the matrix $E\left[\hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\right]$ are no longer guaranteed to be entirely real or to have positive real parts. If an eigenvalue is complex then the mode of convergence associated with this eigenvalue will not converge exponentially, on average, but will have an oscillatory component. Provided the real part of the eigenvalue is positive, however, this oscillation will die away and the algorithm will still be stable. If the real part of one of the eigenvalues is negative, then the adaptation will become unstable. It is not straightforward to relate the difference in the response between the plant model and the physical plant to the sign of the real parts of the eigenvalues of the $E\left[\hat{\boldsymbol{r}}(n)\boldsymbol{r}^{\mathrm{T}}(n)\right]$ matrix. However, it has been shown by Ren and Kumar (1989), see also Wang and Ren (1999), that a sufficient, but not necessary, condition for the stability of an algorithm that is similar to the filtered-reference LMS algorithm, is that the ratio $\hat{G}(e^{j\omega T})/G(e^{j\omega T})$ is strictly positive real (SPR). The phase of this quantity must thus lie within $\pm 90°$ at all frequencies. In the current context, the fact that $\hat{G}(e^{j\omega T})/G(e^{j\omega T})$ is strictly positive real implies that the phase of the plant model must be within $\pm 90°$ of the true phase of the plant at all frequencies. This is a neat extension of the phase condition on the plant model derived for the stability of an adaptive controller operating at a single frequency in Section 2.4 above.

In practice, it is possible to implement stable controllers in which the phase condition on the plant model is not satisfied at some frequencies, even though the reference signal is broadband. In general, however, one does not know the spectrum of the reference signal a priori, and so it is possible that the spectrum has significant energy just at the frequency for which the phase of the plant model was in error by more than 90° for the system to become unstable. In an extreme case, the reference signal could become almost tonal at this frequency, in which case the adaptive algorithm would clearly become unstable. Thus, without a detailed knowledge of the statistics of the reference signal, it would appear that 90° phase accuracy at all frequencies provides the only guaranteed condition on the plant model for the stability of a slowly varying controller. The robustness of the algorithm to modelling error can be improved by using a "leaky" form of the adaptation algorithm (Widrow and Stearns, 1985), but a detailed discussion of the effects of such a leak are postponed until the multichannel case has been considered in Chapter 56.

3.2 Performance Degradation Due to Plant-Modelling Errors

It was noted above that when the filtered-reference LMS algorithm is implemented in practice, with an imperfect plant model, then when the controller is stable it converges to equation (41), which is generally not equal to the optimum solution. The value of the residual mean-square error which remains after the algorithm converges is thus necessarily higher than the minimum mean-square error which could be attained if the optimum filter had been implemented. The degree of this degradation in performance with plant modelling error has been investigated in some simulations performed by Saito and Sone (1996).

These authors considered the case of a filtered-reference LMS algorithm applied to a vehicle cabin, in which the impulse response between the input to a loudspeaker and output from a microphone in the cabin lasted for about 150 samples when sampled at a rate of 2 kHz. The performance of the converged filtered-reference LMS algorithm was then calculated for a broadband reference signal when the plant model used in the algorithm was progressively cruder, which was generated by setting to zero the smaller values of the coefficients in the plant's impulse response. In the example considered by Saito and Sone (1996), the maximum attenuation in the disturbance with a perfect plant model was about 9 dB. As the number of non-zero coefficients in the model of the plant's impulse response was reduced, the final attenuation fell, but was not reduced significantly until the plant had only 20 non-zero coefficients. The performance of the algorithm was severely degraded as the number of non-zero coefficients in the impulse response of the plant model was further reduced from 20 to 10, after which the plant model was not accurate enough to maintain the algorithm's stability.

These results suggest that the performance of the filtered-reference LMS algorithm is relatively robust to differences between the response of the physical plant and that of the plant model, but not quite as robust as the algorithm's stability. This observation has practical implications for many applications in which a very efficient implementation of the plant model is required, to reduce the processing requirements of the algorithm, while maintaining an adequate performance from the active control system. The distinction between *robust performance* and *robust stability* will be further discussed for feedback controllers in Chapter 57.

3.3 Effect of Plant Delays

The maximum convergence coefficient that can be used in the filtered-reference LMS algorithm with a perfect plant model can be inferred from the analogy between equation (42) and the corresponding equation for the LMS algorithm (Widrow and Stearns,

1985). This leads to an estimate of the maximum convergence coefficient which is given by

$$\alpha_{\max} \approx \frac{2}{I\overline{r^2}}, \tag{48}$$

where I is the number of coefficients in the controller and $\overline{r^2}$ is the mean-square value of the filtered reference signal. The fact that the plant response is dynamic, however, means that the analysis outlined above is not valid for rapid adaptation and the value of convergence coefficient given by equation (48) is often too large to ensure a stable system. In a practical system the most important part of the plant dynamics is often its overall *delay*, Δ, and simulations with a white noise reference signal (Elliott and Nelson, 1989) suggest that a more realistic upper bound in the convergence coefficient in this case may be

$$\alpha_{\max} \approx \frac{2}{(I+\Delta)\overline{r^2}}, \tag{49}$$

although even smaller values may need to be used if the reference signal is highly coloured or the plant is very resonant. A simplified form of the filtered-reference LMS algorithm, known as the *Delayed LMS* (DLMS), is obtained when there is a pure delay between the filter output signal and the observed error signal, as occurs in various telecommunications application (Quershi and Newhall, 1973; Kabal, 1983; Long et al., 1989). Values of the maximum convergence coefficient which are similar to equation (49) have been obtained analytically in these applications.

The combined effect of delays in the plant, and phase errors in the estimated plant response, on the convergence behaviour of an adaptive controller have been investigated by Boucher et al. (1991). He assumed a tonal reference signal, synchronously sampled at 4 points per cycle. The value of the convergence coefficient required to give the fastest convergence time for various phase errors in the plant model and various pure delays in the plant are shown in Figure 6a. The optimum convergence coefficient is clearly largest for no phase error, as expected from equation (35), and also for smaller plant delays, as expected from equation (49). The minimum convergence time for the algorithm is obtained using this optimum convergence coefficient, and is plotted in Figure 6b which shows that the convergence time is relatively unaffected by phase errors in the plant model of less than about $\pm 40°$, but the convergence time then increases significantly until it tends to infinity as the phase error approaches $\pm 90°$, and the algorithm

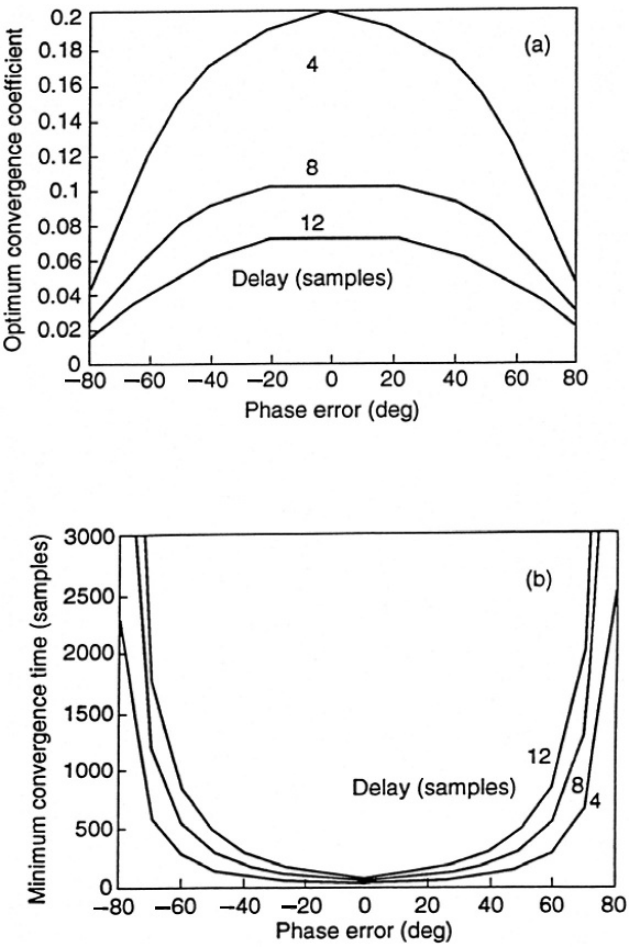


FIGURE 6 The convergence coefficient required to give the fastest convergence time (**a**) and the convergence time resulting from the use of these optimum convergence coefficients (**b**) in a simulation of the filtered-x LMS algorithm with a sinusoidal reference signal having four samples per cycle. Various phase errors were assumed between the estimate of the secondary path $\hat{G}(z)$ and the true secondary path G(z). The three graphs correspond to pure delays in the plant $G(z)$, of 4, 8 and 12 samples.

becomes unstable for any finite value of the convergence coefficient. Snyder and Hansen (1994) have also considered the effect of phase errors for tonal reference signals sampled at different rates and have shown that in general it is difficult to predict what effect the phase error in the plant model will have, particularly for small plant delays, except that the system cannot be guaranteed stable unless the phase error is less than $\pm 90°$.

The filtered-reference LMS algorithm was derived assuming that the coefficients of the adaptive controller change only slowly compared with the timescale of the plant dynamics. In practice it is found that, provided the eigenvalue spread given by equation (47) is not too large and the convergence coefficient is adjusted as in equation (49), then the algorithm will generally converge on a timescale which is similar to that of the plant dynamics. Thus, although some care needs to be taken, the filtered-reference LMS algorithm is found to be surprisingly robust to the rather sweeping assumption of 'slowly varying' controller coefficients.

It is also possible to account for the presence of the plant response in the feedforward control problem by filtering the error signal instead of the reference signal. In this case, the error signal must be filtered by the time-reversed plant response in the single channel case and more generally by the adjoint of the plant response in the multichannel case (Wan, 1993, 1996). In order to implement this algorithm causally, a corresponding delay must be introduced into the reference signal, which tends to reduce the speed of adaptation of the algorithm, because Δ in equation (49) is now determined by the overall duration of the plant's impulse response, rather than just its dominant delays.

4 • Equivalent Feedback Controller

Although the filtered-reference LMS algorithm has been motivated by the need to control stochastic disturbances, it can also provide a very efficient method of adapting a feedforward controller for single frequency disturbances. Such an implementation is shown in Figure 7, in which two quadrature reference signals at the k-th harmonic of a periodic disturbance are multiplied by the single weights w_0 and w_2 to produce the control signal when added together. In this case, both reference signals can be passed through time domain models of the plant response. If the fundamental disturbance frequency varied with time, these plant models would then be able to reproduce the frequency response of the plant over a range of frequencies, and could be implemented as FIR filters. If the disturbance frequency could be measured, however, then the plant model only needs to reproduce the amplitude and phase of the plant response at the specific harmonic frequency under control. Expressing the amplitude and phase in terms of the in phase and quadrature components, it is clear that such a filtered tonal reference signal can be generated for both the adaptive coefficients by suitable weighting of the sine and cosine reference signals and then summing the response. This is represented by the processing of the two reference signals

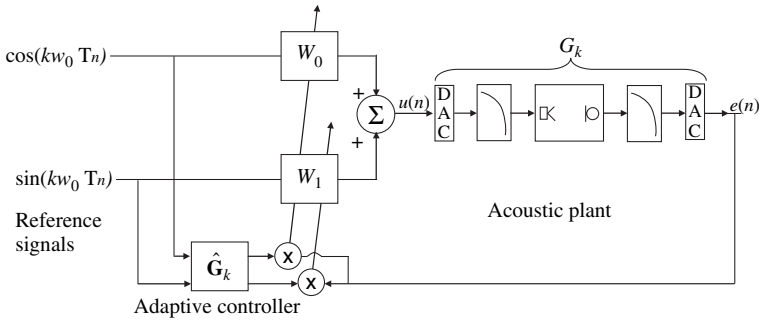


FIGURE 7 Block diagram of the implementation of the k-th harmonic of a time-domain controller for periodic disturbances. Two quadrature reference signals are used, as for the frequency-domain controller of Figure 3, but now instead of the error signal being explicitly demodulated, it is directly multiplied by the filtered version of these reference signals, obtained by multiplying the reference signals by the real and imaginary parts of $\hat{G}\left(e^{jk\omega_o T}\right)$ and summing them together with the 2×2 matrix $\hat{\mathbf{G}}_k$, before adapting the controller weights.

by the 2×2 matrix of real coefficients indicated as $\hat{\mathbf{G}}_k$. Figure 7 thus represents an efficient implementation of the frequency-domain controller shown in Figure 3, because no explicit demodulation of the error signal is required. The multiplication of the filtered reference signal and the error signal is now smoothed by the adaptive algorithm.

Although the operation of this adaptive algorithm can be analysed using the approximate methods described above, it turns out that for this particular form of reference signal another method of analysis, based on an equivalent feedback controller, is available, which is both elegant and exact. This analysis was originally developed for the electrical noise cancellation case with tonal reference signals by Glover (1977), and was then used to analyse feedforward controllers by Elliott et al. (1987). In the single channel case considered here, it can be shown that the relationship between the z transform of the control filter output, $U(z)$, and that of the error signal, $E(z)$, is given *exactly* by the response of a linear time-invariant system with the transfer function

$$-H(z) = \frac{U(z)}{E(z)} = -\alpha A \left[\frac{z \cos(k\omega_o T - \phi) - \cos(\phi)}{1 - 2z \cos(k\omega_o T) + z^2} \right] \qquad (50)$$

where A and ϕ are the magnitude and phase of the plant model at the reference frequency, $\hat{G}\left(e^{jk\omega_o T}\right)$. If the behaviour of the adaptive control algorithm can be exactly represented by

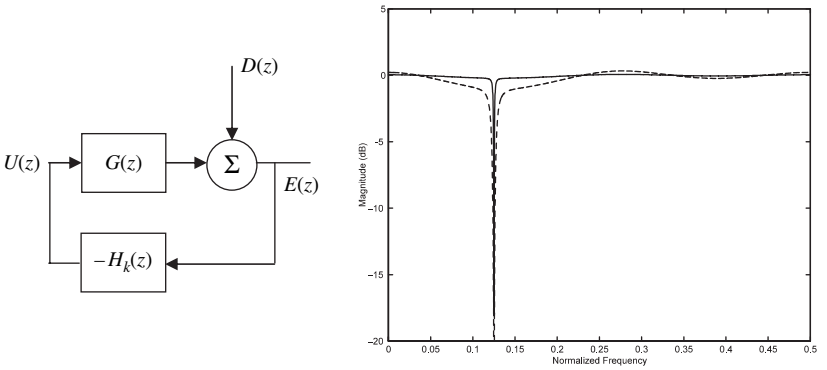


FIGURE 8 Block diagram of the equivalent feedback system that exactly represents the behaviour of the adaptive feedforward tonal controller shown in Figure 7, left-hand side, and the modulus of the frequency response of the complete feedback loop from disturbance signal, $D(z)$, to error signal, $E(z)$ for a simulation in which $\omega_o = \pi/4$, $G = \hat{G} = Z^{-4}$ and $\alpha = 0.01$ (*solid*) and $\alpha = 0.05$ (*dashed*).

equation (50), then the behaviour of the complete interaction between the control algorithm and the plant can be represented by the feedback control system shown on the left-hand side of Figure 8. The overall transfer function of such a system from disturbance input to error output can be written as

$$S(z) = \frac{E(z)}{D(z)} = \frac{1}{1 + G(z)H(z)}, \tag{51}$$

where $S(z)$ is called the sensitivity function in Chapter 57. Using equation (50), this can be explicitly written as

$$S(z) = \frac{1 - 2z\ \cos(k\omega_o)}{1 - 2z\ \cos(k\omega_o \mathrm{T}) + z^2 + \alpha A G(z)\,[z\ \cos(k\omega_o \mathrm{T} - \phi) - \cos\phi]}. \tag{52}$$

The stability and transient properties of the adaptive tonal feedforward controller can thus be analysed exactly by examining the properties of this transfer function. The zeros of the transfer function lie on the unit circle at $z = e^{\pm jk\omega_o \mathrm{T}}$ ensuring perfect steady-state cancellation at this harmonic frequency, if the system is stable. The stability can be analysed by computing the positions of the poles of $S(z)$, and for slow convergence it can be shown that stability is assured if the phase of the plant model is within $\pm 90°$ of that of the plant, as expected from the analysis above. Since

equation (52) exactly represents the response of the adaptive controller, its response to a tonal disturbance at ω_d can be obtained by substituting $z = e^{j\omega_d \mathrm{T}}$ into this expression for $S(z)$. The modulus of the resulting frequency response, for a simulation with $\omega_o = \pi/4$ (8 samples/cycle) and $G = \hat{G} = z^{-4}$, is shown on the right-hand side of Figure 8 for two values of the adaptation coefficient, α. The adaptive feedforward controller clearly acts as a notch filter, with exact cancellation at $\omega_d = k\omega_o$ as expected, but also with a bandwidth in hertz given by Glover (1977)

$$BW\,(Hz) = \frac{\alpha}{\mathrm{T}}, \tag{53}$$

where T is the sample time, and α is the convergence coefficient, which is assumed to be not too close to the value that causes instability. The bandwidth increases with convergence coefficient provided the system remains stable and can thus give some attenuation even if the frequency of the disturbance is not exactly equal to the reference frequency. This is achieved by the weights being adapted so that they vary periodically and so induce a time-dependent phase shift that alters the output frequency. This 'non-Wiener' behaviour is not described by the conventional LMS theory and was one of the motivations for Glover's original work. It remains surprising but true that such time-varying behaviour of an adaptive algorithm can be exactly described by a linear time-invariant feedback loop.

The frequency response shown in Figure 8 also demonstrates that the adaptive algorithm can slightly amplify the disturbance at some frequencies well away from the reference frequency. This 'out of band' behaviour gets more pronounced as the adaptation coefficient is increased towards instability or if there is a significant difference between the response of the physical plant and the plant model (Darlington, 1991; Morgan and Sanford, 1992). Further parallels between the response of harmonic adaptive feedforward controllers and linear resonant feedback controllers have been discussed by Sievers and von Flotow (1992).

References

1. Åström K.J. and Wittenmark B. (1997) *Computer Controlled Systems: Theory and Design*, 3rd Edition, Prentice Hall, Englewood Cliffs, NJ.
2. Bodson M. and Douglas S.C. (1997) Adaptive algorithms for the rejection of sinusoidal disturbances with unknown frequency. *Automatica*, **33**(12), 2213–2221.
3. Boucher C.C., Elliott S.J. and Nelson P.A. (1991) The effect of errors in the plant model on the performance of algorithms for adaptive feedforward control. *Proceedings IEE-F*, **138**, 313–319.

4. Burgess J.C. (1981) Active adaptive sound control in a duct: A computer simulation. *Journal of the Acoustical Society of America*, **70**, 715–726.
5. Chaplin G.B.B. (1983) Anti-sound – The Essex breakthrough. *Chartered Mechanical Engineer*, **30**, 41–47.
6. Conover W.B. (1956) Fighting noise with noise. *Noise Control*, **2**, 78–82.
7. Darlington, 1991 – Only Darlington 1995 in SPAC references (Darlington P. (1995) Performance surfaces of minimum effort estimators and controllers. *IEEE Transactions on Signal Processing*, **43**, 536–539).
8. Elliott S.J. and Darlington P. (1985) Adaptive cancellation of periodic, synchronously sampled interference. *IEEE Transactions on Acoustics Speech and Signal Processing*, **ASSP-33**, 715–717.
9. Elliott S.J. and Nelson P.A. (1989) Multiple point equalisation in a room using adaptive digital filters. *Journal of the Audio Engineering Society*, **37**(11), 899–908.
10. Elliott S.J., Stothers I.M. and Nelson P.A. (1987) A multiple error LMS algorithm and its application to the active control of sound and vibration. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-35**, 1423–1434.
11. Franklin G.F., Powell J.D. and Workman M.L. (1990) *Digital Control of Dynamic Systems*, 2nd Edition, Addison-Wesley, Reading, MA.
12. Glover J.R. (1977) Adaptive noise cancellation applied to sinusoidal noise interferences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-25**, 484–491.
13. Haykin S. (1996) *Adaptive Filter Theory*, 3rd Edition, Prentice Hall, Englewood Cliffs, NJ.
14. Jolly M.R. (1995) On the calculus of complex matrices. *International Journal of Control*, **61**, 749–755.
15. Kabal P. (1983) The stability of adaptive minimum mean-square error equalisers using delayed adjustment. *IEEE Transactions on Communications*, **COM-31**, 430–432.
16. Kuo B.C. (1980) *Digital Control Systems*, Holt, Rinehart and Wilson, New York.
17. Long G., Ling F. and Proakis J.G. (1989) The LMS algorithm with delayed coefficient adaptation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **37**, 1397–1405. See also *IEEE Transactions on Signal Processing*, **SP-40**, 230–232.
18. Meurers A.T. and Veres S.M. (1999) Stability analysis of adaptive FSF-based controller tuning. *Proceedings of ACTIVE99*, Fort Lauderdale, pp. 995–1004.
19. Morgan D.R. (1980) An analysis of multiple correlation cancellation loops with a filter in the auxiliary path. *IEEE Transactions on Acoustics Speech and Signal Processing*, **ASSP-28**, 454–467.
20. Morgan D.R. and Sanford C. (1992) A control theory approach to the stability and transient response of the filtered-*x* LMS adaptive notch filter. *IEEE Transactions on Signal Processing*, **40**, 2341–2346.

21. Owens D.H. and Munde G. (2000) Error convergence in an adaptive iterative learning controller. *International Journal of Control*, **73**, 851–857.
22. Quershi S.K.M. and Newhall E.E. (1973) An adaptive receiver for data transmission of time dispersive channels. *IEEE Transactions on Information Theory*, **IT-19**, 448–459.
23. Ren W. and Kumar P.R. (1989) Adaptive active noise control structures, algorithms and convergence analysis. *Proceeding Internoise 89*, 435–440.
24. Saito N. and Sone T. (1996) Influence of modeling error on noise reduction performance of active noise control systems using filtered-X LMS algorithm. *Journal of the Acoustical Society of Japan*, **17**, 195–202.
25. Sievers L.A. and von Flotow A.H. (1992) Comparison and extension of control methods for narrowband disturbance rejection. *IEEE Transactions on Signal Processing*, **40**, 2377–2391.
26. Smith R.A. and Chaplin G.B.B. (1983) A comparison of some Essex algorithms for major industrial applications. *Proceeding Internoise 83*, **1**, 407–410.
27. Snyder S.D. and Hansen C.M. (1994) The effect of transfer function estimation errors on the filtered-*x* LMS algorithm. *IEEE Transactions on Signal Processing*, **42**, 950–953.
28. Wan E.A. (1993) Finite impulse response neural network with applications in time series prediction. *PhD Thesis*, Stanford University.
29. Wan E.A. (1996) Adjoint LMS: An efficient alternative to the filtered-*x* LMS and multiple error LMS algorithms. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing ICASSP96*, Washington, DC, pp. 1842–1845.
30. Wang A.K. and Ren W. (1999) Convergence analysis of the filtered-u algorithm for active noise control. *Signal Processing*, **73**, 255–266.
31. Widrow B. and Stearns S.D. (1985) *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ.
32. Widrow B., Shur D. and Shaffer S. (1981) On adaptive inverse control, *Proceedings of the 15th ASILOMAR Conference on Circuits, Systems and Computers*, 185–195.

56

Forward Control in Multichannel Systems

Stephen J. Elliott

Institute of Sound and Vibration Research, Southampton, UK

1 • Introduction

Multichannel systems involving the use of multiple actuators and sensors are necessary to control sound over extended regions of space. Such systems are often used to control periodic disturbances, in cars and propeller aircraft for example, that can be broken down into a superposition of sinusoidal components. A control system for periodic disturbances can be implemented using a number of tonal controllers, each of which can operate independently since the system being controlled is very nearly linear. It is the design of these tonal controllers that we will consider in this chapter. The more complicated case of multichannel systems for the control of stochastic disturbances is considered in Elliott (2001). The convergence behaviour of a single-channel LMS controller operating with a broadband reference signal is determined by a number of modes with different time constants that depend on the eigenvalues of the appropriate autocorrelation matrix. The spread of these eigenvalues is related to the *spectral* properties of the filtered reference signal, as seen in equation (47) of Chapter 55, for example. In multichannel systems, however, the reference signal can be a pure tone, and yet the matrix that determines the convergence properties of many algorithms can still have a large spread of eigenvalues, which is now due to the *spatial* distribution of the actuators and sensors.

2 • Optimal Control of Tonal Disturbances

The optimal solution to a control problem will clearly depend on the cost function, or performance index, which one is seeking to minimise. We will begin by considering a cost function proportional to the sum of the mean-square errors, but later we will discuss the significant advantages of incorporating another term, proportional to mean-square control effort, into the cost function. The disturbance is assumed to be tonal, and of normalised frequency $k\omega_o T$, so that provided the plant is linear, control can be achieved by driving the secondary actuators with tonal signals of the same frequency, whose amplitude and phase must be adjusted to minimise the cost function. The amplitude and phase of the L individual disturbance signals and M individual secondary control signals with respect to some arbitrary reference can be described by complex numbers, which we will group together into the vectors $\mathbf{d}\left(e^{jk\omega_o T}\right)$ and $\mathbf{u}\left(e^{jk\omega_o T}\right)$. If the $L \times M$ matrix of complex plant responses at the frequency $k\omega_0 T$ is denoted by $\mathbf{G}\left(e^{jk\omega_o T}\right)$, then the vector of complex residual error signals in the steady state can be written as

$$\mathbf{e}\left(e^{jk\omega_o T}\right) = \mathbf{d}\left(e^{jk\omega_o T}\right) + \mathbf{G}\left(e^{jk\omega_o T}\right)\mathbf{u}\left(e^{jk\omega_o T}\right). \tag{1}$$

For the remainder of this section the explicit dependence of the variables on $k\omega_o T$ will be dropped, so that equation (1) becomes

$$\mathbf{e} = \mathbf{d} + \mathbf{G}\mathbf{u}. \tag{2}$$

The block diagram for such a system is shown in Figure 1. The cost function we will initially consider is the sum of modulus squared error signals,

$$J = \sum_{l=1}^{L} |e_l|^2 = \mathbf{e}^{\mathrm{H}}\mathbf{e}, \tag{3}$$

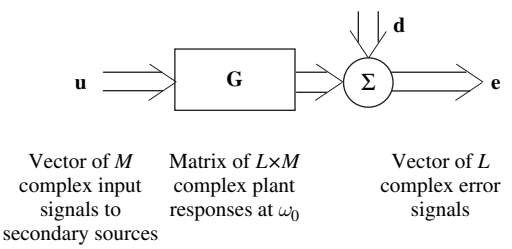


FIGURE 1 Block diagram of a multichannel control system operating at a single frequency of ω_0.

where ^{H} denotes the Hermitian, i.e. complex conjugate, transpose. Substituting equation (2) into (3), we see that the cost function has the Hermitian quadratic form (Nelson and Elliott, 1992):

$$J = \mathbf{u}^{H}\mathbf{A}\mathbf{u} + \mathbf{u}^{H}\mathbf{b} + \mathbf{b}^{H}\mathbf{u} + c, \tag{4}$$

where $\mathbf{A}$ is the Hessian matrix, and in this case

$$\mathbf{A} = \mathbf{G}^{H}\mathbf{G}, \tag{5a}$$

$$\mathbf{b} = \mathbf{G}^{H}\mathbf{d}, \tag{5b}$$

$$c = \mathbf{d}^{H}\mathbf{d}. \tag{5c}$$

The optimal solution we are seeking will depend on the relative numbers of actuators (M) and sensors (L).

2.1 Over-determined Systems

If there are more sensors than actuators ($L > M$), as is usually the case in practice, then the system is described as being *overdetermined* since there are more equations of the form $e_l = d_l + \Sigma G_{lm} u_m$ than there are unknowns, u_m. In general, the smallest value of the cost function that can be obtained by adjusting the elements of $\mathbf{u}$ under these conditions is not zero. We will assume that the matrix $\mathbf{A} = \mathbf{G}^{H}\mathbf{G}$ is positive definite in this case, which will almost always be true in practice, as discussed below. Under these conditions, the set of secondary signals that minimises the cost function of equation (3) can be obtained by setting to zero the derivative of equation (4) with respect to the real and imaginary parts of each component of $\mathbf{u}$, as described in detail in Elliott (2001). The resulting optimal control vector is

$$\mathbf{u}_{\text{opt}} = -\left[\mathbf{G}^{H}\mathbf{G}\right]^{-1}\mathbf{G}^{H}\mathbf{d}, \tag{6}$$

and substituting this expression back into the general cost function gives the expression for its minimum value in this case as

$$J_{\min} = \mathbf{d}^{H}\left[\mathbf{I} - \mathbf{G}\left[\mathbf{G}^{H}\mathbf{G}\right]^{-1}\mathbf{G}^{H}\right]\mathbf{d}. \tag{7}$$

2.1.1 Shape of the Error Surface

The optimal control vector has been obtained, for this ideal case, in a single step, and is thus independent of the shape of the multi-dimensional *error surface* obtained by plotting the cost function, equation (3), as a function of the real and imaginary parts of each element of $\mathbf{u}$. The behaviour of an adaptive algorithm, however, will generally depend very much on the geometric form of the

error surface. The error surface is guaranteed to be quadratic, but its geometric shape will depend strongly on the properties of the matrix $\mathbf{G}^{\mathrm{H}}\mathbf{G}$. To illustrate this, the cost function of equation (4) can also be expressed, using equations (6) and (7), as

$$J = J_{\min} + \left[\mathbf{u} - \mathbf{u}_{\mathrm{opt}}\right]^{\mathrm{H}} \mathbf{A} \left[\mathbf{u} - \mathbf{u}_{\mathrm{opt}}\right]. \tag{8}$$

The matrix $\mathbf{A} = \mathbf{G}^{\mathrm{H}}\mathbf{G}$ is assumed to be positive definite and so has an eigenvalue/eigenvector decomposition of the form

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}}, \tag{9}$$

where $\mathbf{Q}$ is the unitary matrix of normalised eigenvectors, so that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{I}$, and $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues. Substituting equation (9) into (8) and defining the *principal co-ordinates* for the error surface as the translated and rotated set of secondary signals

$$\boldsymbol{\nu} = \mathbf{Q}^{\mathrm{H}} \left[\mathbf{u} - \mathbf{u}_{\mathrm{opt}}\right], \tag{10}$$

the cost function may be written as

$$J = J_{\min} + \boldsymbol{\nu}^{\mathrm{H}} \boldsymbol{\Lambda}\, \boldsymbol{\nu} = J_{\min} + \sum_{m=1}^{M} \lambda_m |\nu_m|^2, \tag{11}$$

where ν_m is the mth element of $\boldsymbol{\nu}$ and λ_m is the mth diagonal element of $\boldsymbol{\Lambda}$, i.e. the mth eigenvalue of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$. Equation (11) clearly shows that the value of the cost function increases quadratically away from the minimum point and varies independently along each of the principal co-ordinates. The rate of increase of the cost function along each of the principal co-ordinates depends on the corresponding eigenvalue λ_m.

Since the matrix $\mathbf{G}^{\mathrm{H}}\mathbf{G}$ is assumed to be positive definite, all its eigenvalues are real and positive, and so the error surface is quadratic with a unique global minimum. The eigenvalues do not, however, have equal magnitudes and in many practical situations the *eigenvalue spread*, which is the ratio of the largest to the smallest eigenvalue, can be very large and the control system is then described as being *ill-conditioned*. The error surface under these conditions has long, thin valleys as illustrated by the contour plot shown in Figure 2, which was computed for an example with two principal components and an eigenvalue spread of 100. The *shape* of the error surface is determined only by the eigenvalues of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$, i.e. only by the response of

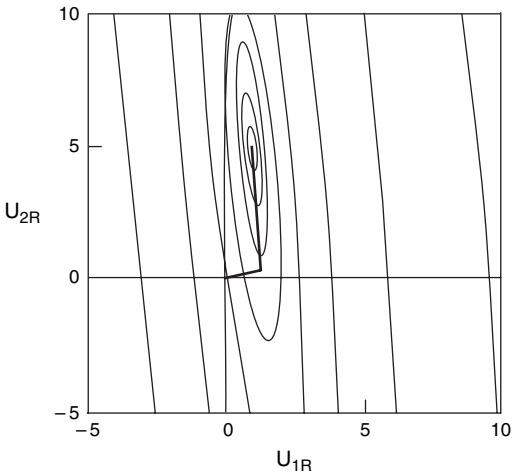


FIGURE 2 Contours of equal cost function, in 5 dB steps, against the real parts of two control signals, u_{1R} and u_{2R}, for an overdetermined control system with an eigenvalue spread of 100.

the plant. The *position* of the error surface with respect to the origin of the co-ordinate system, however, depends on a combination of the disturbance and the plant response, as can be seen from equation (6) for the co-ordinates of the minimum point. The eigenvalue spread of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$, and hence the shape of the error surface, will depend on the positioning of the actuators and sensors and their environment. Generally speaking, the eigenvalue spread will get larger, the more closely grouped the secondary sources or the error sensors become. An extreme case would be when all the actuators or error sensors are positioned at the same point in space, i.e. they are *collocated*, in which case there would be only one non-zero eigenvalue of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$, which would no longer be positive definite. The eigenvalue spread can also become large if there are more secondary sources than significant modes in a lightly damped system, as explained in Nelson and Elliott (1992, Section 12.3). As well as presenting problems for adaptive algorithms, ill-conditioned control problems with large eigenvalue spreads can make the inverse of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$ very sensitive to small perturbations in its elements, which can bring into question the practical usefulness of equation (7) in determining the optimal reductions when the element $\mathbf{G}$ is determined from measurements.

3 • Steepest Descent Algorithms

3.1 Complex Gradient Vector

In this section, we consider control algorithms that iteratively adjust the control actuator's input signals, $\mathbf{u}$, in order to minimise the sum of squared errors, $J = \mathbf{e}^{\mathrm{H}}\mathbf{e}$, using the method of steepest descent. To do this, we need to account for the derivative of the cost function, J, with respect to both the real and the imaginary parts of $\mathbf{u}$, $\mathbf{u}_{\mathrm{R}}$ and $\mathbf{u}_{\mathrm{I}}$, and it is useful to define the *complex gradient vector* as

$$\mathbf{g} = \frac{\partial J}{\partial \mathbf{u}_{\mathrm{R}}} + j\frac{\partial J}{\partial \mathbf{u}_{\mathrm{I}}}, \tag{12}$$

in which $\partial J/\partial \mathbf{u}_{\mathrm{R}}$ and $\partial J/\partial \mathbf{u}_{\mathrm{I}}$ are the vectors of derivatives of J with respect to the elements of $\mathbf{u}_{\mathrm{R}}$ and $\mathbf{u}_{\mathrm{I}}$, which are all entirely real. The steepest descent algorithm for both the real and imaginary parts of the complex vector $\mathbf{u}$ at the $(n+1)$th iteration can therefore be written as

$$\mathbf{u}(n+1) = \mathbf{u}(n) - \frac{\alpha}{2}\mathbf{g}(n), \tag{13}$$

where α is the convergence coefficient and $\mathbf{g}(n)$ is the complex gradient vector at the nth iteration.

It is shown in Nelson and Elliott (1992) that for a Hermitian quadratic function, of the form of equation (4), the complex gradient vector is equal to

$$\mathbf{g} = 2\left[\mathbf{A}\mathbf{u} + \mathbf{b}\right]. \tag{14}$$

Thus, in the case where the Hessian matrix $\mathbf{A} = \mathbf{G}^{\mathrm{H}}\mathbf{G}$, and the vector $\mathbf{b}$ is $\mathbf{G}^{\mathrm{H}}\mathbf{d}$, the complex gradient vector at the nth iteration can be written as

$$\mathbf{g}(n) = 2\left(\mathbf{G}^{\mathrm{H}}\mathbf{G}\mathbf{u}(n) + \mathbf{G}^{\mathrm{H}}\mathbf{d}\right). \tag{15}$$

The vector of optimal control signals in this overdetermined case, equation (6), is obtained by setting this vector of complex gradients to zero.

Assuming that the error signals have time to reach their steady-state value at each iteration, the vector of complex error signals at the nth iteration can be written from equation (2) as

$$\mathbf{e}(n) = \mathbf{d} + \mathbf{G}\,\mathbf{u}(n), \tag{16}$$

so that

$$\mathbf{g}(n) = 2\,\mathbf{G}^{\mathrm{H}}\,\mathbf{e}(n), \tag{17}$$

and the steepest descent algorithm that minimises the sum of the squared error signals can be written as

$$\mathbf{u}(n+1) = \mathbf{u}(n) - \alpha \mathbf{G}^{\mathrm{H}}\, \mathbf{e}(n), \tag{18}$$

The path of the two control parameters in the example that gave rise to the performance surface shown in Figure 2 is also shown in the same figure. It can be seen that the algorithm initially converges in the direction of the first principle axis, and then progresses more slowly along the other principle axis to the minimum value. This behaviour will be quantified below but has a clear geometric interpretation in terms of the steepness of the error surface in the principle axes.

In a practical implementation of this control algorithm, the true plant response $\mathbf{G}$ would not necessarily be perfectly known and an estimate of this quantity, $\hat{\mathbf{G}}$, must be used in the adaptation algorithm. We will postpone a discussion of the effects of this practical complication on the control algorithm until the next section, however, and concentrate here on the stability and convergence properties of equation (18). In a manner exactly analogous to the analysis of the steepest descent algorithm for adaptive digital filters given by Widrow and Stearns (1985), the control signal vector can be transformed into its principal components, defined by equation (10), to derive the stability conditions and convergence properties of the control vector (Elliott et al., 1992). The singular value decomposition (SVD) of the plant response can then be used to establish the effect of the control algorithm on the error signals. In this section, we will use the SVD at the beginning of the analysis, which allows the response of the system, equation (16), and the adaptation algorithm, equation (18), to be written in terms of a set of transformed control signals and error signals. The iterative behaviour of both these signals can then be expressed in a simple and intuitively appealing form.

3.2 Transformation into the Principal Co-ordinates

The SVD of the $L \times M$ matrix of complex plant responses is defined as

$$\mathbf{G} = \mathbf{R}\, \mathbf{\Sigma}\, \mathbf{Q}^{\mathrm{H}}, \tag{19}$$

where $\mathbf{R}$ is the $L \times L$ unitary matrix of complex eigenvectors of $\mathbf{G}\,\mathbf{G}^{\mathrm{H}}$, also known as the left singular vectors of $\mathbf{G}$, so that $\mathbf{R}^{\mathrm{H}}\mathbf{R} = \mathbf{R}\,\mathbf{R}^{\mathrm{H}} = \mathbf{I}$; $\mathbf{Q}$ is the $M \times M$ unitary matrix of complex eigenvectors of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$, also known as the right singular vectors of $\mathbf{G}$, so that

$\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{Q}\,\mathbf{Q}^{\mathrm{H}} = \mathbf{I}$; and the $L \times M$ matrix $\mathbf{\Sigma}$, containing the singular values of $\mathbf{G}$, $\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{M}}$, is defined as

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_M \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{bmatrix}, \tag{20}$$

where the singular values are real and arranged in the order

$$\sigma_1 > \sigma_2 \ldots > \sigma_M. \tag{21}$$

Note that the full form of the SVD must be used here to describe the behaviour of both $\mathbf{u}$ and $\mathbf{e}$, as opposed to the reduced form used by Maciejowski (1989), for example. The SVD of $\mathbf{G}$ can be used to show that

$$\mathbf{G}^{\mathrm{H}}\mathbf{G} = \mathbf{Q}\,\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\,\mathbf{Q}^{\mathrm{H}}, \tag{22}$$

where $\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}$ is a square diagonal matrix with elements σ_m^2. Comparing this expression with equation (9), we can see that the eigenvalues of the matrix $\mathbf{G}^{\mathrm{H}}\mathbf{G}$ are equal to the squared values of the singular values of $\mathbf{G}$, i.e.

$$\lambda_m = \sigma_m^2. \tag{23}$$

The 16 squared singular values of the transfer response matrix measured from 16 loudspeakers to 32 microphones at 88 Hz in a $6\,\mathrm{m} \times 2\,\mathrm{m} \times 2\,\mathrm{m}$ enclosure (Elliott et al., 1992) were found to have a range of about 10^5. This large dynamic range reflects the fact that there are more loudspeakers than significantly excited acoustic modes in the enclosure at this frequency.

Substituting equation (19) into (16), pre-multiplying by $\mathbf{R}^{\mathrm{H}}$ and using the unitary properties of this matrix yields

$$\mathbf{R}^{\mathrm{H}}\mathbf{e}(n) = \mathbf{R}^{\mathrm{H}}\mathbf{d} + \mathbf{\Sigma}\,\mathbf{Q}^{\mathrm{H}}\,\mathbf{u}(n). \tag{24}$$

We now define the transformed set of error and disturbance signals as

$$\mathbf{y}(n) = \mathbf{R}^{\mathrm{H}}\mathbf{e}(n), \tag{25}$$

and

$$\mathbf{p} = \mathbf{R}^{\mathrm{H}}\mathbf{d}, \tag{26}$$

and the transformed control vector as

$$\mathbf{v}(n) = \mathbf{Q}^{\mathrm{H}}\mathbf{u}(n). \tag{27}$$

Equation (16), describing the response of the physical system under control, can now be written in a particularly simple form in terms of these transformed variables as

$$\mathbf{y}(n) = \mathbf{p} + \mathbf{\Sigma}\mathbf{v}(n). \tag{28}$$

The diagonal nature of the $\mathbf{\Sigma}$ matrix in equation (20) means that each of the transformed error signals in equation (28) is either a function of only a single transformed control signal, so that

$$y_l(n) = p_l + \sigma_l\, v_l(n) \quad \text{for} \quad 1 \le l \le M, \tag{29}$$

or is completely unaffected by the control signals, so that

$$y_l(n) = p_l \quad \text{for} \quad M+1 \le l \le L. \tag{30}$$

The lth singular value thus quantifies how well the lth transformed control signal couples into the lth transformed error signal, and so the singular values are also called the *principal gains*, or principal values, of the multichannel system (Maciejowski, 1989). The use of the SVD transforms equation (16), for the physical error signals in the overdetermined case, into a form in which the output signals that are in the rank space of **G**, and can thus be controlled, are clearly divided into those that are in the null space of **G** and thus cannot be controlled.

The first M elements of the transformed error vector can thus be controlled independently by the transformed control signals, and driven to zero if these signals are equal to

$$v_{m:\mathrm{opt}} = -p_m / \sigma_m. \tag{31}$$

A matrix form of this expression can be obtained by substituting equation (19) into equation (6) for the optimum control vector and rearranging to give

$$\mathbf{v}_{\mathrm{opt}} = -\left[\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right]^{-1}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{p}, \tag{32}$$

Using the fact that $\mathbf{R}$ is unitary, the sum of squared errors in the cost function can be written as

$$J = \mathbf{e}^{\mathrm{H}}\mathbf{e} = \mathbf{y}^{\mathrm{H}}\mathbf{y} = \sum_{m=1}^{L} |y_m|^2 . \tag{33}$$

Substituting equation (19) into equation (7) gives the matrix form for the minimum cost function in terms of the transformed primary field,

$$J_{\min} = \mathbf{p}^{\mathrm{H}} \left[I - \mathbf{\Sigma} \left[\mathbf{\Sigma}^{\mathrm{T}} \mathbf{\Sigma} \right]^{-1} \Sigma^{\mathrm{T}} \right] \mathbf{p}, \tag{34}$$

where $\left[I - \mathbf{\Sigma} \left[\mathbf{\Sigma}^{\mathrm{T}} \mathbf{\Sigma} \right]^{-1} \mathbf{\Sigma}^{\mathrm{T}} \right]$ is an $L \times L$ diagonal matrix having the first M diagonal elements equal to zero and the remaining $L-M$ diagonal elements equal to unity. The minimum error can thus be expressed as

$$J_{\min} = \sum_{l=M+1}^{L} |p_l|^2 , \tag{35}$$

and since the cost function without any control can be written as $\sum_{l=1}^{L} |p_l|^2$, the normalised reduction in the cost function can be readily calculated. In general, the cost function will be equal to

$$J = J_{\min} + \sum_{m=1}^{M} |y_m(n)|^2 . \tag{36}$$

3.3 Convergence in the Principal Co-Ordinates

Having expressed the optimal control vector and minimum cost function in terms of the transformed disturbance vector and the singular values of the plant, we can return to the analysis of the steepest descent algorithm, equation (18). Using the SVD given by equation (19) in this equation, the algorithm can be expressed in terms of the transformed variables as

$$\mathbf{v}(n+1) = \mathbf{v}(n) - \alpha \mathbf{\Sigma}^{T} \mathbf{y}(n), \tag{37}$$

and because of the properties of $\mathbf{\Sigma}$, each element of $\mathbf{v}$ will only be affected by the corresponding element in $\mathbf{y}(n)$, so that

$$v_m(n+1) = v_m(n) - \alpha \, \sigma_m \, y_m(n) \quad \text{for} \quad 1 \le m \le M. \tag{38}$$

The elements of both the physical system, equation (28), and the control algorithm, equation (37), are thus uncoupled when

expressed in terms of the transformed variables, and so the transformed elements of the control signal converge independently of one another.

Substituting equation (29) for $y_m(n)$ into equation (38) and using the definition of $v_{m:\text{opt}}$ in equation (31), we obtain

$$\left(v_m(n+1) - v_{m:\text{opt}}\right) = \left[1 - \alpha\sigma_m^2\right]\left(v_m(n) - v_{m:\text{opt}}\right), \tag{39}$$

so that using the fact that $\sigma_m^2 = \lambda_m$, the eigenvalues of the matrix $\mathbf{G}^{\text{H}}\mathbf{G}$, we can write

$$\left(v_m(n) - v_{m:\text{opt}}\right) = \left[1 - \alpha\lambda_m\right]^n \left(v_m(0) - v_{m:\text{opt}}\right). \tag{40}$$

The steepest descent algorithm will converge to the optimum result, provided

$$|1 - \alpha\lambda_m| < 1 \quad \text{for} \quad 1 \le m \le M, \tag{41}$$

so that

$$0 < \alpha < \frac{2}{\lambda_m} \quad \text{for all } \lambda_m. \tag{42}$$

The convergence coefficient must thus be positive and its magnitude is limited to

$$\alpha < \frac{2}{\lambda_{\max}}, \tag{43}$$

where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{G}^{\text{H}}\mathbf{G}$, although the fastest convergence rate is obtained with $\alpha = 1/\lambda_{\max}$. Some care must be taken with this estimate of the maximum convergence coefficient, however, because it has been derived using equation (16), which inherently assumes that any transients have died away in the error signals and they have reached their steady-state sinusoidal waveforms before being measured. If the time between each iteration of the steepest descent algorithm is not long compared with the transient response time of the system under control, then high values of the convergence coefficient will invalidate the assumptions inherent in equation (16), and equation (43) may give an optimistically high maximum value for the convergence coefficient.

If $\alpha\lambda_m \ll 1$ and $v_m(0) = 0$, then equation (40) can be written as

$$v_m(n) \approx \left(1 - e^{-\alpha\lambda_m n}\right) v_{m:\text{opt}}, \tag{44}$$

which shows that the transformed control signals converge exponentially towards the optimum solution in the principal coordinates, with time constants determined by $1/(\alpha\lambda_m)$. By substituting $v_m(n)$ in equation (44) into equation (29) for $y_l(n)$ and using equations (31) and (36), the sum of squared errors can also be expressed for $\alpha\lambda_m \ll 1$ as

$$J(n) \approx J_{\min} + \sum_{m=1}^{M} |p_m|^2 e^{-2\alpha\lambda_m n}, \tag{45}$$

and each $y_m(n)$, which converges independently and exponentially, can be referred to as the amplitude of a *mode* of the control system.

The convergence of the individual modes and the overall convergence of the cost function are shown in Figure 3 for a simulation of the steepest descent algorithm operating in the environment discussed above with 16 loudspeakers and 32 microphones at 88 Hz in a small enclosure (Elliott et al., 1992). The time constant of the slowest mode is $1/(\alpha\lambda_{\min})$ samples, where $\lambda_{\min}$ is the smallest eigenvalue of $\mathbf{G}^H\mathbf{G}$. Since for rapid convergence α must be about $1/\lambda_{\max}$, however, the slowest mode of the steepest descent algorithm must have a time constant that is about $\lambda_{\max}/\lambda_{\min}$ samples, i.e. the eigenvalue spread of the matrix $\mathbf{G}^H\mathbf{G}$. This eigenvalue spread can be very large for some

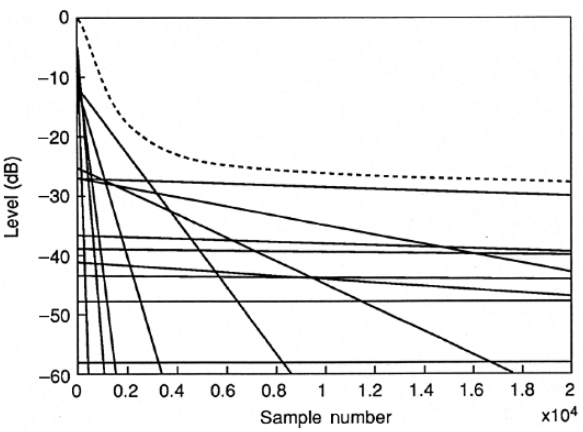


FIGURE 3 The convergence of the sum of squared error signals, normalised by the sum of squared primary disturbances, together with the individual "modes" of convergence for a steepest descent control system operating with 16 loudspeakers and 32 microphones in a small enclosure, after Elliott et al. (1992).

practical multichannel control systems, as in this case where $\lambda_{\max}/\lambda_{\min} \approx 10^5$.

3.4 Convergence of the Control Effort

The control effort required to achieve a given reduction in the error can be written, using the fact that **Q** is unitary, as

$$\mathbf{u}^{\mathrm{H}}\mathbf{u} = \mathbf{v}^{\mathrm{H}}\mathbf{v} = \sum_{m=1}^{\mathrm{M}} |v_m|^2 . \tag{46}$$

Using equations (44) and (31), and the fact that $\sigma_m^2 = \lambda_m$, the control effort at the nth iteration can thus be written for $\alpha\lambda_m \ll 1$ as

$$\mathbf{u}^{\mathrm{H}}(n)\mathbf{u}(n) \approx \sum_{m=1}^{\mathrm{M}} \frac{|p_m|^2}{\lambda_m} \left(1 - e^{-\alpha\lambda_m n}\right)^2 . \tag{47}$$

In controlling the mth component of the transformed disturbance, p_m, the control algorithm thus requires an additional control effort of $|p_m|^2/\lambda_m$. Many active control problems are approximately *well balanced*, in that $|p_m|^2/\lambda_m$ is reasonably constant with m, so that the secondary sources can couple into each of the control modes with a gain that is approximately proportional to the excitation of these modes by the primary source.

Thus, if the control system is well balanced, the same control effort is required to control each of the modes. If the primary source particularly excites some control modes that the secondary sources do not couple into efficiently, however, then $|p_m|^2/\lambda_m$ will be large for these modes, and the system will not be well balanced. This is particularly likely to happen if λ_m is very small, so that the convergence of this mode is also very slow. Also, as the mode converges, the contribution to the control effort from this mode in equation (47) becomes particularly high, which can lead to excessive steady-state actuator driving signals.

The solid line in Figure 4, for example, shows the way in which the control effort changes with the attenuation of the sum of squared error signals during the course of a numerical simulation of the steepest descent algorithm. This was performed using a set of data measured from 16 loudspeakers to 32 microphones in the enclosure discussed above, in which the primary source was arranged such that the system was *not* well balanced. In this case, although 20 dB of attenuation in the mean-square error is obtained with only 0.1 units of effort quite early in the convergence process, over 2 units of effort are then required to achieve a further 4 dB or so of attenuation when the algorithm is

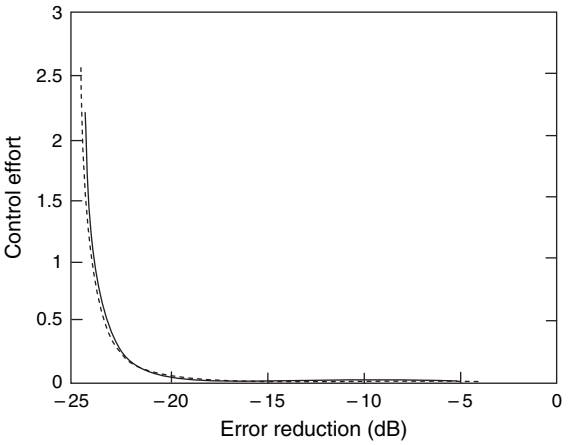


FIGURE 4 The locus of the total control effort with the attenuation in the sum of squared errors for a control system that is not "well balanced" during the convergence of the steepest descent algorithm (*solid line*) and for the steady-state solution calculated for various values of the effort weighting parameter β (*dashed line*).

allowed to converge for some time. The significance of the dashed curve in Figure 4 will be explained in the following section. At a single frequency, the control effort in an active sound control system, in which the electrical impedance of the loudspeakers is relatively insensitive to their acoustic environment, is approximately proportional to the electrical power required to drive the loudspeakers. If the units of control effort in Figure 4 were hundreds of electrical watts, for example, then the additional control effort required for the final 4 dB of attenuation at the microphones is probably not justified, particularly since practical experience has shown that this 4 dB of further reduction at the error microphone positions is typically accompanied by an increase in the sound field away from these positions.

4 • Trade-Off Between Error Attenuation and Control Effort: The Leaky Algorithm

In order to prevent large increases in control effort, a simple modification can be made to the steepest descent algorithm so that it minimises a cost function

$$J = \mathbf{e}^{\mathrm{H}}\mathbf{e} + \beta\, \mathbf{u}^{\mathrm{H}}\mathbf{u}, \tag{48}$$

which includes the sum of squared errors and a component of control effort proportional to the parameter β. The complex gradient vector, equation (48), is now equal to

$$\mathbf{g}(n) = 2\left(\mathbf{G}^{\mathrm{H}}\mathbf{e}(n) + \beta\mathbf{u}(n)\right). \tag{49}$$

The steepest descent algorithm, equation (13), thus becomes

$$\mathbf{u}(n+1) = \gamma\mathbf{u}(n) - \alpha\mathbf{G}^{H}\mathbf{e}(n), \tag{50}$$

where $\gamma = 1 - \alpha\beta$ provides a level of "leakage" in the algorithm.

The effort weighting term in the cost function modifies the equation for the convergence of the cost function, with $\alpha\lambda_m \ll 1$, to (Elliott et al., 1992)

$$J(n) = J_{\min} + \sum_{m=1}^{M} \frac{\sigma_m^2}{\sigma_m^2 + \beta} |p_m|^2 e^{-2\alpha\lambda_m n}. \tag{51}$$

The time constants of convergence are now determined by $1/\lambda_m$, where $\lambda_m = \sigma_m^2 + \beta$ are the eigenvalues of the matrix $\mathbf{G}^{\mathrm{H}}\mathbf{G} + \beta\,\mathbf{I}$, and now

$$J_{\min} = \sum_{m=1}^{M} \frac{\beta}{\sigma_m^2 + \beta} |p_m|^2 + \sum_{m=M+1}^{L} |p_m|^2. \tag{52}$$

The main effect of the effort weighting parameter β is to discriminate against components of the transformed disturbance for which σ_m^2 is small compared with β, in which case two effects occur. First, the term $\sigma_m^2/(\sigma_m^2 + \beta)$ in equation (51) becomes much less than unity for these singular values, and the control system does not attempt to control these modes. Second, the very small values of σ_m^2 that led to very long time constants above now all become approximately equal to β, so there are no longer any very slow modes of convergence. The time constant of the slowest mode is now limited to $\lambda_{\max}/\beta$, assuming $\beta \gg \lambda_{\min}$. The effect of the parameter β on the behaviour of the control effort is even more dramatic, as may be expected, and the control effort at the nth iteration is now given by (Elliott et al., 1992)

$$P(n) = \mathbf{u}(n)^{\mathrm{H}}\mathbf{u}(n) \approx \sum_{m=1}^{M} \frac{\sigma_m^2 |p_m|^2}{(\sigma_m^2 + \beta)^2} \left(1 - e^{-\alpha\lambda_m n}\right)^2, \tag{53}$$

where again $\lambda_m = \sigma_m^2 + \beta$. The steady-state control effort required for modes that have a small value of σ_m^2 compared with β is now proportional to $\sigma_m^2 |p_m|^2/\beta^2$, which is much less than the value of

$|p_m|^2/\sigma_m^2$ given by equation (47) without any effort weighting. A small value of the effort weighting parameter β can thus have a dramatic effect on limiting the control effort, with only a small effect on the residual mean-square error. To illustrate this point, the optimal trade-offs between the attenuation in the sum of square errors and the control effort have been calculated for various values of β in the active noise control example described above, which was not well balanced, and are plotted against each other in the dashed curve of Figure 4. A good practical choice for the value of β would thus be one that ensured that the steady-state solution was at the knee of this curve, giving about 20 dB of attenuation for only about 0.1 units of control effort.

It is interesting to note the similarity between the two curves in Figure 4, even though the solid line has been obtained by plotting the instantaneous values of the attenuation and effort as the steepest descent algorithm evolves in time, whereas the dashed line represents the optimum steady-state values of attenuation and effort as the parameter β is varied. This similarity suggests that at every time during its convergence, the steepest descent algorithm is nearly optimal for a particular value of effort weighting, and this value of the "effective effort weighting" gradually decreases with time. A similar observation has been made by Sjöberg and Ljung (1992) in connection with the training of neural networks. If a finite value of β is included in the steepest descent algorithm, the system will converge in a way very similar to the normal steepest descent algorithm until the value of β in the control algorithm is equal to the gradually decreasing "effective effort weighting", when the convergence will stop. The convergence for the sum of squared errors as a function of time with a finite value of β thus looks very much like the initial part of Figure 3 except that the attenuation no longer continues to increase once the weighted effort term becomes dominant in the cost function.

5 • Robustness to Plant Uncertainties and Plant Model Errors

The steady-state harmonic behaviour of the physical system under control at the nth iteration was described in Section 2 by the equation

$$\mathbf{e}(n) = \mathbf{d} + \mathbf{G}\,\mathbf{u}(n), \tag{54}$$

where $\mathbf{G}$ is the matrix of transfer responses for the physical plant. If an internal *model* of these responses, $\hat{\mathbf{G}}$, is used to adjust the

control vector using the steepest descent algorithm with effort weighting, then equation (50) becomes

$$\mathbf{u}(n+1) = (1-\alpha\beta)\,\mathbf{u}(n) - \alpha\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{e}(n). \tag{55}$$

In the previous section, it was assumed that this plant model was perfect, i.e. $\hat{\mathbf{G}} = \mathbf{G}$. This is clearly a rather sweeping assumption for practical control systems, and one which will be examined in more detail in this section.

There are two important and distinct problems which cause $\hat{\mathbf{G}}$ to be different from $\mathbf{G}$ in practical systems. The first, which has been studied particularly for active control systems (Boucher et al., 1991; Elliott et al., 1992), is when there are modelling errors in $\hat{\mathbf{G}}$, but the plant response itself, $\mathbf{G}$, is assumed to be fixed. The second problem, which is more familiar from the control literature (Morari and Zafiriou, 1989), is when the plant model accurately represents the physical plant in some nominal state, so that $\hat{\mathbf{G}} = \mathbf{G}_0$ say, but the response of the physical plant can then vary about this nominal state. In both cases, we are interested in the stability and performance of the control algorithm. Some progress can be made in this analysis using only equations (54) and (55), but later in this section more specific assumptions will have to be made about the relationship between $\mathbf{G}$ and $\hat{\mathbf{G}}$.

5.1 Convergence Conditions

A general expression for the behaviour of a practical control algorithm is obtained by substituting equation (54) into equation (55) to give

$$\mathbf{u}(n+1) = \mathbf{u}(n) - \alpha\left[\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{d} + \left(\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G} + \beta\mathbf{I}\right)\mathbf{u}(n)\right]. \tag{56}$$

If the control algorithm is stable, it will reach a steady-state solution when $\mathbf{u}(n+1)$ is equal to $\mathbf{u}(n)$. The term in square brackets in equation (56) must be zero under these steady-state conditions and the control vector will thus be equal to

$$\mathbf{u}_\infty = -\left[\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G} + \beta\mathbf{I}\right]^{-1}\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{d}. \tag{57}$$

Note that even if the control algorithm is stable it does not generally converge to the true optimal solution in this case. Only in the special case in which $\beta = 0$ and there are the same number of actuators as sensors, so that $\mathbf{G}$ and $\hat{\mathbf{G}}$ are square and are assumed to be non-singular, does $\mathbf{u}_\infty$ equal the true optimal control vector, which is not a function of $\hat{\mathbf{G}}$.

An iterative mapping of the control vector, offset by the assumed steady-state solution, can now be obtained from equation (57) and equation (56), as

$$(\mathbf{u}(n+1) - \mathbf{u}_\infty) = \left[\mathbf{I} - \alpha\left(\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G} + \beta\mathbf{I}\right)\right](\mathbf{u}(n) - \mathbf{u}_\infty). \tag{58}$$

The convergence of the algorithm is determined by the eigenvalues of the matrix $\left[\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G}+\beta\mathbf{I}\right]$ (Elliott et al., 1992). This matrix is not Hermitian and so, in contrast to the case when $\hat{\mathbf{G}}=\mathbf{G}$, the eigenvalues may be complex and the convergence may have oscillatory as well as exponential modes. Even though the eigenvalues may be complex, the convergence of equation (58) can still be analysed using the principal co-ordinates. If the eigenvalues of $\left[\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G}+\beta\mathbf{I}\right]$ are denoted by λ'_m, then the condition for stability is that

$$|1-\alpha\lambda'_m|<1 \quad \text{for all } m. \tag{59}$$

By direct analogy with the single-channel case described in Section 2.4 of Chapter 55, the convergence condition given by equation (59) can also be written as a bound on the convergence coefficient, α, as

$$0<\alpha<\frac{2\mathrm{Re}(\lambda'_m)}{|\lambda'_m|^2} \quad \text{for all } m, \tag{60}$$

so that even in the limit of slow convergence, $\alpha \ll 1$, the real parts of all the eigenvalues must be positive for the system to converge. The real parts of the small eigenvalues of $\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G}$, λ_m say, can rather easily become negative if the eigenvalue spread is large (Boucher et al., 1991). Since the eigenvalues of $\left[\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G}+\beta\mathbf{I}\right]$ are equal to

$$\lambda'_m=\lambda_m+\beta, \tag{61}$$

it is clear that small value of the effort weighting parameter, β, can counteract the effect of eigenvalues with small negative real parts and ensure that the algorithm remains stable. In general, it is the modes with small values of λ_m that must be stabilised and the effect of a small value of β on the convergence of modes with larger values of λ_m will be small, as demonstrated in equations (51) and (52).

We now consider the form of the eigenvalues for the matrix $\hat{\mathbf{G}}^{\mathrm{H}}\mathbf{G}$ when there is either plant uncertainty or modelling errors. To begin with the case of plant uncertainty, the response to the physical plant, $\mathbf{G}$, is assumed to be equal to that of a *nominal* plant, $\mathbf{G}_0$, plus a matrix of additive plant uncertainties, $\mathbf{\Delta G}_{\mathrm{p}}$, whereas the plant model, $\hat{\mathbf{G}}$, is assumed to be an accurate model of the nominal plant, so that

$$\mathbf{G}=\mathbf{G}_0+\Delta\mathbf{G}_{\mathrm{p}} \text{ and } \hat{\mathbf{G}}=\mathbf{G}_0. \tag{62}$$

In the case of modelling uncertainty, the true plant response is assumed to be constant at its nominal value, $\mathbf{G}_0$, and the plant model is subject to additive modelling errors, $\Delta\mathbf{G}_m$, so that

$$\mathbf{G} = \mathbf{G}_0 \text{ and } \hat{\mathbf{G}} = \mathbf{G}_0 + \Delta\mathbf{G}_m. \tag{63}$$

We have seen that the stability of multichannel steepest descent algorithm depends on the sign of the real parts of the eigenvalues of the matrix $\hat{\mathbf{G}}^H\mathbf{G}$, equation (60). In the case of additive plant uncertainties, equations (62), the real parts of these eigenvalues are

$$\mathrm{Re}\left[\mathrm{eig}\left(\hat{\mathbf{G}}^H\mathbf{G}\right)\right] = \mathrm{Re}\left[\mathrm{eig}\left(\mathbf{G}_0{}^H\mathbf{G}_0 + \mathbf{G}_0{}^H\Delta\mathbf{G}_p\right)\right]. \tag{64}$$

In the case of additive modelling errors, equations (63), the real parts of the eigenvalues are

$$\mathrm{Re}\left[\mathrm{eig}\left(\hat{\mathbf{G}}^H\mathbf{G}\right)\right] = \mathrm{Re}\left[\mathrm{eig}\left(\mathbf{G}^H\hat{\mathbf{G}}\right)\right] = \mathrm{Re}\left[\mathrm{eig}\left(\mathbf{G}_0{}^H\mathbf{G}_0 + \mathbf{G}_0{}^H\Delta\mathbf{G}_m\right)\right], \tag{65}$$

in which we have used the fact that the eigenvalues of $\mathbf{A}^H$ are the conjugates of the eigenvalues of $\mathbf{A}$. The additive matrices of plant uncertainties and modelling errors thus play an identical role in determining the *stability* of a multichannel steepest descent algorithm. The same thing cannot, however, be said of their effect on the *performance* of the algorithm (Omoto and Elliott, 1996, 1999). These authors also discuss the effect of the *structure* of the additive modelling error matrix on the performance of the control system. The structure of this matrix can be seen most clearly if it is decomposed using the SVD of the *nominal* plant matrix, which is given by

$$\mathbf{G}_0 = \mathbf{R}_0\ \mathbf{\Sigma}\ \mathbf{Q}_0{}^H. \tag{66}$$

The eigenvector matrices $\mathbf{R}_0$ and $\mathbf{Q}_0$, obtained for the nominal plant, $\mathbf{G}_0$, can now be used to decompose the matrix of perturbations in the plant, so that

$$\Delta\mathbf{\Sigma} = \mathbf{R}_0{}^H\ \Delta\mathbf{G}_p\ \mathbf{Q}_0, \tag{67}$$

is a matrix of perturbed singular values that, in general, will *not* be diagonal.

Different kinds of perturbation in the plant give rise to different distributions in the magnitudes of the elements of the $\Delta\mathbf{\Sigma}$ matrix, with, for example, changes in the location of objects in the enclosure giving rise mainly to larger elements in the top left-hand quadrant. The performance of the control system turns out

to be most affected by large values of the $\mathbf{\Delta\Sigma}$ matrix in the lower right-hand quadrant, however (Omoto and Elliott, 1999). Thus even though changes in the positions of objects in the enclosure, e.g. people in an aircraft cabin, cause significant changes to the elements of the plant matrix, $\mathbf{G}$, these changes have a *structure* that does not significantly affect the performance of a multichannel control system, which is fortunate.

6 • Applications: Controlling Propeller Noise in Aircraft

In this final section, we describe the application of multichannel feedforward control algorithms to the reduction of tonal components of the sound in the passenger cabins of propeller aircraft. In this case, the amplitude and phase of the tones only change relatively slowly with time during normal cruise conditions. The fact that they change at all is the main reason why the control system must be made adaptive, although there are also changes in the plant response that must be compensated for, due to people moving around in the cabin for example. The control of sound at harmonics of the blade passage frequencies (BPFs) of the rotors inside a helicopter has also been studied, but it is rather more challenging (Boucher et al., 1996; Elliott et al., 1997). One problem is the potential need for large loudspeakers to control the high sound levels at the fundamental frequency of the main rotor, which is typically 10–20 Hz. A more serious problem in this application is the fact that the disturbance can have a considerable short-term amplitude and phase modulation, due to the recirculation of turbulent air round the rotors, which the control system must track. Although active sound control has been demonstrated on helicopters, it is in propeller aircraft that it has progressed most rapidly to commercial implementation.

Many short-haul aircraft, having up to about 50 seats, use propellers instead of jet engines because they can be considerably more efficient at speeds below about 300 mph. The development of prop-fan engines to power larger faster aircraft has also been suggested because of their improved fuel efficiency. The spectrum of the sound pressure in the passenger cabins of such aircraft contain strong tonal components at the BPF of the propellers, which are difficult to attenuate using passive absorption (Wilby et al., 1980; Metzger, 1981). Active control of these tones has been considered since the early 1980s (Chaplin, 1983; Ffowcs-Williams, 1984; Bullmore et al., 1987) and appears to be a good solution to this problem since active sound control works particularly well at low frequencies, where the acoustic wavelengths are

large. Also, with lightweight loudspeakers, active control potentially carries a significantly smaller weight penalty than passive methods of noise control. In this section, we will briefly describe the results of some of the original flight trials of a practical active control system operating in the passenger cabin, in a British Aerospace 748 turboprop aircraft during a series of flight trials undertaken in early 1988 (Elliott et al., 1989, 1990; Dorling et al., 1989) and the subsequent development of this technology.

The 50-seat aircraft used for the trials had a fully trimmed passenger cabin about 9 m long by 2.6 m diameter, and was flown at an altitude of about 10,000 feet under straight and level cruise conditions with the engines running at a nominal 14,200 rpm. As a result of the gearing in the engine and the number of propeller blades, this produces a BPF of 88 Hz. The control system used in the experiments discussed here employed a tachometer on one of the engines to generate reference signals at the BPF and its second and third harmonics, i.e. 88, 176 and 264 Hz. These reference signals were passed through an array of adaptive digital filters, as described in Chapter 55, and their outputs were used to drive 16 loudspeakers. The coefficients of the digital filters were adjusted to implement the steepest descent algorithm described above, in order to minimise the sum of the square values of the 32 microphone signals. Many different configurations of loudspeaker and microphone position were investigated. The results presented here are with the loudspeakers and microphones distributed reasonably uniformly on the floor and on the luggage racks throughout the cabin, although there is a greater concentration at the front of the aircraft since this is close to the plane of the propellers.

The levels of reduction achieved in the sum of the mean-squared pressures at all 32 control microphones, ΔJ (dB), are listed in Table 1 at the three control frequencies. These were achieved by using the control system with all the loudspeakers operating on the pressure field from either the port or starboard propeller alone.

At the second and third harmonics, the attenuations in the sum of squared microphone signals are somewhat smaller than at the BPF. This reflects the greater complexity of the acoustic field within the aircraft cabin at these higher frequencies, which makes the sound field more difficult to control. By moving some of the loudspeakers to a circumferential array in the plane of the propellers, at the front of the aircraft, somewhat greater reductions than those listed in Table 1 could be achieved at the second and third harmonics. This improvement is achieved by the ability of these secondary loudspeakers to more closely match the fuselage vibration caused by the propellers, and thus reduce the

TABLE 1 Changes in the level of the sum of the squared output of 32 control microphones measured during flight trials of a 16-loudspeaker, 32-microphone control system in a British Aerospace 748 aircraft

Configuration	Propeller	88 Hz ΔJ (dB)	176 Hz ΔJ (dB)	264 Hz ΔJ (dB)
Loudspeakers distributed throughout the cabin	Port	−13.8	−7.1	−4.0
	Starboard	−10.9	−4.9	−4.8

Elliott et al. (1990).

overall acoustic excitation of the cabin. Some care must be taken in the interpretation of the results from the error microphones at these higher frequencies, however, since the greater complexity in the sound field makes it more likely that the pressure level has been substantially reduced only at the locations of the error microphones, and that the average pressure level in the rest of the cabin has remained substantially unaltered. More recent flight trials have included a separate array of monitoring microphones, which are not used in the cost function minimised by the active control system, so that an independent measure of attenuation can be achieved (Borchers et al., 1994).

Although the aircraft control system did not have to respond very quickly under straight and level cruise conditions if the speeds of the two propellers were locked together, this was not possible on the aircraft used for the flight trials. The results presented in Table 1 were obtained with only one engine operating at 14,200 rpm and the other detuned to 12,700 rpm, so that its contribution to the sound field was ignored by the control system. As the speeds of the two engines are brought closer together, the sound fields due to the two propellers beat together and the amplitude tracking properties of the adaptive algorithm allow the control system to follow these beats, provided they are not quicker than about two beats per second. Reductions in overall A-weighted sound pressure level of up to 7 dB(A) were measured with the active control system operating at all three harmonics (Elliott et al., 1990). These measurements include a frequency-dependent A-weighting function to account for the response of the ear (Kinsler et al., 1982).

Since this early work, which demonstrated the feasibility of actively controlling the internal propeller noise using loudspeakers, a number of commercial systems have been developed (Emborg and Ross, 1993; Billout et al., 1995) and are now in service on a number of aircraft. Instead of using loudspeakers

as secondary sources, it is also possible to use structural actuators attached to the fuselage to generate the secondary sound field. These have a number of potential advantages over loudspeakers, since they are capable of reducing both the cabin vibration and the sound inside the aircraft. It may also be possible to integrate structural actuators more easily into the aircraft manufacturing process, since no loudspeakers would need to be mounted on the trim panels. Early work in this area using piezoceramic actuators on the aircraft fuselage (Fuller, 1985) demonstrated that relatively few secondary actuators may be needed for efficient acoustic control. More recent systems have used inertial electromagnetic actuators mounted on the aircraft frame (Ross and Purver, 1997) or actively tuned resonant mechanical systems (Fuller et al., 1995; Fuller, 1997) to achieve a more efficient mechanical excitation of the fuselage at the BPF and its harmonics.

References

Billout G., Norris M.A., and Rossetti D.J. (1995) System de controle actif de bruit Lord NVX pour avions d'affaire Beechcraft Kingair, un concept devanu produit. *Proceedings of Active Control Conference*, Cenlis.

Borchers I.U., et al. (1994) Advanced study of active noise control in aircraft (ASANCA), in *Advances in Acoustics Technology*, Ed. J.M.M. Hernandez, John Wiley and Sons, New York, **1**, 80.

Boucher C.C., Elliott S.J. and Nelson P.A. (1991) The effect of errors in the plant model on the performance of algorithms for adaptive feedforward control. *Proceedings IEE-F*, **138**, 313–319.

Boucher C.C., Elliott S.J. and Baek K-H. (1996) Active control of helicopter rotor tones. *Proceedings of Internoise 96*, Liverpool, pp. 1179–1182.

Bullmore A.J., Nelson P.A., Elliott S.J., Evers J.F. and Chidley B. (1987) Models for evaluating the performance of propeller aircraft active noise control systems. *Proceedings of the AIAA 11th Aeroacoustics Conference*, Palo Alto, CA, Paper AIAA-87-2704.

Chaplin G.B.B. (1983) Anti-sound – The Essex breakthrough. *Chartered Mechanical Engineer*, **30**, 41–47.

Dorling C.M., Eatwell G.P., Hutchins S.M., Ross C.F. and Sutcliffe S.G.C. (1989) A demonstration of active noise reduction in an aircraft cabin. *Journal of Sound and Vibration*, **128**, 358–360.

Elliott S.J. (2001) *Signal Processing for Active Control*, Academic Press, New York.

Elliott S.J., Nelson P.A., Stothers I.M. and Boucher C.C. (1989) Preliminary results of in-flight experiments on the active control of propeller-induced cabin noise. *Journal of Sound and Vibration*, **128**, 355–357.

Elliott S.J., Nelson P.A., Stothers I.M. and Boucher C.C. (1990) In-flight experiments on the active control of propeller-induced cabin noise. *Journal of Sound and Vibration*, **140**, 219–238.

Elliott S.J., Boucher C.C. and Nelson P.A. (1992) The behaviour of a multiple channel active control system. *IEEE Transactions on Signal Processing*, **40**, 1041–1052.

Elliott S.J., Boucher C.C. and Sutton T.J. (1997) Active control of rotorcraft interior noise. *Proceedings of the Conference on Innovations in Rotorcraft Technology*, Royal Aeronautical Society, London, 15.1–15.6.

Emborg U. and Ross C.F. (1993) Active control in the SAAB 340. *Proceedings of Recent Advances in the Active Control of Sound and Vibration*, Blacksburg, VA, pp. 567–573.

Ffowcs-Williams J.E. (1984) Review lecture: Anti-sound. *Proceedings of the Royal Society of London*, **A395**, 63–88.

Fuller C.R. (1985) Experiments on reduction of aircraft interior noise using active control of fuselage vibration. *Journal of the Acoustical Society of America*, **78**(S1), S88.

Fuller C.R. (1997) Active control of cabin noise – lessons learned? *Proceedings of Fifth International Congress on Sound and Vibration*, Adelaide.

Fuller C.R., Maillard J.P., Meradal M. and von Flotow A.H. (1995) Control of aircraft interior noise using globally detuned vibration absorbers. *Proceedings of the First Joint CEAS/AIAA Aeroacoustics Conference*, Munich, Germany, Paper CEAS/AIAA-95-082, 615–623.

Kinsler L.E., Frey A.R., Coppens A.B. and Sanders J.V. (1982) *Fundamentals of Acoustics*, 3rd Edition, John Wiley, New York.

Maciejowski J.M. (1989) *Multivariable Feedback Design*. Addison-Wesley, Reading, MA.

Metzger F.B. (1981) Strategies for reducing propeller aircraft cabin noise. *Automotive Engineering*, **89**, 107–113.

Morari M. and Zafiriou E. (1989) *Robust Process Control*, Prentice Hall, Englewood Cliffs, NJ.

Nelson P.A. and Elliott S.J. (1992) *Active Control of Sound*, Academic Press, New York.

Omoto A. and Elliott S.J. (1996) The effect of structured uncertainty in multichannel feedforward control systems. *Proceedings of International Conference on Acoustics, Speech and Signal Processing ICASSP96*, Washington, DC, pp. 965–968.

Omoto A. and Elliott S.J. (1999) The effect of structured uncertainty in the acoustic plant on multichannel feedforward control systems. *IEEE Transactions on Speech and Audio Processing*, **SAP-7**, 204–213.

Ross C.F. and Purver M.R.J. (1997) Active cabin noise control. *Proceedings of ACTIVE 97*, Budapest, Hungary, pp. xxxix–xlvi.

Sjöberg J. and Ljung L. (1992) Overtraining, regularisation and searching for minimum in neural networks. *Proceedings of the Symposium on Adaptive in control and Signal Processing*, Grenoble, 73–78.

Widrow B. and Stearns S.D. (1985) *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ.

Wilby J.F., Rennison D.C., Wilby E.G. and Marsh A.H. (1980) Noise control prediction for high speed propeller-driven aircraft. *Proceedings of American Institute of Aeronautics and Astronautics 6th Aeroacoustic Conference*, Palo Alto, CA Paper AIAA-80-0999.

57

Feedback Control Systems

Stephen J. Elliott

Institute of Sound and Vibration Research, Southampton, UK

1 • Introduction

Feedback control systems are distinguished from feedforward systems in that they have no reference sensor to give time-advanced information about the disturbance being controlled. They are used in a variety of applications for the active control of sound and vibration, particularly where the primary sources that generate the disturbance cannot be directly observed, or in which there are too many primary sources to economically obtain reference signals from each one. The most important example of an active feedback sound control system is in active headsets.

In our discussion of feedback control, we will continue to characterise the plant under control using an input–output approach, rather than using a state variable model. State variable methods for active feedback control have been reviewed, for example, by Franklin et al. (1994), Fuller et al. (1996), Preumont (1997) and Clark et al. (1998), and such methods can provide a complete model of the global response of the system under control. They are particularly applicable to the control of the first few modes of a structure, and are widely used for flight control systems in aircraft, the control of low-frequency vibration in distributed structures (see, for example, Meirovitch (1990) and Rubenstein et al. (1991) and the control of sound radiated by a lightly damped structures (see, for example, Baumann

et al. (1991) and Petitjean and Legrain (1994)). When sound or vibration is controlled over a broad bandwidth with an active system, the plant can be subject to significant delays and be of high order, with a response that is characterised by many well-damped and overlapping modes, whose natural frequencies can change over rather short timescales. A state variable model for such a system may be very complicated to identify, whereas an input–output model can be determined from direct measurement of the plant's frequency response, for example. A frequency response approach also allows some of the design methods and design compromises that are inherent in a feedback control system to be illustrated in a straightforward fashion, and one that is consistent with conventional signal processing formulations.

In the control literature, analogue systems are more accurately referred to as continuous-time systems and digital systems as sampled-time systems, although we will generally retain the less formal titles here, since they are more widely used by the signal processing community. Active systems are generally designed for *disturbance rejection*, and most of the discussion in the following two chapters will concentrate on this application of feedback control. In the conventional control literature, however, more emphasis is generally placed on feedback controllers designed to ensure that the plant output follows some command signal, which are called *servo-control* systems. There is a large and well-developed literature on the design of feedback control systems Franklin et al. (1994), Morari and Zafiriou (1989), Skogestad and Postlethwaite (1996).

The design of analogue feedback controllers is described in Section 2, in which frequency response methods such as those of Bode and Nyquist are particularly emphasised. The application of such "classical" design methods to the problem of feedback control of sound in a headset is then discussed. The complications involved in sampling the plant output and using a digital feedback controller are outlined in Section 3 and, using this sampled-time example, the advantages of one particular architecture of feedback controller are emphasised in Section 4. This architecture uses an internal model of the plant response to transform the feedback control problem into a feedforward control problem, and thus allows the insight and analytical tools developed in the previous sections to be brought to bear on the design of feedback controllers.

1.1 Disturbance Rejection

A physical representation of an active headset using an analogue feedback loop was shown in Figure 3 of Chapter 55, in which the signal from a single microphone is fed back to the secondary loudspeaker, with input signal, via an analogue

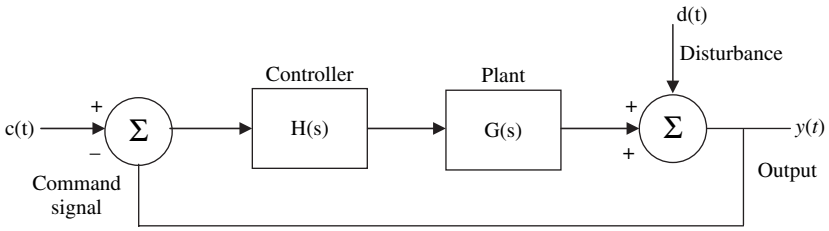


FIGURE 1 Block diagram of a negative feedback control system having both a disturbance, $d(t)$, and a command signal, $c(t)$.

feedback controller with the Laplace-domain transfer function – $H(s)$ as discussed, for example, by Franklin et al. (1994) The phase inversion in the feedback controller reminds us that we are hoping to design a negative feedback system. The equivalent block diagram for this feedback control system was also shown in Figure 3 of Chapter 55 and is redrawn in Figure 1, in which we assume, for the time being, that the command signal is zero. The Laplace-domain transfer function of the plant, from the loudspeaker input to the microphone output, is denoted by $G(s)$, and the output of the microphone in the absence of control is denoted by $d(t)$, which is the disturbance in this case. The Laplace transform of the output signal can be written as

$$Y(s) = D(s) - G(s)H(s)E(s), \tag{1}$$

so that the transfer function from disturbance to output signal is equal to

$$\frac{Y(s)}{D(s)} = \frac{1}{1 + G(s)H(s)} = S(s). \tag{2}$$

This transfer function is also known as the *sensitivity function*, $S(s)$, of the system. Since the objective of this control system is good rejection of the disturbances, we would like the sensitivity function to be small. The conditions required for this to happen are most clearly seen in the frequency domain, and letting $s = j\omega$ in equation (2), the frequency response of the sensitivity function can be written as

$$S(j\omega) = \frac{1}{1 + G(j\omega)H(j\omega)}, \tag{3}$$

where $G(j\omega)$ is the frequency response of the plant and $H(j\omega)$ is the frequency response of the controller, and the feedback system is assumed to be stable.

If the frequency response of the plant were relatively flat and free from phase shift below some frequency ω_0, and the electrical controller were a high-gain amplifier, so that $G(j\omega)H(j\omega)$, the *loop gain*, was large at all frequencies below ω_0, then

$$|1 + G(j\omega)H(j\omega)| \gg 1 \quad , \quad \omega < \omega_0, \tag{4}$$

in which case

$$S(j\omega) \ll 1 \quad , \quad \omega < \omega_0. \tag{5}$$

The action of the control system at these frequencies can be described in simple terms by imagining a small perturbation being introduced as the disturbance in Figure 1. As the perturbation begins to increase, it is immediately amplified and its waveform is inverted before being fed back to the summing junction. The waveform of the perturbation is thus very nearly balanced by the signal fed back through the plant.

The condition described by equation (4) can be closely fulfilled in purely electrical systems for which negative feedback circuits were first invented (Black, 1934). For operational amplifier circuits, such a feedback loop is said to have a *virtual earth*, since the input signal is driven down to earth potential by the feedback loop, and this term is sometimes carried over to describe acoustical feedback systems. In acoustic systems, however, the plant generally exhibits considerably more dynamic response than an electronic amplifier. The electroacoustic response of a moving coil loudspeaker, for example, introduces significant phase shift near its mechanical resonance frequency, which is typically between about 10 and 100 Hz. The acoustic path from loudspeaker to microphone will also inevitably involve some delay, due to the acoustic propagation time, which will introduce an increasing phase shift with frequency. As the net phase shift in the plant approaches 180°, the negative feedback action described above becomes positive feedback and the control system can become unstable.

1.2 Following a Command Signal

Many applications of feedback control are concerned with the manipulation of the plant output to follow some prescribed set point or *command signal*. Such a feedback control system is described above as having a servo action. A command signal can be used to mix a communication channel into the type of active headset described in Section 3. Figure 1 shows how such a command signal can be introduced. If we now assume that the disturbance is zero, the Laplace transform of the output of the system is equal to

$$Y(s) = G(s)H(s)\,[C(s) - Y(s)]\,, \tag{6}$$

so that the transfer function from command signal to plant output is equal to

$$\frac{Y(s)}{C(s)} = \frac{G(s)H(s)}{1+G(s)H(s)} = T(s). \tag{7}$$

This transfer function is called the *complementary sensitivity function*, $T(s)$, because the sum of this and the sensitivity function, equation (2), is equal to unity,

$$S(s)+T(s) = 1. \tag{8}$$

The frequency response of the complementary sensitivity function is equal to

$$T(j\omega) = \frac{G(j\omega)H(j\omega)}{1+G(j\omega)H(j\omega)}. \tag{9}$$

If we again assume that below some frequency, ω_0, the loop gain is large, so that equation (4) again holds, we find that the complementary sensitivity function would be close to unity in this case, i.e. for excitation frequencies below those where the plant dynamics became significant, a high-gain controller can achieve good servo control.

The complementary sensitivity function plays an important role in determining the robustness of the control system to variations in the plant response, as will be explained in more detail in the next section. Such variations have an effect at the output of the plant, which is similar to that of a command signal.

If a command signal and disturbance are both present in Figure 1, the net output will, by superposition, be equal to

$$Y(s) = S(s)D(s)+T(s)C(s). \tag{10}$$

1.3 Bandwidth Limitations Due to a Delay

The bandwidth over which feedback control can be applied is partly determined by the *delay* in the plant response. Assume, for example, that the plant response corresponds to that of a pure delay of τ seconds, so that

$$G(j\omega) = e^{-j\omega\tau}. \tag{11}$$

The best a servo-control system could then achieve would be to track the required signal with the same delay (Skogestad and Postlethwaite, 1996), so that

$$T(j\omega) = e^{-j\omega\tau}. \tag{12}$$

If we now consider the rejection of disturbances in such a control system, which is determined by the sensitivity function $S(j\omega)$ of equation (2), then since $S(j\omega)+T(j\omega)=1$, in this case we have

$$S(j\omega)=1-e^{-j\omega\tau}. \tag{13}$$

The disturbances will be attenuated at all frequencies for which $|S(j\omega)|<1$, which corresponds to

$$\omega\tau<\frac{\pi}{3}. \tag{14}$$

The bandwidth, in hertz, over which the disturbances can be attenuated for a plant with a delay of τ seconds can thus be written as

$$\text{Bandwidth (Hz)}<\frac{1}{6\times\text{Delay (in seconds)}}. \tag{15}$$

A more detailed description of the performance of feedback control systems will be presented in the next section, but equation (15) is a useful "rule of thumb" for calculating the bandwidth of control in many active control applications, in which the plant response has significant delay.

2 • Analogue Controllers

In this section, the stability and design of single-channel continuous-time feedback controllers will be discussed. The *stability* of a feedback system such as that shown in Figure 1 is most easily established in terms of the positions of the poles of the closed-loop transfer function. The poles of both the sensitivity function and complementary sensitivity function are given by the roots of the *characteristic equation*

$$1+G(s)H(s)=0. \tag{16}$$

If the output of the control system is to remain bounded (finite) for any bounded input, then all of the poles of the closed-loop transfer function must be in the open left-hand half of the s plane, i.e. with none on the imaginary axis (see, for example, Franklin et al. (1994)). Strictly speaking, this definition of stability is only valid provided the plant and controller are themselves stable, since otherwise "hidden" instabilities could occur within the control loop, which are not observed at the output. A stronger stability condition is provided by the concept of *internal stability* (see, for example, Morari and Zafiriou (1989)), which is only

satisfied if bounded inputs injected at *any* point in the control system generate bounded responses at any other point. In the discussion of active control systems below, we will thus assume that both the plant and the controller are stable, since this considerably simplifies the analysis. The more general analysis is described, for example, in Morari and Zafiriou (1989).

2.1 Nyquist Stability Criterion

In many applications of active control, an explicit pole/zero model of the plant, which is needed to calculate the roots of equation (16), is not available. This is either because it would be of very high order, or because the plant response involves delays, or because of the considerable variability that can exist in such systems. It is thus more appropriate to continue with the frequency response approach adopted in other parts of this book and consider the *Nyquist* definition of stability for a feedback system (as also described, for example, in Franklin et al. (1994)). This involves an examination of the polar plot of the open-loop frequency response $G(j\omega)H(j\omega)$. If, at some frequency, $G(j\omega)H(j\omega)$ were equal to exactly -1, then it is clear from equation (3) that the response of the feedback control system at this frequency would become unbounded. Since we are assuming that the plant under control and the controller are themselves stable, the definition of stability provided by the Nyquist criterion is that the polar plot of the open-loop frequency response $G(j\omega)H(j\omega)$ must not enclose the *Nyquist point* $(-1,0)$ as ω varies from $-\infty$ to ∞ (Nyquist, 1932), as illustrated in Figure 2. The locus of $G(j\omega)H(j\omega)$ for negative ω is shown as a dashed line in this figure, which is the mirror image of the locus for positive ω about the real axis since

$$G(-j\omega)H(-j\omega) = G^*(j\omega)H^*(j\omega), \tag{17}$$

where * denotes complex conjugation.

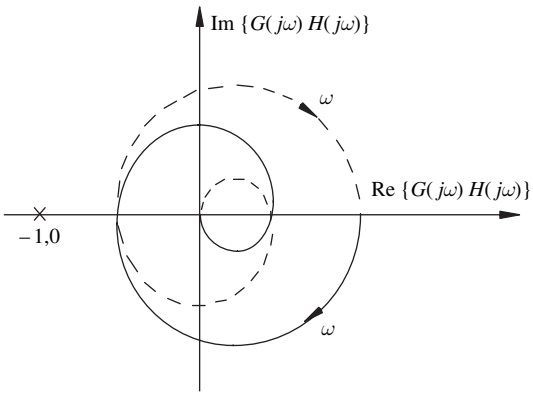


FIGURE 2 The polar plot of a stable open-loop frequency response, for $\omega > 0$ (*solid line*) and $\omega < 0$ (*dashed line*), which does not encircle the Nyquist point $(-1,0)$ and so has a stable closed-loop response.

Note that the magnitude of the frequency response of most physical systems tends to zero at high frequencies (they are said to be *strictly proper* in the control literature), in which case the polar plots of the open-loop frequency response converge on the origin as ω or $-\omega$ tends to ∞. As ω or $-\omega$ tends to zero, the open-loop frequency response also tends to the same, real, value equal to the open-loop "dc gain" of the system. By adopting a bounded-input, bounded-output definition of stability and assuming that both the plant and controller are stable according to this definition, we have excluded the possibility of there being an integrator in the loop, which would cause the response of the system to become progressively larger as the frequency tends to zero. The lack of an integrator would limit the steady-state performance of a servo system, but we are primarily interested here in the rejection of relatively high-frequency disturbances and so such an assumption is not overly restrictive for our purposes. The complete locus in Figure 2 can be deduced from frequency response *measurements* made on the open-loop system, and the stability of the closed-loop system can thus be predicted from these measurements before the loop is closed.

2.2 Gain and Phase Margins

Apart from determining the absolute stability of a feedback system, the Nyquist plot of the open-loop frequency response can also provide useful information about the *relative stability* of such a system, i.e. its stability margins. Two parameters have commonly been used to quantify the relative stability of a feedback control system, and they both give information about the changes that can be allowed in the response of the plant whilst maintaining closed-loop stability. The first of these parameters is called the *gain margin* and is defined to be the increase that can be tolerated in the overall gain of the plant without the feedback loop becoming unstable. The other widely used relative stability parameter is the *phase margin*, defined to be the additional phase shift that could be tolerated in the plant, without the closed-loop system becoming unstable.

One problem with describing the relative stability using the gain and phase margins is that they do not indicate whether the system will be unstable if there is a change in *both* gain and phase. The causes of the changes in the open-loop frequency response that can give rise to instability are generally perturbations in the response of the plant. These perturbations can be due to a multitude of reasons, for example: a change in temperature or operating condition, or a change in the amplitude of the signal driving the plant if the system is weakly non-linear. If the control system remains stable despite such changes, it is described as being *robustly stable*.

2.3 Robust Stability

If the robust stability of a system is to be determined, a model of the perturbations, or *uncertainties*, in the plant response is required. The response of the plant under normal operating conditions, i.e. with no uncertainty, is called the *nominal* plant response. The effect of changes to the plant response on the Nyquist plot of the open-loop frequency response is to introduce regions of uncertainty round the nominal open-loop frequency response at each frequency. The shapes of these regions of uncertainty are determined by the types of change that occur in the plant, but they can always be enclosed by a *disc of uncertainty*. We will be concerned below with designing controllers whose stability is robust to plant uncertainties that are described by such discs, and although the design may be rather conservative if the true uncertainty region is not well described by a disc, a robustly stable controller can always be designed by making this assumption.

Such a disc of uncertainty is generated by an *unstructured* model of uncertainty in which we assume that the true plant response at a given frequency is given by

$$G(j\omega) = G_0(j\omega)\,[1 + \Delta_G(j\omega)]\,, \tag{18}$$

where $G_0(j\omega)$ is the nominal plant response and $\Delta_G(j\omega)$ is the multiplicative plant uncertainty. The uncertainty, $\Delta_G(j\omega)$, is generally complex and with unknown phase, but it is assumed that its magnitude is bounded at each frequency to be less than $B(\omega)$, so that

$$|\Delta_G(j\omega)| \leq B(\omega) \text{ for all } \omega. \tag{19}$$

At a given frequency, all plant responses that are described by equations (18) and (19) lie within a disc in the complex plane whose centre is equal to the nominal plant response, $G_0(j\omega)$, and whose radius is $B(\omega)\,|G_0(j\omega)|$, and the open-loop frequency response of a system with unstructured uncertainty could, at a given frequency, be located anywhere in a disc in the Nyquist plane whose centre is $G_0(j\omega)H(j\omega)$ and whose radius is $B(\omega)\,|G_0(j\omega)H(j\omega)|$, as shown on the left-hand side of Figure 3. For robust stability, the closed-loop system cannot be unstable for any plant whose response satisfies equations (18) and (19), so that the disc of uncertainty must not enclose the Nyquist point. This means that the distance from the centre of the circle to the Nyquist point, $|1 + G_0(j\omega)H(j\omega)|$, must be greater than the radius of the circle, so that

$$|1 + G_0(j\omega)H(j\omega)| > B(\omega)\,|G_0(j\omega)H(j\omega)|\,, \tag{20}$$

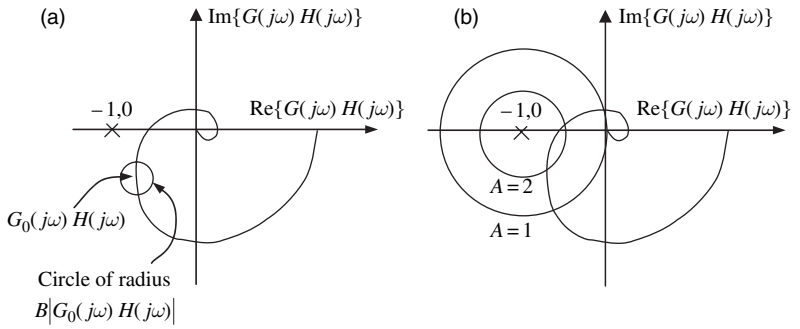


FIGURE 3 The conditions for robust stability and disturbance enhancement can be represented by geometric constraints in the Nyquist plot. (**a**) For robust stability, the uncertainty disc, whose radius is proportional to the upper bound on the multiplicative uncertainty, B, must not encircle the $(-1,0)$ point. (**b**) In order for the system not to enhance the disturbance by more than a factor of A, the Nyquist plot must not pass through a circle drawn about the $(-1,0)$ point having a radius of $1/A$.

and so

$$\frac{|G_0(j\omega)H(j\omega)|}{|1+G_0(j\omega)H(j\omega)|} < \frac{1}{B(\omega)}. \tag{21}$$

The left-hand side of equation (21) is equal to the modulus of the complementary sensitivity function for the nominal plant, $T_0(j\omega)$. The condition for robust stability can thus be written as

$$|T_0(j\omega)| < \frac{1}{B(\omega)} \quad \text{for all } \omega, \tag{22}$$

or

$$|T_0(j\omega)B(\omega)| < 1 \quad \text{for all } \omega. \tag{23}$$

If a control system were required to be robustly stable to changes in the plant response described by a multiplicative uncertainty bounded by $B(\omega) = 0.5$, for example, the control system would have a gain margin of about 3.5 dB and a phase margin of about 30°, as can be confirmed by a simple geometric construction in the Nyquist plane.

The maximum value of $|T_0(j\omega)B(\omega)|$ over frequency (or, strictly speaking, the least upper bound of this function) is known as its *supremum* and the condition for robust stability can also be written as

$$\sup_{\omega} |T_0(j\omega)B(\omega)| = \|T_0B\|_\infty < 1, \tag{24}$$

where $\|T_0B\|_\infty$ denotes the ∞ norm of T_0B, the weighted complementary sensitivity function, and this form of the robust stability condition is widely used in the H_∞ control literature (Doyle et al., 1992; Morari and Zafiriou, 1989; Skogestad and Postlethwaite, 1996). Thus we see that robust stability involves a *constraint* on the complementary sensitivity function, which must be satisfied at each frequency. The symbol H_∞ is used to denote that the norm is defined for the Hardy space of transfer functions that are stable and proper, i.e. their response is finite as $\omega \to \infty$. Later we will also use the H_2 norm of a transfer function, which is proportional to the square root of the integrated modulus squared response over all frequencies, and this is defined for the Hardy space of transfer functions that are stable and strictly proper, i.e. their response tends to zero as $\omega \to \infty$ (see also, for example, Skogestad and Postlethwaite (1996)). In active control, the cost function that we generally want the control system to minimise is proportional to the mean-square value of a function, or the square of its H_2 norm, and the controller that minimises such a cost function is said to be H_2 optimal. It is also possible to design controllers that minimise the H_∞ norm of a function, or the maximum value of the frequency response (Zames, 1981; Doyle et al., 1992), which are called H_∞-optimal controllers. Feedback systems for active control should thus ideally be designed to minimise an H_2 cost function while maintaining multiple H_∞ constraints, and such controllers are said to perform H_2/H_∞ control (Rafaely and Elliott, 1999).

If the multiplicative uncertainty in the plant is too great, it may not be possible to obtain good suppression of disturbances while ensuring robust stability and any design method inevitably involves a compromise between performance and robustness.

2.4 Disturbance Enhancement

We can also make some rather general statements about the conditions under which *enhancement* or amplification of the disturbance can take place using the Nyquist plot. Suppose we are interested in the conditions under which the modulus of the sensitivity function, and hence the amplification of the disturbance, is greater than some value A. Since $S(j\omega) = [1 + G(j\omega)H(j\omega)]^{-1}$, we can see that if

$$|S(j\omega)| > A \tag{25}$$

then

$$|1 + G(j\omega)H(j\omega)| < \frac{1}{A} \tag{26}$$

Geometrically, equation (26) implies that the disturbance will be amplified by a factor of at least A if the open-loop frequency response falls within a disc round the Nyquist point of radius $1/A$. This is illustrated on the right-hand side of Figure 3 for $A = 1$ and 2, i.e. amplifications of 0 dB, or no disturbance enhancement, and 6 dB enhancement of the disturbance. The more robustly stable the control system is made, and hence the further the open-loop frequency response remains from the Nyquist point, the smaller will be the maximum amplification of the disturbance.

A phase lag of 180° in the open-loop frequency response will always occur at high frequencies if the magnitude response falls off at least as fast as 40 dB per decade, in which case a very interesting general result can be demonstrated. If this condition holds, which it almost inevitably does in practice, and the plant and controller are stable, then Bode (1945) showed that the integral of the log sensitivity function, which is linearly proportional to the disturbance attenuation in decibels, over the whole frequency range must be zero, i.e.

$$\int_0^\infty 20 \log_{10} |S(j\omega)| \, d\omega = 0, \tag{27}$$

which is called *Bode's sensitivity integral*. This important result implies that the active attenuation, in decibels, integrated over the whole frequency range, is zero.

This profound limitation on performance appears at first sight to impose a severe constraint on all controllers. On closer inspection, however, we realise that good attenuation, $|S(j\omega)| \ll 1$, over a small bandwidth, where the disturbance has significant energy, can be traded for small enhancements, $|S(j\omega)| > 1$, over large ranges of frequency, where the disturbance has little energy. Unfortunately, this disturbance enhancement can only be made arbitrarily small for minimum-phase plants, and for non-minimum-phase plants, a disturbance attenuation over a finite range of frequencies inevitably gives rise to a certain minimum level of disturbance enhancement at another frequency (Freudenberg and Looze, 1985). If the disturbance is "pushed down" at some frequencies, then for a non-minimum-phase plant, the disturbance at other frequencies will inevitably "pop up", which is described as the *waterbed effect* (see, for example, Doyle et al. (1992)). One of the advantages of a feedforward control system, using a time-advanced reference signal, is that the performance is *not* then limited by the Bode sensitivity integral, as discussed by Hon and Bernstein (1998). In this case, broadband attenuation of a disturbance can be achieved without any out-of-band enhancements.

2.5 Analogue Compensators

Apart from analysing the potential stability of a feedback control system, the frequency-domain methods introduced above can also provide a useful tool in the design of such systems. Analogue feedback control systems have traditionally used rather simple controllers. Known as *compensators*, first- and second-order RC circuits have traditionally been used to improve the performance of feedback controllers or increase their relative stability. The circuit of one such compensator, together with its frequency response when $R_1 = 9R_2$, is shown in Figure 4. The frequency response of this circuit can be written as

$$\frac{V_{\text{out}}(j\omega)}{V_{\text{in}}(j\omega)} = \frac{1 + j\omega R_2 C}{1 + j\omega(R_1 + R_2)C}. \tag{28}$$

At low frequencies, the gain of this circuit is unity and it has little phase shift. The useful aspect of the circuit is that at high frequencies the phase shift is again small, but the gain is now $R_2/(R_1 + R_2)$, and so the loop gain can be considerably attenuated at frequencies around the Nyquist point, for example, with little phase shift. In the frequency region from $\omega = 1/(R_1 + R_2)C$ to $\omega = 1/R_2C$, the phase shift is negative and can be significant if $R_1 \gg R_2$. This phase lag can destabilise a feedback system by reducing its phase margin if it occurs too near the Nyquist point. The best design of this *phase-lag compensator* thus involves choosing R_1 and R_2 relative to C and the overall gain

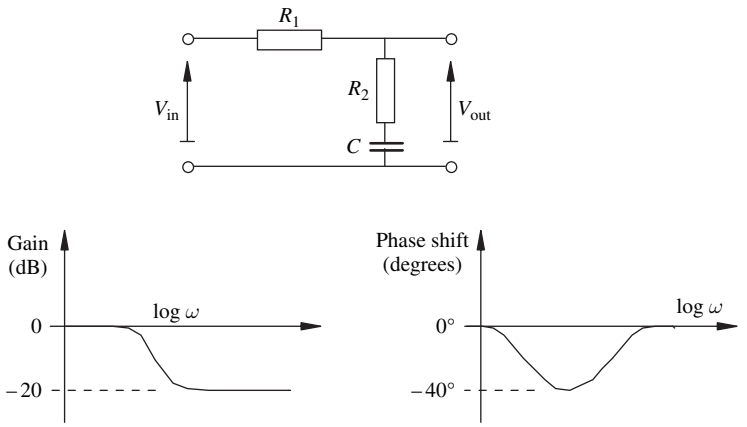


FIGURE 4 An analogue plant controlled by a digital controller, shown together with the data converters and analogue filters generally required in active control.

of the controller, such that the loop gain is reduced near the Nyquist point, without excessive phase lag being introduced in this vicinity. The use of phase-lag compensators in increasing the gain and phase margins in servo systems is fully described by Dorf (1990), for example. Its use in improving the robust stability of a feedback control system used in an active headset will be described in the next section.

3 • Application: Active Headsets

There are two main types of headset that are used for rather different acoustical purposes. The "circumaural", or "closed-backed" headset, or *ear defender* is illustrated in Figure 5 and has a rigid earshell, which is tensioned against the ear by a band over the head, and a soft cushion that is designed to seal the headset against the side of the face. The main purpose of such headsets is generally hearing protection, although they may be fitted with small loudspeakers for communication if the headset is to be worn by air crew for example. The other type of headset is called "superaural" or "open-backed", and has an open foam element between a small loudspeaker and the ear. The main purpose of such headsets is generally sound reproduction, as in portable tape and CD players. It is possible to use active control in either of these headsets to reduce the external noise heard when wearing the headset. The application of active feedback control to the ear defender type of headset was historically investigated first, and we will consider the design and performance of fixed feedback control systems for these headsets. Figure 6 shows

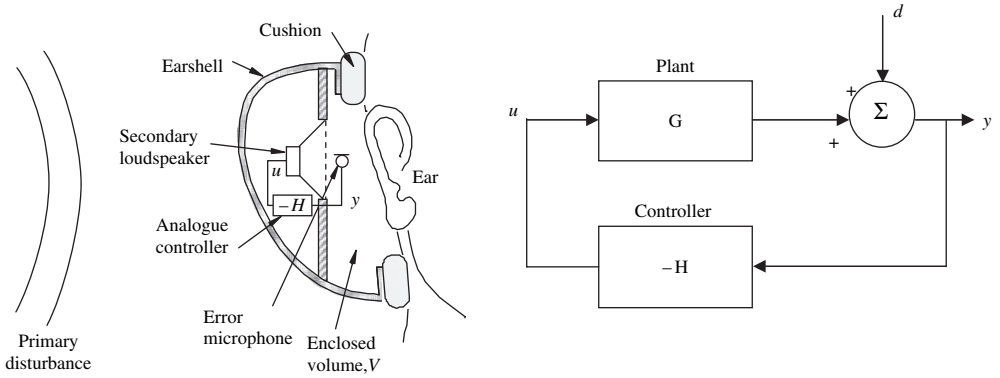


FIGURE 5 An active circumaural headset.

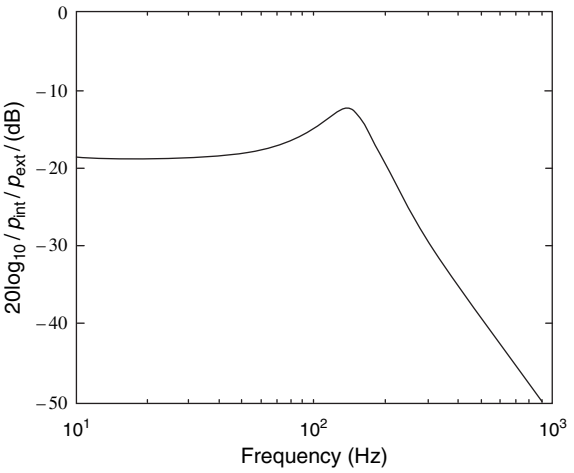


FIGURE 6 The modulus of the internal to external pressure, in decibels, for a typical circumaural headset without active attenuation.

the typical passive performance of such a headset expressed in terms of the ratio of the internal to external pressure. Although the passive performance above about 500 Hz is very good, this is limited at about 100 Hz by the resonance of the earshell mass on the cushion stiffness and at lower frequencies by the compliance of the cushion compared with the compliance of the air cavity (Shaw and Thiessen, 1962; Elliott, 2001).

The use of a feedback loop to improve the low-frequency performance of a headset was originally suggested in the 1950s (Simshauser and Hawley,1955; Olson, 1956). A rather complete theoretical and experimental investigation of an active headset using a fixed analogue controller was reported soon afterwards (Meeker, 1958, 1959). The simplicity and the lack of delay inherent in analogue controllers have made them widely used in commercial systems up to the present day.

The frequency response of the plant, between the input to the loudspeaker and the output of the microphone, in an active closed-back headset fitted to a subject's head has been measured by Wheeler (1986) and is shown in Figure 7. The physical proximity of the secondary source and feedback microphone, which are within about 1 cm of each other, means that there is very little acoustic delay in the plant response. The dynamics are dominated by the mechanical resonance of the loudspeaker at low frequencies (about 200 Hz) and resonances associated with acoustic modes inside the earshell at higher frequencies. It is important to have an appropriate amount of foam filler in the earshell so that these acoustic resonances are well damped but does not cause excessive phase shifts at high frequencies, which can lead to instability. It is also important to ensure that the plant

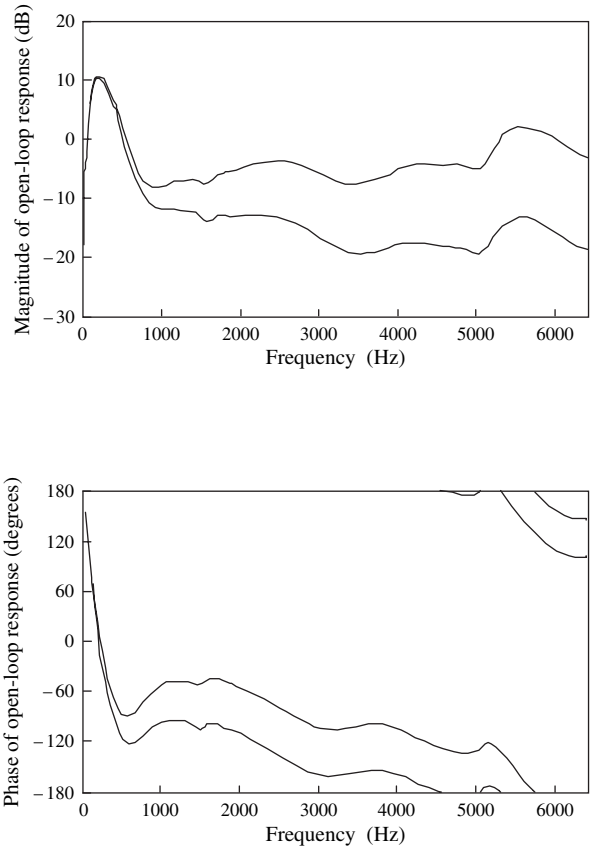


FIGURE 7 Modulus (*upper*) and phase (*lower*) of the open-loop frequency response of an active headset, measured by [34]. The response is shown both before (*upper curves*) and after (*lower curves*) the insertion of a phase-lag compensating network in the loop.

response is as close to being minimum phase as possible, since this will reduce the disturbance enhancement that may occur due to the "waterbed effect", described above.

3.1 Analogue Phase-Lag Compensation

Also shown in Figure 7 is the open-loop frequency response of the active headset when a phase-lag compensator, of the type shown in Figure 4, is included as the analogue controller. The compensator makes little difference to the response at low frequencies, but at high frequencies gives an extra 15 dB or so of attenuation at the price of only about 40° extra phase lag at 6 kHz. The gain margin of the system can be calculated from the open-loop gain at the frequency where the phase shift falls through 180°, which is at about 5.7 kHz without the compensator and 4.5

and 5.2 kHz after the compensator has been included. It can be seen from Figure 7 that the gain margin of the system has been improved by about 15 dB by the incorporation of the phase-lag compensator.

One of the main causes of uncertainty in the response of the active headset is the variation observed when the headset is worn by different subjects. This is due to differences in the fit and sealing of the helmet from person to person, and also because the dynamic response of the ear, and hence the acoustic impedance presented to the headset by the ear, varies from subject to subject. Wheeler (1986) conducted a series of experiments in which he measured the response of an active headset on a number of different subjects and found that the variation in the amplitude of the plant response from 1 to 6 kHz was about $\pm$ 3 dB, and the variation in the phase of the plant response was about $\pm 20^\circ$. The active headset thus has to be designed with gain and phase margins of at least these figures. It is interesting to note, in view of the discussion Section 2, that both the amplitude and phase variation could be fairly accurately accounted for in this case by assuming an unstructured multiplicative uncertainty with an upper bound of $B(\omega) \approx 0.3$ over this frequency range.

An additional 20 dB attenuation of noise was measured inside the headset when this active control system was operating in the third octave band at 200 Hz, although a small degree of amplification of the noise also occured in the third octave bands about 4 kHz (Wheeler, 1986). It should be remembered, however, that the frequency range spanned by a third octave band at 4 kHz is considerably greater than a third octave band at 200 Hz. The attenuation integrated over a linear frequency range thus does not violate the Bode limit discussed above. Crabtree and Rylands (1992) have measured the performance of five different types of active headset and all show a maximum attenuation of about 20 dB at 100–200 Hz, falling to zero below about 30 Hz and above about 1 kHz.

Finally, Figure 8 shows the noise levels in a typical military vehicle in third octave bands, together with the noise levels experienced at the ear of a person wearing either a conventional passive ear defender or an active one, as quoted by Wheeler (1987). The range of levels and frequencies important in speech communication are also shown. The passive attenuation provided by the headset is better above about 125 Hz than it is below this frequency, as predicted by the passive transmission ratio curve in Figure 6. The active control system clearly reduces the external noise levels at low frequencies so they are below the typical speech levels in this important frequency range.

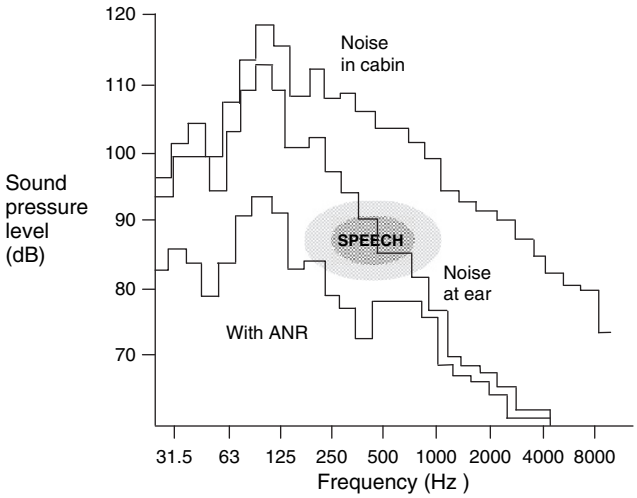


FIGURE 8 Noise levels in the cabin of a typical military vehicle, in third-octave bands, together with the noise levels experienced at the ear by a person wearing a conventional headset (noise at ear) and a headset using feedback active control (with ANR). The range of levels and frequencies important in speech communication are also shown, reproduced from Wheeler (1987).

4 • Digital Controllers

In this section, we will discuss the limitations of implementing the controller "digitally", i.e. with a sampled-time system. The definition of the plant response "seen" by the digital controller has been introduced in Section 1 of Chapter 55 and is further discussed for digital feedback controllers by Kuo (1980) and Franklin et al. (1990). Using this digital plant response, the block diagram of the digital feedback loop is then identical to that of the analogue feedback loop, in Figure 1, but with the Laplace-domain transfer functions replaced with z-domain transfer functions, and continuous-time signals replaced by sampled-time signals.

The sensitivity function of the digital feedback controller can thus be written as

$$S(z) = \frac{E(z)}{D(z)} = \frac{1}{1 + G(z)H(z)}. \tag{29}$$

Although the stability of this transfer function is now determined by whether its poles are inside the unit circle in the z plane, rather than whether they were on the left-hand side of the s plane, the Nyquist criterion for stability remains the same. In the sampled-time case, the open-loop frequency response need only be plotted

in the Nyquist plane for frequencies up to half the sample rate, since the frequency response of a sampled system is periodic. All of the results regarding robust stability in the previous section, for example, can thus be carried directly over into the analysis of digital controllers (Morari and Zafiriou, 1989), on which we will concentrate for the remainder of this chapter.

4.1 Internal Model Control

In this section, a reformulation of the feedback control problem will be outlined that allows a strong link to be made with the work described in the previous chapters. This is achieved by transforming the feedback design problem into an equivalent feedforward one. We will assume that the plant is stable, which is most often the case in active control problems, but provided an unstable plant is first stabilised by any stable feedback controller, exactly the same methods of design as those described below can then be used on the stabilised plant (Morari and Zafiriou, 1989; Maciejowski, 1989). Consider the internal structure of a feedback controller shown in Figure 9, in which the complete negative feedback controller, $-H(z)$, is contained within the dashed lines shown in this figure. The feedback controller contains an *internal model*, $\hat{G}(z)$, of the plant response, $G(z)$, which is fed by the input to the plant and whose output is subtracted from the observed error, $e(n)$. The resulting output, $\hat{d}(n)$, acts as the input to the *control filter*, $W(z)$, whose output, $u(n)$, drives the plant. The transfer function of the complete feedback controller is thus

$$H(z) = \frac{-W(z)}{1 + \hat{G}(z)W(z)}. \tag{30}$$

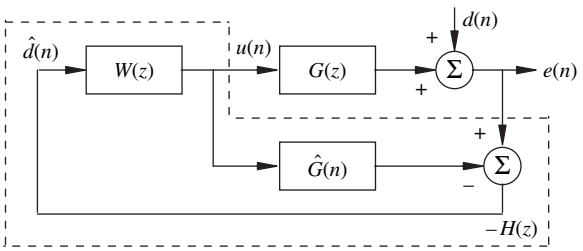


FIGURE 9 Block diagram of a feedback control system in which the feedback controller, $H(z)$, is implemented using an internal model $\hat{G}(z)$ of the plant $G(z)$. This arrangement of feedback controller is known as internal model control (IMC).

Substituting this expression into equation (29) for the sensitivity function of the complete feedback control system gives

$$S(z) = \frac{E(z)}{D(z)} = \frac{1 + \hat{G}(z)W(z)}{1 - \left[G(z) - \hat{G}(z)\right]W(z)}. \tag{31}$$

4.2 Perfect Plant Model

We assume for the time being that the internal plant model perfectly represents the response of the plant, so that $\hat{G}(z) = G(z)$. Morari and Zafiriou (1989) show that with a perfect plant model, the control system shown in Figure 9 is internally stable, that is, any bounded input will only produce a bounded output, provided both $G(z)$ and $W(z)$ are stable. If the plant model is perfect, then the contribution to the observed error from the plant output is completely cancelled by that from the internal model, and so the overall transfer function for the complete feedback control system is

$$S(z) = \frac{E(z)}{D(z)} = 1 + G(z)W(z). \tag{32}$$

The equivalent block diagram of the feedback control system thus becomes entirely *feedforward*.

The immediate advantage of the internal model structure for the feedback controller is that the control filter that minimises the mean-square error can now be designed using the standard Wiener techniques discussed in Chapter 55. This was the original motivation for such a controller, which was proposed for analogue controllers by Newton et al. (1957). These authors showed that the optimum least-squares controllers could then be designed "analytically", rather than using the rather ad hoc methods outlined in Section 2. This approach to the design of the controller was extended by Youla et al. (1976a,b), to the case in which the plant is unstable. For a stable plant, it should be emphasised that *any* stabilising controller can be put into the form of Figure 9, and so the control filter, $W(z)$, "parameterises all stabilising controllers". This structure of feedback controller is sometimes known as *Youla parameterisation* or *Q parameterisation*. In process control, this arrangement of feedback controller is widely used for both the design and implementation of feedback controllers and is known as *internal model control* (IMC), which is how it will be described here. An excellent review of the properties of IMC systems is given by Morari and Zafiriou (1989). The simple physical interpretation of the IMC architecture allows it to be readily extended to deal with non-linear systems, as discussed by Hunt and Sbarbaro (1991) for example,

and multichannel systems. If the control filter, $W(z)$, is implemented as an FIR filter, the mean-square error is a quadratic function of the coefficients of this filter, exactly as in Section 3 of Chapter 55.

4.3 Robust Stability Constraint

The constraint imposed by robust stability becomes particularly simple with an IMC architecture since it depends on the complementary sensitivity function, which following equation (8) is equal to

$$T(z) = 1 - S(z). \tag{33}$$

For an IMC system with perfect plant modelling, the sensitivity function is given by equation (32), and so the complementary sensitivity function takes the simple form

$$T(z) = -G(z)W(z). \tag{34}$$

The complementary sensitivity function is thus also a linear function of the coefficients of an FIR control filter. The condition for robust stability, equation (22), can now be written as

$$\left|G_0\left(e^{j\omega T}\right) W\left(e^{j\omega T}\right)\right| < \frac{1}{B\left(e^{j\omega T}\right)}, \quad \text{for all } \omega T, \tag{35}$$

where $G_0\left(e^{j\omega T}\right)$ is the frequency response of the nominal plant.

Our initial assumption that the plant model was perfect in the IMC system at first sight appears to oversimplify the feedback design problem, by throwing away the feedback. It can be seen above, however, that the reason why feedback is needed in the first place is because of uncertainties in the system under control, and that these uncertainties can be accounted for in a very direct way using the IMC architecture, by making the stability of the system robust to such uncertainty.

The parameterisation of the feedback controller provided by the IMC architecture thus enables both the error signal, equation (32), and the robustness condition, equation (35), to be expressed as linear functions of the frequency response of the control filter. Standard convex optimisation techniques (Boyd and Barratt, 1991) can thus be used to design the feedback controller that minimises the mean-square error while maintaining the robust stability condition, as in the H_2/H_∞ controller designed for an active headset by Rafaely and Elliott (1999).

5 • Hybrid Feedback Controllers

One significant disadvantage of digital feedback controllers for active sound systems is that the delay introduced by the converters and anti-aliasing filters can limit the bandwidth over which significant attenuation can be achieved, as described in Section 1.3. Very high sampling rates could be used to reduce this delay, but then the control filter becomes complicated and the system is expensive to implement. An important advantage of digital feedback systems implemented using the IMC architecture shown in Figure 9, however, is that because it reduces the control problem to a feedforward one, adaptive methods such as those described in Chapter 55 can be used to make the system react to changes in disturbance spectrum. Some care needs to be taken to ensure that the controller is always robust, but the robust stability condition (equation 35) can be enforced with a leakage term in the adaptive algorithm, which limits the magnitude of the control filter's response. Thus if the disturbance spectrum is narrowband, but non-stationary, an adaptive digital feedback controller could achieve good control even without an external reference signal. The reference signal is effectively generated internally in the IMC arrangement as the estimated disturbance signal in Figure 9.

The features of an analogue controller, however, are that it can give good broadband performance, because of its low delay,

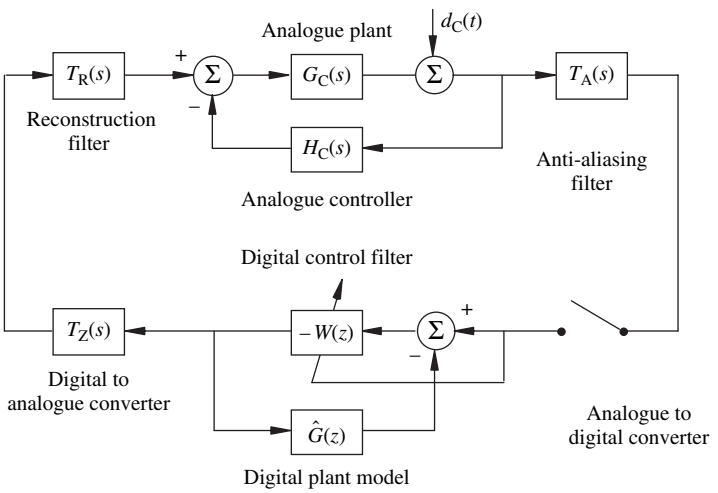


FIGURE 10 Block diagram of a combined feedback control system that has a fixed analogue inner loop, via the analogue controller $H_a(s)$, and an adaptive digital outer loop, via the adaptive digital internal mode control (IMC) system with control filter $W(z)$.

but is limited to stationary disturbances since it is difficult to make it adaptive. In order to use the best features of both the analogue and adaptive digital feedback controllers, a hybrid controller could be used, as shown in Figure 10 (Doelman, 1991; Rafaely and Elliott, 1996; Clark and Bernstein, 1998). In this block diagram, the analogue plant, $G_c(s)$, which describes the acoustic system under control, is first controlled by an analogue feedback system, $H_c(s)$, and then, after anti-alias filtering and sampling, is controlled by an IMC digital controller with adaptive control filter $W(z)$. The digital plant model in this arrangement, $\hat{G}(z)$, must now model the sampled plant response with the inner analogue feedback loop in place. This can be easier to model than the uncontrolled analogue plant since the analogue controller will reduce the magnitude of any resonances in the acoustic system, and will also reduce the variability of the overall response due to plant uncertainty. This in turn will improve the transient response of the adaptive digital controller and allow it to be made robustly stable with a smaller leakage factor. The suppression of resonances in the physical system using analogue feedback techniques may also be used to reduce the convergence time of adaptive digital feedforward controllers. Although the implementation of analogue feedback systems is well established for active control, the practical use of such hybrid controllers is still in its infancy but does offer the possibility of effective control of both broadband and narrowband disturbances.

References

1. Baumann W.T., Saunders W.R. and Robertshaw H.H. (1991) Active suppression of acoustic radiation from impulsively excited structures. *Journal of the Acoustical Society of America*, **88**, 3202–3208.
2. Black M.W. (1934) Stabilised feedback amplifiers. *Bell Syst. Tech.Journal*. **13**, 1–18.
3. Bode M.W. (1945) Network Analysis and Feedback Amplifier Design, Van Nostrand.
4. Boyd S.P. and Barratt C.H. (1991) *Linear Controller Design, Limits of Performance*, Prentice Hall, Englewood Cliffs, NJ.
5. Clark R.L. and Bernstein D.S. (1998) Hybrid control: separation in design. *Journal of Sound and Vibration*, **214**, 784–791.
6. Clark R.L., Saunders W.R. and Gibbs G.P. (1998) *Adaptive Structures, Dynamics and Control*. John Wiley and Sons, New York.
7. Crabtree R.B. and Rylands J.M. (1992) Benefits of active noise reduction to noise exposure in high-risk environments. *Proceedings of Inter-Noise '92*, Poughkeepsie, NY, 295–298.
8. Doelman N.J. (1991) A unified control strategy for the active reduction of sound and vibration. *Journal of Intelligent Materials, Systems and Structures*, **2**, 558–580.

9. Dorf R.C. (1990) *Modern Control Systems*, 6th Ed., Addison-Wesley, Reading, MA.
10. Doyle J.C., Francis B.A. and Tannenbaum A.R. (1992) *Feedback Control Theory*, Maxwell MacMillan International, New York.
11. Elliott S.J. (2001) *Signal Processing for Active Control*, Academic Press, New York.
12. Franklin G.F., Powell J.D. and Workman M.L. (1990) *Digital Control of Dynamic Systems*, 2nd Ed., Addison-Wesley, New York.
13. Franklin G.F., Powell J.D. and Emani-Naeini A. (1994) *Feedback Control of Dynamic Systems*, 3rd Ed., Addison-Wesley, Reading, MA.
14. Freudenberg J.S. and Looze D.P. (1985) Right hand plane poles and zeros and design trade-offs in feedback systems. *IEEE Transactions on Automatic Control*, **AC-30**, 555–565.
15. Fuller C.R., Elliott S.J. and Nelson P.A. (1996) *Active Control of Vibration*, Academic Press, New York.
16. Hon J. and Bernstein D.S. (1998) Bode integral constraints, collocation and spillover in active noise and vibration control. *IEEE Transactions on Control Systems Technology*, **6**(1), 111–120.
17. Hunt K.J. and Sbarbaro D. (1991) Neural networks for nonlinear internal model controller. *IEE Proceedings-D*, **138**, 431–438.
18. Kuo B.C. (1980) *Digital Control Systems*, Holt, Rinehart and Wilson, New York.
19. Maciejowski J.M. (1989) *Multivariable Feedback Design*. Addison-Wesley, New York.
20. Mecker W.F. (1958, 1959) Active ear defender systems, Part I and II. *WADC Technical Reports*, **57–368(I)** and **57–368(II)**.
21. Meirovitch L. (1990). *Dynamics and Control of Structures*, John Wiley, New York.
22. Morari M. and Zafiriou E. (1989) *Robust Process Control*, Prentice Hall, Englewood Cliffs, NJ.
23. Newton G.C., Gould L.A. and Kaiser J.F. (1957) *Analytical Design of Linear Feedback Controls*, John Wiley and Sons, New York.
24. Nyquist M. (1932) Regeneration theory. *Bell Syst.Tech.Journal*, **11**, 126–147.
25. Olson H.F. (1956) Electronic control of noise, reverberation, and vibration. *Journal of the Acoustical Society of America*, **28**, 966–972.
26. Petitjean B. and Legrain I. (1994) Feedback controllers for broadband active noise reduction. *Proceedings of Second European Conference on Smart Structures and Materials*, Glasgow.
27. Preumont A. (1997) *Vibration Control of Active Structures, an introduction*. Kluwer Academic Publishers, Dodrecht.
28. Rafaely B. and Elliott S.J. (1996) Adaptive plant modelling in an internal model controller for active control of sound and vibration *Proc. Identification in Engineering Systems Conference.*
29. Rafaely B. and Elliott S.J. (1999) H_2/H_∞ active control of sound in a headrest: design and implementation. *IEEE Transactions on Control System Technology*, **7**, 79–84.

30. Rubenstein S.P., Saunders W.R., Ellis G.K., Robertshaw H.H. and Baumann W.T. (1991) Demonstration of a LQG vibration controller for a simply-supported plate. *Proceedings of First Conference on Recent Advances in Active Control of Sound and Vibration*, Virginia Tech. pp. 618–630.
31. Shaw E.A.G. and Thiessen G.J. (1962) Acoustics of circumaural ear phones. *Journal of the Acoustical Society of America*, **34**, 1233–1246.
32. Simshauser E.D. and Hawley M.E. (1955) The noise cancelling headset – an active ear defender. *Journal of the Acoustical Society of America*, **27**, 207 (Abstract).
33. Skogestad S. and Postlethwaite I. (1996) *Multivariable Feedback Control, Analysis and Design*, John Wiley and Sons, New York.
34. Wheeler P.D. (1986) Voice communications in the cockpit noise environment – the role of active noise reduction. *PhD Thesis*, University of Southampton, England.
35. Wheeler P.D. (1987) The role of noise cancellation techniques in aircrew voice communications systems. *Proceedings of the Royal Aeronautical Society Symposium on Helmets and Helmet-Mounted Displays, London.*
36. Youla D.C., Bongiorno J.J. and Jabr M.A. (1976a) Modern Wiener-Hopf design of optimal controllers. Part I – the single input-output case. *IEEE Transactions on Automatic Control*, **AC-21**, 3–13.
37. Youla D.C., Jabr M.A.and Bangiorno J.J. (1976b) Modern Wiener-Hofp design of optimal controllers. Part II – the multivariable case. *IEEE Transactions on Automatic Control*, **AC21**, 319–338.
38. Zames G. (1981) Feedback and Optimum Sensitivity; model reference transformations, multiplicative seminorms and approximate inverses. *IEEE Transactions on Automatic Control*, **AC-26**, 301–320.

PART X STRUCTURAL ACOUSTICS AND VIBRATION

Jean-Louis Guyader

Laboratoire Vibrations-Acoustique, Institut National des Sciences Appliquées de Lyon, Villeurbanne, France

58 Intensity Techniques

Finn Jacobsen 1109

Introduction • Sound Intensity and Sound Power • Active and Reactive Sound Fields • Measurement of Sound Intensity • Applications of Sound Intensity

59 Nearfield Acoustical Holography

S.I. Hayek 1129

Introduction • Planar Nearfield Acoustical Holography • Cylindrical Nearfield Acoustical Holography • Spherical and General Surface Nearfield Acoustical Holography

60 Actuator Design and Deployment

Colin H. Hansen 1141

Introduction • Actuator Design • Actuator Deployment Using Genetic Algorithms

61 Materials Testing

Walter Lauriks and Philippe Leclaire 1167

Introduction • Sound Propagation in Rigid Porous Materials • Sound Propagation in Elastic Porous Materials

58

Intensity Techniques

Finn Jacobsen

Technical University of Denmark, Lyngby, Denmark,
e-mail: fja@elektro.dtu.dk

1 • Introduction

Energy considerations are of considerable importance in acoustics. Their usefulness is mainly due to the fact that a statistical approach where the energy of the sound field is considered gives extremely useful approximations. The value of knowing the sound power radiated by a source is due to the fact that this quantity is largely independent of the surroundings of the source in the audible frequency range and makes it possible to predict the sound pressure generated by the source with very little additional information. By comparison, solving the wave equation is infinitely more complicated and requires much more information.

2 • Sound Intensity and Sound Power

Sound intensity is a measure of the flow of acoustic energy in a sound field, defined as a vector that describes the time average of the net flow of sound energy per unit area in a direction perpendicular to the area. This quantity is measured as power per unit area (W/m^2). The instantaneous sound intensity $\mathbf{I}(t)$ is the product of the sound pressure $p(t)$ and the particle velocity $\mathbf{u}(t)$,

$$\mathbf{I}(t) = p(t)\mathbf{u}(t). \tag{1}$$

By combining the equation of conservation of mass, the adiabatic relation between changes in the pressure and in the density, and

Euler's equation of motion, one can show that the divergence of the instantaneous intensity equals the negative rate of change of the sum of the potential and kinetic energy density, $w(t)$,

$$\nabla \cdot \mathbf{I}(t) = -\frac{\partial w(t)}{\partial t}. \tag{2}$$

This is the equation of conservation of sound energy [1,2], which expresses the simple fact that a net flow of energy away from a point in a sound field is accompanied by a corresponding reduction of the sound energy density at the point because the energy is conserved. Integrating this equation over a volume V enclosed by the surface S gives, with Gauss's theorem,

$$\int_S \mathbf{I}(t) \cdot d\mathbf{S} = -\frac{\partial}{\partial t}\left(\int_V w(t) dV\right) = -\frac{\partial E(t)}{\partial t}, \tag{3}$$

in which $E(t)$ is total sound energy within the surface as a function of time. The left-hand term is the net outflow of sound energy through the surface and the right-hand term is the negative rate of change of the total sound energy within the surface.

In practice, we are often concerned with stationary sound fields and the time-averaged sound intensity rather than the instantaneous intensity, that is,

$$\langle \mathbf{I}(t) \rangle_t = \langle p(t)\mathbf{u}(t) \rangle_t . \tag{4}$$

For simplicity, the symbol $\mathbf{I}$ is used for this quantity in what follows. If the sound field is harmonic with angular frequency $\omega = 2\pi f$, the complex representation of the sound pressure and the particle velocity can be used, which leads to the expression

$$\mathbf{I} = \tfrac{1}{2}\mathrm{Re}\{p\mathbf{u}^*\}, \tag{5}$$

where $\mathbf{u}^*$ denotes the complex conjugate of $\mathbf{u}$.

A consequence of (3) is that the integral of the normal component of the time-averaged sound intensity over a closed surface is zero,

$$\int_S \mathbf{I} \cdot d\mathbf{S} = 0, \tag{6}$$

when there is neither generation nor dissipation of sound energy within the surface, also in the presence of steady sources *outside* the surface. By contrast, if the surface encloses a steady source,

then the surface integral of the time-averaged intensity equals the sound power emitted by the source P_a, that is,

$$\int_S \mathbf{I} \cdot d\mathbf{S} = P_a, \tag{7}$$

irrespective of the presence of other steady sources outside the surface. This important equation is the basis for intensity-based sound power measurements.

3 • Active and Reactive Sound Fields

A typical characteristic of the sound field near a source is that the sound pressure and the particle velocity are partly out of phase. To describe such a phenomenon, one may introduce the concepts of *active* and *reactive* sound fields.

The reactive intensity is a vector defined as the imaginary part of the product of the complex pressure and the complex conjugate of the particle velocity,

$$\mathbf{J} = \tfrac{1}{2} \operatorname{Im}\{p\mathbf{u}^*\} \tag{8}$$

(with the $e^{j\omega t}$ convention) [3]; cf. (5). The physical significance of this quantity can be illustrated by noting that the sound pressure near a monopole is 90° out of phase with the particle velocity,

$$p(r) = \frac{j\omega\rho Q}{4\pi r} e^{j(\omega t - kr)}, \tag{9}$$

$$u_r(r) = \frac{p(r)}{\rho c}\left(1 + \frac{1}{jkr}\right), \tag{10}$$

where ρ is the density of the medium, c the speed of sound, $k = \omega/c$ the wavenumber, Q the volume velocity of the monopole, and r the distance to the source. From (5), (8), (9), and (10) it follows that

$$I_r(r) = \frac{|p(r)|^2}{2\rho c}, \tag{11}$$

$$J_r(r) = \frac{|p(r)|^2}{2\rho c k r}, \tag{12}$$

and these expressions show that the sound field is strongly reactive, $J_r >> I_r$, when $kr << 1$.

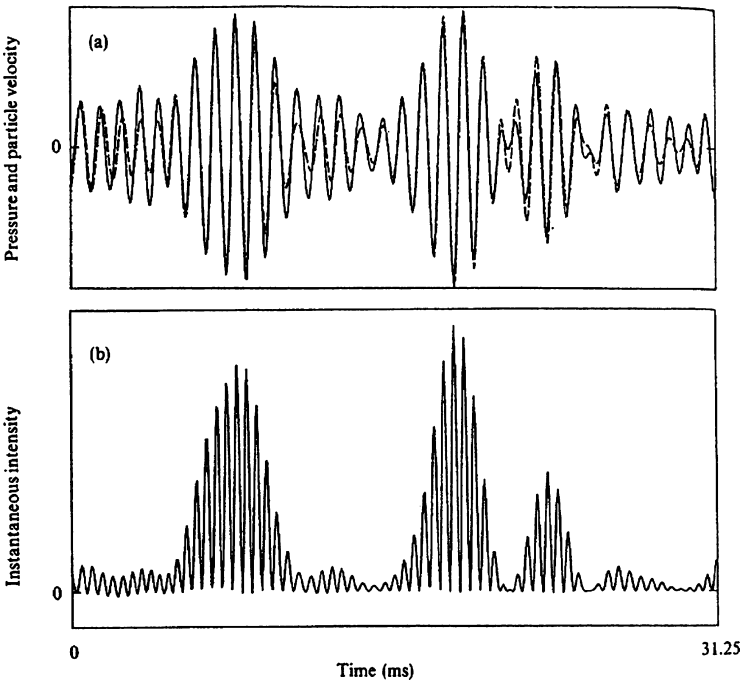


FIGURE 1 Measurement 30 cm from a loudspeaker driven with one-third octave noise with a center frequency of 1 kHz. (**a**) Instantaneous sound pressure (*solid line*) and instantaneous particle velocity multiplied by ρc (*dashed line*); (**b**) instantaneous sound intensity [4].

Figures 1 and 2 illustrate the phenomenon. Figure 1 shows the result of a measurement at a position 30 cm from a small loudspeaker driven with a band of one-third octave noise with a center frequency of 1 kHz. The sound pressure and the particle velocity (normalized by ρc) are almost identical; therefore the instantaneous intensity is positive (cf. (4)). This is an active sound field. In Figure 2 the result of a similar measurement at a distance of a few centimeters from the cone of a loudspeaker driven with one-third octave noise with a center frequency of 250 Hz is shown. In this case the sound pressure and the particle velocity are almost 90° out of phase, and therefore the instantaneous intensity fluctuates about zero. This is an example of a reactive sound field.

In general, vibrating structures generate more complicated sound fields. For example, an infinite plate in the plane $z = 0$ carrying a propagating bending wave with the normal vibrational velocity

$$U(x,y,t) = Ae^{j(\omega t - k_b x)} \tag{13}$$

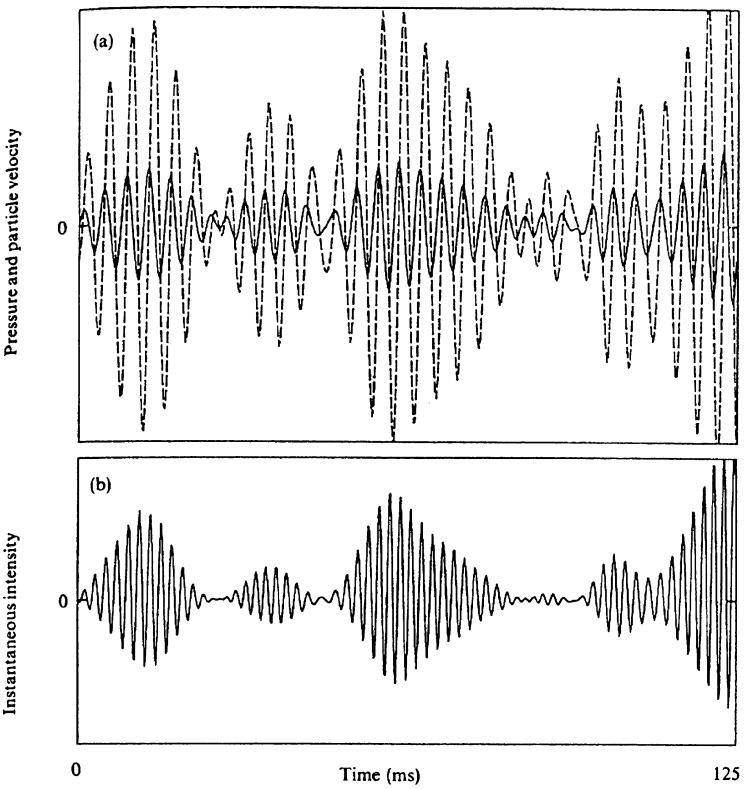


FIGURE 2 Measurement in the near field of a loudspeaker driven with one-third octave noise with a center frequency of 250 Hz. (**a**) Instantaneous sound pressure (s*olid line*) and instantaneous particle velocity multiplied by ρc (*dashed line*); (**b**) instantaneous sound intensity [4].

generates the sound pressure [2]

$$p(x,y,z) = A\rho c \frac{k}{k_z} e^{j(\omega t - k_b x - k_z z)}, \tag{14}$$

and the particle velocity

$$u_z(x,y,z) = \frac{p(x,y,z)}{\rho c} \frac{k_z}{k}, \tag{15}$$

$$u_x(x,y,z) = \frac{p(x,y,z)}{\rho c} \frac{k_b}{k}, \tag{16}$$

in the half-space above the plate, $z \geq 0$. The z-component of the wavenumber is

$$k_z = \begin{cases} \sqrt{k^2 - k_b^2} & \text{for } k_b \leq k \\ -j\sqrt{k_b^2 - k^2} & \text{for } k_b > k. \end{cases} \tag{17}$$

Inspection of these equations leads to the conclusion that the generated sound field is purely active (the pressure and the particle velocity are in phase and $\mathbf{J} = \mathbf{0}$) when $k_w \leq k$, corresponding to the bending wave traveling with supersonic speed (faster than the speed of sound), whereas the sound field is reactive ($I_z = 0$, $J_z > 0$) when the bending wave travels with subsonic speed (slower than the speed of sound). In the latter case the vibrating plate generates an evanescent wave in which the sound pressure and the particle velocity in the normal direction are 90° out of phase. The evanescent wave dies out exponentially with the distance to the plate. There is no flow of energy out of

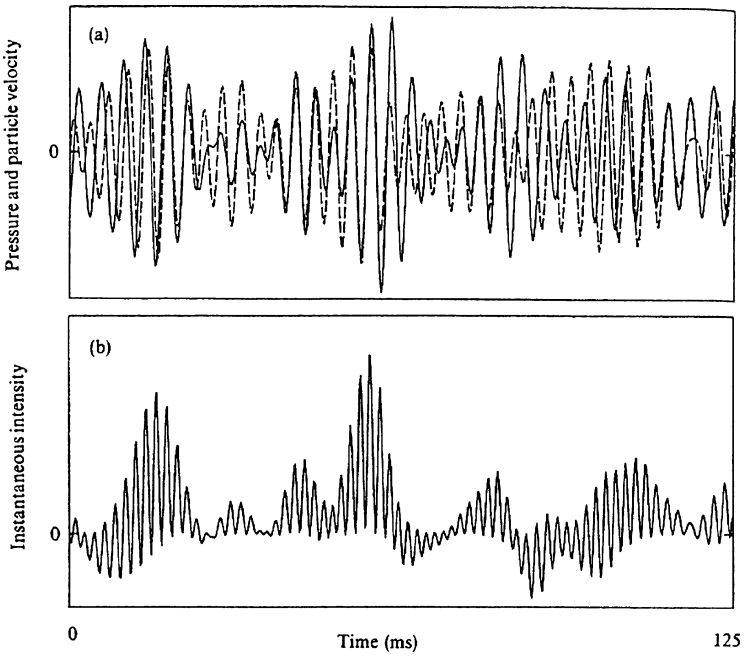


FIGURE 3 Measurement 30 cm from a vibrating steel box driven with one-third octave noise with a center frequency of 250 Hz. (**a**) Instantaneous sound pressure (s*olid line*) and instantaneous particle velocity multiplied by ρc (*dashed line*); (**b**) instantaneous sound intensity [4].

the plate, but there is nevertheless a local region with a mean flow of sound energy parallel to the plate.

More realistic finite vibrating panels generate still more complicated sound fields, but they tend to be predominantly active or reactive depending on whether the speed of free bending waves in the panel is supersonic or subsonic.

Figure 3 shows data measured about 30 cm from a vibrating box made of 3-mm steel plates. It is apparent that the vibrating structure generates a complicated sound field that is both active and reactive.

4 • Measurement of Sound Intensity

In a plane propagating wave the sound pressure p and the particle velocity u_r are in phase and related by the characteristic impedance of air,

$$u_r(t) = \frac{p(t)}{\rho c}. \tag{18}$$

Under such conditions the intensity is

$$I_r = \langle p(t)u_r(t)\rangle_t = \langle p^2(t)\rangle_t/\rho c = p_{\text{rms}}^2/\rho c = \frac{|p|^2}{2\rho c}, \tag{19}$$

where p_{rms}^2 is the mean square pressure and p in the rightmost expression is the complex amplitude of the pressure in a harmonic plane wave. In this particular case the sound intensity is seen to be simply related to the rms sound pressure, which can be measured with a single microphone. Under free-field conditions the sound field generated by any source of finite extent is "locally plane" sufficiently far from the source (the "Sommerfeld radiation condition" [1]), and thus (18) and (19) are good approximations. This is the basis for the free-field method of sound power determination (which requires an anechoic room); the sound pressure is measured at a number of points on a sphere that encloses the source [5]. A practical consequence of (19) is the simple relation between the sound intensity level (re $I_{\text{ref}} = 1\,\text{pW/m}^2$) and the sound pressure level (re $p_{\text{ref}} = 20\ \mu\text{Pa}$), $L_I \simeq L_p$, which is due to the fact that $\rho c \simeq p_{\text{ref}}^2/I_{\text{ref}} = 400\ \text{kgm}^{-2}\text{s}$ in air under normal ambient conditions. At a static pressure of 101.3 kPa and a temperature of 23°C the error is only 0.1 dB. However, in the general case there is no simple relation between the sound intensity and the sound pressure, and both the sound pressure and the particle velocity must be measured simultaneously and their instantaneous product time-averaged as indicated

by (4). This requires the use of a more complicated device than a single microphone; at least two transducers are needed.

One method known as the "*p–p* method" is based on Euler's equation of motion,

$$\nabla p = -\rho \frac{\partial \mathbf{u}}{\partial t}. \tag{20}$$

The gradient of the sound pressure ∇p is approximated by a finite difference measured with two closely spaced pressure microphones, and the average of the two microphone signals is used as the pressure [2,6]. An alternative method known as the "*p–u* method" combines a pressure microphone with a particle velocity transducer [2]. The latter method has been hampered by the absence of reliable particle velocity transducers until recently, but with the advent of the Microflown particle velocity sensor [7,8], it may have a future.

4.1 The *p–p* Measurement Principle

The *p–p* method based on two closely spaced pressure microphones (or hydrophones) has dominated sound intensity measurement since the early 1980s, and both the IEC standard on instruments for the measurement of sound intensity and the corresponding North American ANSI standard deal exclusively with *p–p* measurement systems [9,10]. The success of this method can be explained by the fact that condenser microphones are more stable and reliable than any other acoustic transducer. A commercial example of a sound intensity probe based on this

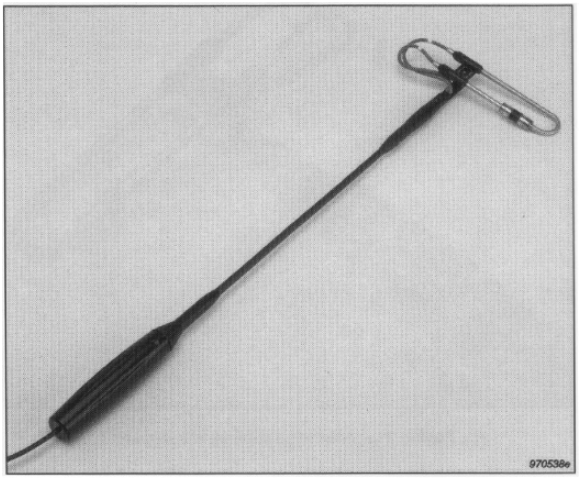


FIGURE 4 Sound intensity probe with the microphones in the "face-to-face" arrangement produced by Brüel and Kjær [11].

measurement principle with the two microphones "face to face" is shown in Figure 4. The "spacer" between the microphones tends to stabilize the "acoustic distance" between them.

4.1.1 Sources of Error in Measurement of Sound Intensity with *p–p* Measurement Systems

The most important limitations of sound intensity measurement systems based on the *p–p* approach are caused by the finite-difference approximation, scattering and diffraction, and instrumentation phase mismatch.

The accuracy of the finite-difference approximation obviously depends on the separation distance and the wavelength. For a plane wave of axial incidence the finite-difference error can be shown to be

$$\hat{I}_r/I_r = \frac{\sin k\Delta r}{k\Delta r}, \tag{21}$$

where $\hat{I}_r$ is the estimate, I_r the "true" intensity, and Δr the separation distance [12]. The upper frequency limit of *p–p* intensity probes has generally been considered to be the highest frequency at which the finite-difference error for axial plane wave incidence is acceptably small. With 12 mm between the microphones (a typical value), this gives an upper limiting frequency of about 5 kHz. However, in the late 1990s it was discovered that the effect of scattering and diffraction in combination with the resonance of the small cavity between the spacer and the diaphragm of each microphone not only tends to counterbalance the finite-difference error but in fact for a certain length of the spacer very nearly cancels it in any sound field [13]. Figure 5 shows the error of the resulting sound

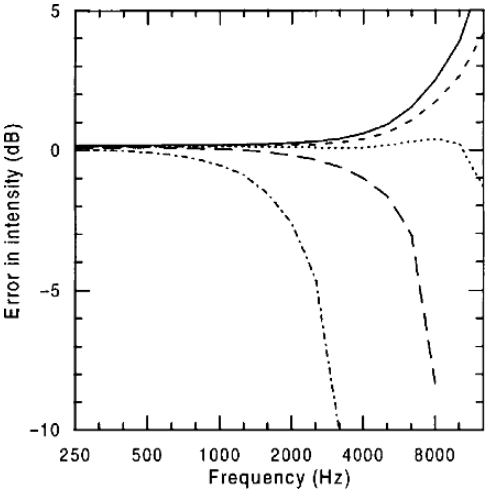


FIGURE 5 Error of a sound intensity probe with 1/2-in. microphones in the face-to-face configuration in a plane wave of axial incidence for different spacer lengths: 5 mm (*solid line*); 8.5 mm (*dashed line*); 12 mm (*dotted line*); 20 mm (*long dashes*); 50 mm (*dash-dotted line*) [13].

intensity estimate. A practical consequence of this finding is that the upper frequency limit of a sound intensity probe with 1/2-in. microphones separated by a 12-mm spacer in the face-to-face arrangement is about 10 kHz, almost an octave higher than the frequency limit determined by the finite-difference approximation. The combination of 1/2-in. microphones and a 12-mm spacer is now regarded as optimal, and longer spacers are only used when the focus is on low frequencies.

Microphones for measurement of sound intensity with the *p–p* method have to be phase matched extremely well, and state-of-the-art sound intensity microphones are matched to a maximum phase response difference of 0.05° below 250 Hz and a phase difference proportional to the frequency above 250 Hz (say, 0.2° at 1 kHz) [14]. However, phase mismatch between the two measurement channels is nevertheless the most serious source of error in measurement of sound intensity with *p–p* measurement systems, even with such equipment. It can be shown that a small phase mismatch error ϕ_{pe} gives rise to a bias error that can be approximated by the following expression [15],

$$\hat{I}_r \simeq I_r - \frac{\varphi_{\mathrm{pe}}}{k\Delta r}\frac{p_{\mathrm{rms}}^2}{\rho c} = I_r\left(1 - \frac{\varphi_{\mathrm{pe}}}{k\Delta r}\frac{p_{\mathrm{rms}}^2/\rho c}{I_r}\right). \tag{22}$$

This expression shows that the effect of a given phase error is inversely proportional to the frequency and the microphone separation distance and is proportional to the ratio of the mean square sound pressure to the sound intensity. If the pressure–intensity index,

$$\delta_{pI} = 10\log\left(\frac{p_{\mathrm{rms}}^2 \rho \mathrm{c}}{|I_r|}\right), \tag{23}$$

is large, say, more than 12 dB, then the "true" phase difference between the two pressure signals is small and even the small phase errors mentioned above will give rise to significant bias errors. Because of phase mismatch, it will rarely be possible to make reliable measurements below, say, 80 Hz unless a longer spacer than the usual 12-mm spacer is used.

The phase error can be measured by exposing the two pressure microphones to the same pressure p_0. Under such conditions the "true" intensity is zero and (22) gives

$$I_0 \simeq -\frac{\varphi_{\mathrm{e}}}{k\Delta r}\frac{p_0^2}{\rho c}. \tag{24}$$

The "residual intensity" I_0 should obviously be as small as possible, or, otherwise expressed, the pressure–residual intensity index,

$$\delta_{pI_0} = 10\log\left(\frac{p_0^2\rho c}{|I_0|}\right) = 10\log\left(\frac{k\Delta r}{|\varphi_e|}\right), \tag{25}$$

should be as large as possible. As can be seen, this quantity depends on Δr. With 12 mm between the microphones and phase errors as quoted above (25), the pressure–residual intensity index is 18 dB from 250 Hz and upward, 15 dB at 125 Hz, and 12 dB at 63 Hz.

A typical value for state-of-the art equipment with 12 mm between the microphones is 20 dB in most of the frequency range.

The pressure–residual intensity index of the measurement system should be checked occasionally. Combined with the pressure–intensity index of the actual measurement, it makes it possible to estimate the error. The condition $\delta_{pI} < \delta_{pI_0} - 7\,\text{dB}$ ensures that the error due to phase mismatch is less than 1 dB ("survey accuracy" [16]). The stronger requirement $\delta_{pI} < \delta_{pI_0} - 10\,\text{dB}$ corresponds to the error being less than 0.5 dB ("engineering accuracy" [16]). These values correspond to the phase error of the equipment being 5 and 10 times less than the actual phase angle in the sound field, respectively.

Most applications of sound intensity measurements involve integrating the normal component of the intensity over a surface. The global version of (22) is found by integrating over a surface that encloses a source. The result is the expression

$$\hat{P}_a = \int_S \hat{\mathbf{I}}\cdot d\mathbf{S} \simeq P_a\left(1 - \frac{\varphi_{pe}}{k\Delta r}\frac{\int_S (p_{rms}^2\rho c)\,dS}{\int_S \mathbf{I}\cdot d\mathbf{S}}\right), \tag{26}$$

where P_a is the "true" sound power of the source within the surface [15]. It is still possible to supervise the bias error by comparing the pressure–intensity index with the pressure-residual intensity index, although measuring the former quantity now involves averaging over the measurement surface.

Sources outside the measurement surface do not contribute to the surface integral of the "true" intensity (the denominator of the second term on the right-hand side of (26)), but they invariantly increase the surface integral of the mean square pressure (the numerator of the second term), from which it follows that even a very small phase error imposes restrictions on the amount of extraneous noise that can be tolerated in sound power measurements with a *p–p* sound intensity measurement system.

It can be difficult to avoid that the intensity probe is exposed to airflow, for example in measurements near air-cooled machinery. Turbulence generates flow noise (unrelated with the sound field) that contaminates the signals from the two microphones, and at low frequencies these signals are correlated and thus interpreted by the measurement system as intensity. The "false" intensity is mainly a problem below 200 Hz [17]. A longer spacer helps reducing the problem, but the most efficient expedient is to use of a windscreen of porous foam.

Yet another source of error is the electrical self-noise from the microphones and preamplifier circuits. Although the level of such noise is very low in modern ½-in. condenser microphones, it can have a serious influence on sound intensity measurement at low levels. The noise increases the random error associated with a finite averaging time at low frequencies, in particular if the pressure–intensity index takes a large value. The problem reveals itself by poor repeatability [18].

4.1.2 Calibration of *p–p* Sound Intensity Measurement Systems

Calibration of *p–p* sound intensity measurement systems is fairly straightforward: the two pressure microphones are calibrated with a pistonphone in the usual manner. However, because of the serious influence of phase mismatch the pressure–residual intensity index should also be determined. The IEC standard for sound intensity instruments and its North American counterpart [9,10] specify minimum values of the acceptable pressure–residual intensity index for the probe as well as for the processor; according to the results of a test the instruments are classified as being of "class 1" or "class 2". The test involves subjecting the two microphones of the probe to identical pressures in a small cavity driven with wideband noise. A similar test of the processor involves feeding the same signal to the two channels. The pressure and intensity response of the probe should also be tested in a plane propagating wave as a function of the frequency, and the directional response of the probe is required to follow the ideal cosine law within a specified tolerance.

4.2 The *p–u* Measurement Principle

A *p–u* sound intensity measurement system combines two different transducers: a pressure microphone and a velocity transducer. The sound intensity is simply the time average of the instantaneous product of the pressure and particle velocity signal,

$$I_r = \langle p u_r \rangle_t = \tfrac{1}{2}\,\mathrm{Re}\{p u_r^*\}, \tag{27}$$

where the latter expression is based on the complex exponential representation.

A micromachined hot-wire anemometer called the "Microflown" has recently become available for measurement

of the particle velocity [7], and an intensity probe based on this device in combination with a small pressure microphone is now in commercial production [8]; see Section 9.1.2. The Microflown probe is very small – much smaller than a standard *p–p* probe. Thus it is possible to measure very close to a vibrating surface with this device. At the time of writing (2004) there was still relatively little experimental evidence of the practical utility of the Microflown intensity probe, but it seems to have potential [19].

A variant of the *p–u* method, the "surface intensity method", combines a pressure microphone with a transducer that measures the vibrational displacement, velocity, or acceleration of a solid surface, for example with an accelerometer or a laser vibrometer [20,21]. Obviously, this method can only give the normal component of the sound intensity near vibrating surfaces. A disadvantage of the surface intensity method is that sound fields near complex sources often are very complicated, which makes it necessary to measure at many points.

4.2.1 Sources of Error in Measurement of Sound Intensity with *p–u* Measurement Systems

Evidently, some of the limitations of any *p–u* intensity probe must depend on the particulars of the particle velocity transducer. Some general problems are described in the following.

Irrespective of the measurement principle used in determining the particle velocity, there is one fundamental problem: the pressure and the particle velocity transducer will invariably have different phase responses. One must compensate for this "*p–u* phase mismatch", otherwise the result is likely to be meaningless. Even a small residual *p–u* mismatch error can have serious consequences under certain conditions, as can be seen by introducing such a small phase error, ϕ_{ue}, in (27). The result is

$$\hat{I}_r = {}^1\!/_2 \,\mathrm{Re}\left\{p u_r^* e^{-j\varphi_{ue}}\right\} = \mathrm{Re}\left\{(I_r + jJ_r)\right.$$
$$\left.(\cos\varphi_{ue} - j\sin\varphi_{ue})\right\} \simeq I_r + \varphi_{ue} J_r. \qquad (28)$$

Equation (28) demonstrates that a small uncompensated *p–u* phase mismatch error will give rise to a significant bias error when $J_r >> I_r$. On the other hand, it also shows that substantial *p–u* phase errors can be tolerated if $J_r << I_r$; even phase mismatch of 35° gives a bias error of less than 1 dB under such conditions. In other words, the phase calibration is critical when measurements are carried out under near-field conditions, but not at all critical if the measurements are carried out in the far-field of a source. The "reactivity" (the ratio of the reactive to the active intensity) indicates whether this source of error is of concern or not.

The global version of (28) is found by integrating over a surface that encloses a source,

$$\hat{P}_{\mathrm{a}} = \int_S Re\left\{(\mathbf{I}+j\mathbf{J})\left(\cos\varphi_{\mathrm{ue}} - j\sin\varphi_{\mathrm{ue}}\right)\right\}\cdot \mathrm{d}\mathbf{S} \simeq P_{\mathrm{a}} + \varphi_{\mathrm{ue}}\int_S \mathbf{J}\cdot\mathrm{d}\mathbf{S} = P_{\mathrm{a}}\left(1+\varphi_{\mathrm{ue}}\frac{\int_S \mathbf{J}\cdot\mathrm{d}\mathbf{S}}{\int_S \mathbf{I}\cdot\mathrm{d}\mathbf{S}}\right), \tag{29}$$

and this shows that uncompensated *p–u* phase mismatch is a source of error when the reactivity (the ratio of the surface integral of the reactive intensity to the surface integral of the active intensity) is large. This will typically occur at low frequencies when the measurement surface is close to the source. Thus the reactivity is an important error indicator for *p–u* probes [19]. By contrast, the pressure–intensity index is not relevant for *p–u* probes.

Airflow, and in particular unsteady flow, is a more serious problem for *p–u* than for *p–p* sound intensity measurement systems, irrespective of the particulars of the particle velocity transducer, since the resulting velocity cannot be distinguished from the velocity associated with the sound waves. Windscreens of porous foam reduce the problem.

4.2.2 Calibration of *p–u* Sound Intensity Measurement Systems

One cannot calibrate a particle velocity transducer in a coupler; it is necessary to expose the probe to a sound field with a known relation between the pressure and the particle velocity, for example a plane propagating wave a simple spherical wave, or a standing wave [7,8,19]. There are no standardized calibration procedures and no standards for testing such instruments. However, recent experimental results seem to indicate the possibility of calibrating the particle velocity channel of the Microflown intensity probe in the far field of a loudspeaker in an anechoic room, in the near field of a very small source (sound emitted from a small hole in a large baffle), and in a rigidly terminated standing-wave tube.

5 • Applications of Sound Intensity

Some of the most important practical applications of sound intensity measurements are now briefly discussed.

5.1 Sound Power Determination

One of the most important applications of sound intensity measurements is the determination of the sound power of operating machinery in situ. Sound power determination using intensity measurements is based on (7). The analysis of errors and

limitations presented in Section 4 leads to the conclusion that the sound intensity method is suitable for determining the sound power of stationary sources in stationary background noise if the pressure–intensity index is within acceptable limits (with *p–p* measurement systems) or if the reactivity is within acceptable limits (with *p–u* measurement systems). On the other hand, the method is *not* suitable for sources that operate in long cycles (because the sound field will change during the measurement), in nonstationary background noise (for the same reason), and (with *p–p* measurement systems) for very weak sources of low-frequency noise (because of large random errors caused by electrical noise in the microphone signals). The standards for sound power determination [16,22,23] only consider measurements with the *p–p* technique.

The surface integral can be approximated either by sampling at discrete points [22] or by scanning manually or with a robot over the surface [16,23]. With the "scanning" approach, the intensity probe is moved continuously over the measurement surface. Two perpendicular scanning paths on a partial surface are shown in Figure 6. This measurement procedure was regarded with much skepticism for more than a decade [24], but it is now generally regarded as more accurate and very much faster and more convenient than the procedure based on fixed points. A moderate scanning rate, say 0.5 m/s, and a "reasonable" scan line density should be used, say 5 cm between adjacent lines if the surface is very close to the source, 20 cm if it is further away. Usually, the measurement surface is divided into a number of segments that are convenient to scan. In practice, one uses a surface of a simple shape at some distance, say 25–50 cm, from the source. If there is

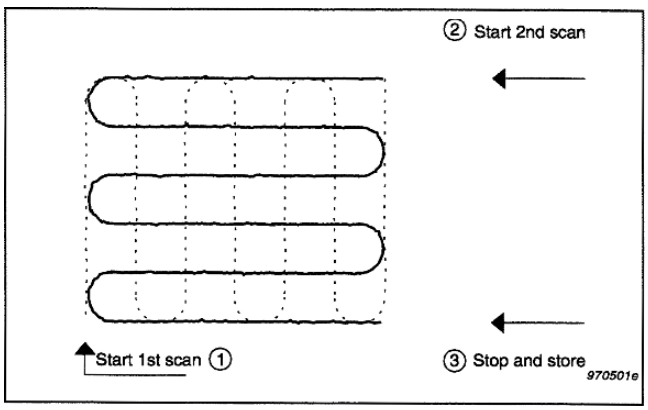


FIGURE 6 Two scanning paths on a measurement surface [11].

a strong reverberant field or significant ambient noise from other sources, the measurement surface should be somewhat closer to the source under study.

5.2 Radiation Efficiency of Structures

The radiation efficiency of a structure is a measure of how effectively it radiates sound. This dimensionless quantity is defined as

$$\sigma = \frac{P_a}{\rho c \langle v_n^2(t) \rangle_{t,s} S}, \tag{30}$$

where P_a is the sound power radiated by the structure, S its surface area, v_n normal velocity of the surface, and $\langle\rangle_{t,s}$ indicates averaging over time and space. This quantity can be measured directly with a sound intensity probe if it can be placed sufficiently close to the structure [25]. The Microflown *p–u* probe has an advantage in this respect.

5.3 Noise Source Identification

A noise reduction project usually starts with the identification and ranking of noise sources and transmission paths. Intensity measurements make it possible to determine the sound power contribution of the various components separately. Plots of the sound intensity normal to a measurement surface (noise mapping) can be used in locating noise sources. Figure 7 shows an example of intensity measurements near a ship window.

Figure 8 shows calculated vector plots of the sound intensity in the sound field generated by a plywood plate carrying a bending wave. It is apparent that the generated sound field is very complicated (cf. the considerations in Section 3). Note in particular the recirculation of sound energy; there are regions with negative intensity – in other words, some parts of the plate act as active absorbers of sound. This phenomenon obviously makes it difficult to locate regions with a net flow of energy to the far field. Such observations have let Williams to introduce the concept

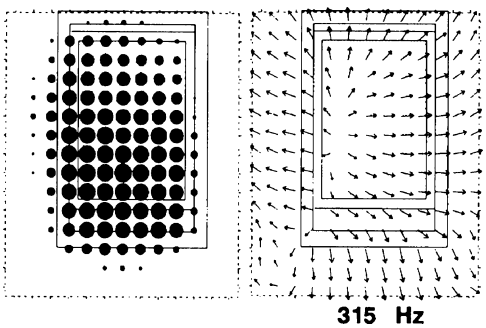


FIGURE 7 Sound intensity distribution near a single cabin window. *Left figure*: component perpendicular to the measurement plane; *right figure*: tangential component [26].

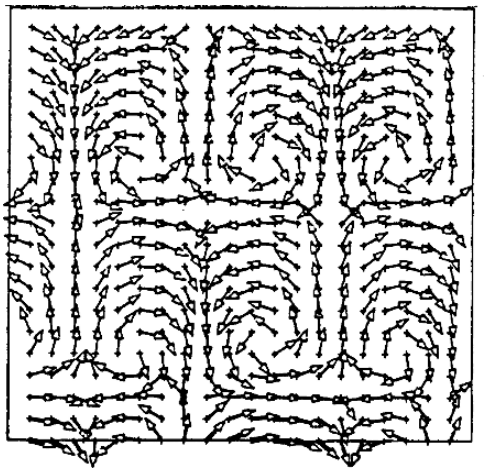


FIGURE 8 Sound intensity distribution above a 10-mm plywood plate carrying a bending wave at 1 kHz [27].

of supersonic intensity; subsonic parts are removed by filtering near-field holographic data in the wavenumber domain [28]. The supersonic intensity is undoubtedly useful for this purpose, but it cannot be obtained from ordinary intensity measurements; holographic processing of data obtained with microphone arrays is needed.

5.4 Transmission Loss of Structures

The transmission loss of a partition is the ratio of incident to transmitted sound power in logarithmic form. The traditional method of measuring this quantity requires a transmission suite consisting of two vibration-isolated reverberation rooms. The sound power incident on the partition under test in the source room is deduced from the spatial average of the mean square sound pressure on the assumption that the sound field is diffuse, and the transmitted sound power is determined from a similar measurement in the receiving room where, in addition, the reverberation time must be determined. The sound intensity method makes it possible to measure the transmitted sound power directly [29,30]. By contrast, one cannot measure the incident sound power in the source room using sound intensity, since the method gives the *net* intensity. If the intensity method is used for determining the transmitted sound power, it is not necessary that the sound field in the receiving room is diffuse, from which it follows that only one reverberation room is needed. However, the main advantage of the intensity method is that it makes it possible to evaluate the contributions of individual transmission paths. Obviously, each sound power measurement must satisfy the conditions for reliable sound power measurements analyzed in Section 4.

5.5 Other Applications

The fact that the sound intensity level is considerably lower than the sound pressure level in a diffuse, reverberant sound field has led to the idea of replacing a measurement of the "emission sound pressure level" generated by machinery at the operator's position by a measurement of the sound intensity level, because the latter is less affected by diffuse background noise [31]. This method has recently been standardized [32].

References

1. A.D. Pierce: *Acoustics: An Introduction to Its Physical Principles and Applications*, 2nd ed. (Acoustical Society of America, New York, 1989)
2. F.J. Fahy: *Sound Intensity*, 2nd ed. (E & FN Spon, London, 1995)
3. J.A. Mann III, J. Tichy, A.J. Romano: Instantaneous and time-averaged energy transfer in acoustic fields, J. Acoust. Soc. Am. **82**, 17–30 (1987)
4. F. Jacobsen: A note on instantaneous and time-averaged active and reactive sound intensity, J. Sound Vib. **147**, 489–496 (1991)
5. ISO 3745 Acoustics—Determination of sound power levels of noise sources using sound pressure—Precision methods for anechoic and hemi-anechoic rooms (2003)
6. M.P. Waser, M.J. Crocker: Introduction to the two-microphone cross-spectral method of determining sound intensity, Noise Control Eng. J. **22**, 76–85 (1984)
7. H.-E. de Bree: The Microflown: An acoustic particle velocity sensor, Acoust. Aust. **31**, 91–94 (2003)
8. R. Raangs, W.F. Druyvesteyn, H.E. de Bree: A low-cost intensity probe, J. Audio Eng. Soc. **51**, 344–357 (2003)
9. IEC (International Electrotechnical Commission) 1043 Electroacoustics—Instruments for the Measurement of Sound Intensity—Measurements with Pairs of Pressure Sensing Microphones (1993)
10. ANSI (American National Standards Institute) S1.9-1996 Instruments for the Measurement of Sound Intensity (1996)
11. Brüel & Kjær Technical Documentation, Sound intensity software BZ 7205 (1998)
12. G. Rasmussen, M. Brock: Acoustic intensity measurement probe, *Proceeding of Recent Developments in Acoustics Intensity Measurement*, Senlis, France, pp. 81–88 (1981)
13. F. Jacobsen, V. Cutanda, P.M. Juhl: A numerical and experimental investigation of the performance of sound intensity probes at high frequencies, J. Acoust. Soc. Am. **103**, 953–961 (1998)
14. Anonymous: Product Data, Sound Intensity Pair – Type 4197 (Brüel & Kjær, Nærum, Denmark)
15. F. Jacobsen: A simple and effective correction for phase mismatch in intensity probes, Appl. Acoust. **33**, 165–180 (1991)

16. ISO (International Organization for Standardization) 9614-2 Acoustics—Determination of Sound Power Levels of Noise Sources Using Sound Intensity—Part 2: Measurement by Scanning (1996)
17. F. Jacobsen: Intensity measurements in the presence of moderate airflow, *Proceedings of Inter-Noise 94*, Yokohama, Japan, pp. 1737–1742 (1994)
18. F. Jacobsen: Sound intensity measurement at low levels, J. Sound Vib. **166**, 195–207 (1993)
19. F. Jacobsen, H.-E. de Bree: A comparison of two different sound intensity measurement principles, submitted to J. Acoust. Soc. Am.
20. M.C. McGary and M.J. Crocker: Surface intensity measurements on a diesel engine, Noise Control Eng. J. **16**, 27–36 (1981)
21. Y. Hirao, K. Yamamoto, K. Nakamura, S. Ueha: Development of a hand-held sensor probe for detection of sound components radiated from a specific device using surface intensity measurements, Appl. Acoust. **65**, 719–735 (2004)
22. ISO (International Organization for Standardization) 9614-1 Acoustics—Determination of Sound Power Levels of Noise Sources Using Sound Intensity—Part 1: Measurement at Discrete Points (1993)
23. ANSI (American National Standards Institute) S12.12-1992 Engineering Method for the Determination of Sound Power Levels of Noise Sources Using Sound Intensity (1992)
24. M.J. Crocker: Sound power determination from sound intensity – To scan or not to scan, Noise Control Eng. J. **27**, 67 (1986)
25. B. Forssen, M.J. Crocker: Estimation of acoustic velocity, surface velocity, and radiation efficiency by use of the two-microphone technique, J. Acoust. Soc. Am. **73**, 1047–1053 (1983)
26. S. Weyna: The application of sound intensity technique in research on noise abatement in ships, Appl. Acoust. **44**, 341–351 (1995)
27. M. Bockhoff: Sound power determination by intensity measurements in the nearfield of a vibrating panel, *Proceedings of Inter-Noise 85*, Munich, Germany, pp. 1135–1138 (1985)
28. E. Williams, Supersonic intensity on planar sources, J. Acoust. Soc. Am. **104**, 2845–2850 (1998)
29. ISO 15186-1 Acoustics—Measurement of sound insulation in buildings and of building elements using sound intensity—Part 1: Laboratory measurements (2000)
30. ISO 15186-2 Acoustics—Measurement of sound insulation in buildings and of building elements using sound intensity—Part 2: Field measurements (2003)
31. H.G. Jonasson: Determination of emission sound pressure level and sound power level in situ (SP Report 39, 1998)
32. ISO 11205 Acoustics—Noise emitted by machinery and equipment—Engineering method for the determination of emission sound pressure levels in situ at the work station and at other specified positions using sound intensity (2003)

59

Nearfield Acoustical Holography

Sabih I. Hayek

Department of Engineering Science and Mechanics, University Park, PA 16802, USA

Abstract Nearfield acoustical holography (NAH) is a method by which a set of acoustic pressure measurements at points located on a specific surface (called a hologram) can be used to image sources on vibrating surfaces on the acoustic field in three-dimensional space. NAH data are processed to take advantage of the evanescent wavefield to image sources that are separated less that one-eighth of a wavelength.

1 • Introduction

Acoustical holography is a process by which the measurement of the pressure at a set of points on a surface in the vicinity of an acoustically radiating object is analyzed to yield the velocity vector field, the acoustic pressure, and the acoustic intensity map in three-dimensional space. If the measurement points are recorded on a surface that is located at a distance of at least one acoustic wavelength from the actively vibrating surface, then processing these data results in an acoustic image that can only resolve sources that are separated by at least one acoustic wavelength. However, if the measurement points are taken in the nearfield of the source region, then the spatial resolution can be reduced to a fraction of a wavelength, resulting in what is known as nearfield acoustical holography (NAH). The measurement

surface is termed the hologram surface and the area subtending these points constitute the aperture. These measured pressures are then processed by using backward propagators to image the structure's surface pressure as well as the vector velocity and intensity fields. In addition, using forward propagators can image the acoustical field in 3D in the region outside the hologram surface and the source area. The need for a nearfield measurement, well within an acoustical wavelength, ensures that sources which are separated less than one wavelength can be resolved [1–7]. If the measurement hologram plane is located outside one wavelength, the only sources that are separated by one or more wavelengths can be resolved [8].

The imaging of the vector velocity and pressure field starts with recording the acoustic pressure amplitude and phase over an enclosing surface. These are filtered, digitized, and processed to compute the pressure, the velocity vector, and the vector intensity fields on the surface of the vibrator or at any point in 3D space enclosing the vibrating structure.

2 • Planar Nearfield Acoustical Holography

Planar nearfield acoustical holography is a method of imaging the source field of a planar or near planar vibrating structure and the 3D half-space in front of the structure. The hologram is thus a plane of measurement points.

2.1 Imaging Analysis

The imaging analysis requires the use of Fast Fourier Transform (FFT) techniques on the measured acoustic pressure hologram. Defining the Fourier transform integrals as

$$\tilde{u}(k_x,k_y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x,y)e^{-i(k_x x+k_y y)}\,dx\,dy, \tag{1}$$

$$u(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{u}(k_x,k_y)e^{i(k_x x+k_y y}\,dk_x\,dk_y, \tag{2}$$

then the transform of the wave equation on the acoustic pressure $p(x,y,z)$ is

$$\frac{d^2\tilde{p}}{dz^2} + k_z^2\tilde{p} = 0, \text{ where } k_z^2 = k^2 - k_r^2, k_r^2 = k_x^2 + k_y^2, \tag{3}$$

where the $e^{-i\omega t}$ factor has been omitted. For outgoing waves in z,

$$\tilde{p} = Ae^{ik_z z}, \tag{4}$$

where

$$k_z = \sqrt{k^2 - k_r^2} \text{ if } k^2 > k_r^2 \text{ and } k_z = i\sqrt{k_r^2 - k^2} \text{ if } k^2 < k_r^2 \tag{5}$$

and A is an arbitrary complex function of k_x and k_y. Note that the evanescent waves (imaginary k_z) are exponentially attenuated. Choosing the measurement plane (the hologram) to be located at $z = z_H$, the source plane to be located at z_s, z to be at any imaging plane, and requiring that $z_H > z_s > 0$ and $z > z_s$, then the transform of the hologram pressure field $p_H(x,y,z_H)$ is given by

$$\tilde{p}_H(k_x,k_y,z_H) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_H(x,y,z_H)e^{-i(k_x x+k_y y)}\,dx\,dy, \tag{6}$$

so that the transform of the pressure at any field point is

$$\tilde{p}(k_x,k_y,z) = \tilde{p}_H e^{ik_z(z-z_H)}. \tag{7}$$

The inverse transform of the pressure at any field point z can then be written as

$$p(x,y,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{p}_H e^{ik_z(z-z_H)}e^{i(k_x x+k_y y)}\,dk_x\,dk_y. \tag{8}$$

The velocity vector field at z is given by

$$\begin{aligned}
\vec{v} &= \frac{1}{i\omega\rho}\vec{\nabla}p = \frac{1}{2\pi i f\rho}\vec{\nabla}p,\\
v_x &= \frac{1}{(2\pi)^2}\frac{1}{f\rho}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}k_x\tilde{p}_H e^{ik_z(z-z_H)}e^{i(k_x x+k_y y)}\,dk_x\,dk_y,\\
v_y &= \frac{1}{(2\pi)^2}\frac{1}{f\rho}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}k_y\tilde{p}_H e^{ik_z(z-z_H)}e^{i(k_x x+k_y y)}\,dk_x\,dk_y,\\
v_z &= -\frac{1}{(2\pi)^2}\frac{1}{f\rho}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}k_z\tilde{p}_H e^{ik_z(z-z_H)}e^{i(k_x x+k_y y)}\,dk_x\,dk_y,
\end{aligned} \tag{9}$$

where $\int$ is the frequency in hertz and ρ is the density. The intensity field at z is given by

$$\tilde{I} = \frac{1}{2}\Re(p\vec{v}^*), I_x = \frac{1}{2}\Re(pv_x^*),\ I_y = \frac{1}{2}\Re(pv_y^*),\ I_z = \frac{1}{2}\Re(pv_z^*), \tag{10}$$

where v^* is the complex conjugate of v and $\Re$ is the real part of a complex number.

To generate a discrete form of the Fourier transforms used in the previous equations, the spatial improper integrals are converted to definite integrals over the aperture of the measurement plane or surface. Let $2L_x$ and $2L_y$ be the dimensions of the aperture, i.e., $-L_x \le x \le L_x$ and $-L_y \le y \le L_y$. Thus, the spatial increments dx and dy are

$$dx \to \Delta x = \frac{L_x}{M} \text{ and } dy \to \Delta y = \frac{L_y}{N} \tag{11}$$

where the integers M and N represent the discrete number of measurement points in L_x and L_y, respectively, so that

$$\begin{aligned} x &= q\Delta x = \frac{qL_x}{M} \text{ with } -M \le q \le M, \\ y &= r\Delta y = \frac{rL_y}{M} \text{ with } -N \le r \le N. \end{aligned} \tag{12}$$

Since we will also need discrete points in wavenumber-space, we let

$$dk_x \to \Delta k_x = \frac{\pi}{L_x} \text{ and } dk_y \to \Delta k_y = \frac{\pi}{L_y} \tag{13}$$

so that

$$\begin{aligned} k_x &= m\Delta k_x = \frac{m\pi}{L_x} \text{ with } -M \le m \le M, \\ k_y &= n\Delta k_y = \frac{n\pi}{L_y} \text{ with } -N \le n \le N. \end{aligned} \tag{14}$$

The integral Fourier transforms, Eqs. (1) and (2), are now written in their discrete form for u a 2D array of $2M$ by $2N$ elements:

$$\begin{aligned} \tilde{u}(m,n) &= \frac{L_x L_y}{2\pi NM} \sum_{q=-M}^{M} \sum_{r=-N}^{N} u(q,r) e^{-2i\pi\left(\frac{mq}{2M} + \frac{nr}{2N}\right)}, \\ &= \frac{2L_x L_y}{\pi} \mathrm{FFT}[u], \end{aligned} \tag{15}$$

$$\begin{aligned} u(q,r) &= \frac{\pi}{2L_x L_y} \sum_{m=-M}^{M} \sum_{n=-N}^{N} \tilde{u}(m,n) e^{2i\pi\left(\frac{mq}{2M} + \frac{nr}{2N}\right)}, \\ &= \frac{\pi}{2L_x L_y} \mathrm{IFFT}[\tilde{u}]. \end{aligned} \tag{16}$$

The symbols FFT and IFFT represent the typical implementation of the forward and inverse discrete Fourier transform (DFT) using the FFT algorithm. The factor of $(1/4NM)$ in Eq. (15) could also be absorbed into the definition of the inverse transform algorithm if done so consistently. The integral calculation of the pressure at any field point, Eq. (8), can now be written in its discrete form as

$$p(q,r,z) = \mathrm{IFFT}[\mathrm{FFT}[p_H(q,r,z_H)]G(m,n)] \tag{17}$$

with the terms which act like a Green's function collected into G,

$$G(m,n) = e^{2i\pi K_z(z-z_H)} \tag{18}$$

where

$$\begin{aligned} K_z &= \sqrt{K^2 - K_r^2} \text{ if } K^2 > K_r^2, K_r^2 = K_x^2 + K_y^2 \\ &= i\sqrt{K_r^2 - K^2} \text{ if } K^2 < K_r^2, \end{aligned} \tag{19}$$

and

$$K = \frac{1}{\lambda} = \frac{f}{c},\ K_x = \frac{m}{2L_x},\ K_y = \frac{n}{2L_y}, \tag{20}$$

with λ being the acoustic wavelength. The discrete form of the calculation of the velocity field becomes

$$\begin{aligned} v_x &= \frac{1}{\rho f}\text{IFFT}[K_x\ \text{FFT}[p_H]G(m,n)], \\ v_y &= \frac{1}{\rho f}\text{IFFT}[K_y\ \text{FFT}[p_H]G(m,n)], \\ v_z &= -\frac{1}{\rho f}\text{IFFT}[K_z\ \text{FFT}[p_H]G(m,n)], \end{aligned} \tag{21}$$

The discrete form of the calculation of the intensity can be obtained by a straightforward substitution of Eqs (17) and (21) into Eq. (10).

2.2 Processing of Evanescent Wave Data

In wavenumber-space, the evanescent regions will always contain some "noise" due to acoustical noise in the test facility, electronic noise in the data acquisition hardware, quantization errors, and the computer's limited numerical accuracy. Thus, to help ensure that the experimental evanescent data are above this evanescent noise, the hologram is recorded at a distance less than a quarter of a wavelength at the highest frequency used.

When reconstructing evanescent waves at the source ($z = z_s$), the Green's function, Eq. (18), becomes an exponential amplifier that grows for the most extreme case encountered to 284 dB. Most recording systems cannot detect evanescent waves that have been attenuated by more than about 12 dB. Thus, those regions in wavenumber-space that call for a reconstruction amplification of more than about 12 dB represent regions that are probably noise and that should be attenuated rather than amplified. Hence, the

Green's function, Eq. (18), is replaced with a windowed version $G_W(m,n)$ such that

$$\begin{aligned} G_W(m,n) &= (1-\frac{1}{2}e^{-s(1-K_r/K_c)})G(m,n), \text{ for } K_r < K_c, \\ &= \frac{1}{2}G(m,n), \text{ for } K_r = K_c, \\ &= \frac{1}{2}e^{-s(K_r/K_c-1)}G(m,n), \text{ for } K_r > K_c. \end{aligned} \tag{22}$$

The factor s was chosen as 6.666 to yield a relatively sharp slope. The cutoff point of the filter, K_c, was chosen so that at this wavenumber the value of G_W represents an amplification of 12 dB or 3.98. It should be noted that this evanescent filter deletes a significant percentage of wavenumber-space.

2.3 Measurement System and Data Collection and Processing

For accurate reconstruction of the noise source field and for propagating the acoustic farfield, it is best to have a measurement hologram of six to nine times the size of the vibrating surface [9,10]. Using a hologram size smaller than the recommended size of six to nine times the size of the vibrating structure, e.g., same size as the vibrating surface, still allows for accurate imaging of surface sources. However, details outside the aperture area are completely lost, and the acoustic field reconstruction outside that area is also poor and void of details [11].

Data recording can be performed with a movable line array of receivers or with a full rectangular array of stationary receivers. Each receiver would record the amplitude and phase of the acoustic pressure. Each recording channel must be individually calibrated relative to a reference source. Each recording channel contains its own programmable gain amplifier, a filter (flat, full octave, or 1/3 octave), A/D converter (at least 12 bit), and 1024 words of 12-bit memory. The filter center frequency and the A/D converter, which are the same for all channels, can be preset. The sample rate should be four times the experimental frequency so that one can average 1024 samples recorded at each location and then translate the data into a single amplitude and phase. With such filtering and averaging, one can attain repeatability to within 0.1 dB in amplitude and 0.1 degree in phase.

Let each hologram be made of a set of $N \times N$ points. Each set of hologram data is calibrated, then doubled in size to $2N \times 2N$ points by adding zeros in order to avoid FFT wraparound. In wavenumber-space, the hologram size is again doubled to $4N \times 4N$ points by adding zeros in order to obtain a doubling of the

resolution of the final image. A computer program running in batch mode can create the reconstructed pressure, velocity, and intensity fields for a number of selected reconstruction planes.

3 • Cylindrical Nearfield Acoustical Holography

Cylindrical nearfield acoustical holography is a method of imaging the source field on a cylindrical or near-cylindrical vibrating structure and the field outside the structure. The hologram is thus a cylindrical surface of measurement points. The digital algorithms for the reconstructions were derived in various forms by Luce and Hayek [12], Luce [13], and Williams et al. [14], and those from Ref. [13] will be adapted here for completeness.

3.1 Imaging Analysis

Consider a cylindrical vibrating surface whose dimensions are length L and diameter $(2a)$. The acoustic pressure field is recorded on a cylindrical hologram surface of "infinite" length and radius $r_0 > a$. The reconstruction of the acoustic pressure $\tilde{p}$ and the velocity vector $\vec{v}$ at location (r, ϕ, z) are

$$\begin{Bmatrix} \tilde{p}(r,\phi,z) \\ \vec{v}(r,\phi,z) \end{Bmatrix} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{in\phi} \int_{-\infty}^{\infty} \bar{P}_n(r_0,k_z) Q_n(r,r_0,k_r) e^{ik_z z} dk_z \tag{23}$$

where the propagator Q_n is given by

$$Q_n = \begin{cases} \dfrac{Z_n(k_r r)}{Z_n(k_r r_0)} \text{ for pressure} \\ \dfrac{i}{k\rho c}\dfrac{1}{2\pi}(kr)\dfrac{Z_n'(k_r r)}{Z_n(k_r r_0)}\vec{e}_r \text{ for } v_r \\ \dfrac{i}{k\rho c}\dfrac{1}{2\pi}\left(\dfrac{-in}{r}\right)\dfrac{Z_n(k_r r)}{Z_n(k_r r_0)}\vec{e}_\phi \text{ for } v_\phi \\ \dfrac{i}{k\rho c}\dfrac{1}{2\pi}(-ik_z)\dfrac{Z_n(k_r r)}{Z_n(k_r r_0)}\vec{e}_z \text{ for } v_z \end{cases} \tag{24}$$

with primes over functions indicate differentiation with their argument. The radial wavenumber k_r or k_r' are related to the axial wavenumber k_z by

$$\text{(a) } k_r = \sqrt{k^2 - k_z^2} \ |k| > |k_z| \quad \text{(b) } k_r' = \sqrt{k_z^2 - k^2} \ |k| < |k_z| \tag{25}$$

The function Z_n in the propagators Q_n in Eq. (24) are given by

$$Z_n = H_n^{(1)}(k_r r) \text{ for (25a)}, = K_n(k_r' r) \text{ for (25b)} \tag{26}$$

where $H_n^{(1)}$ and K_n are the Hankel function of the first kind and the modified Bessel function, both of order n, representing the propagating and the evanescent waves, respectively. When carrying out the integration in Eq. (23) for the reconstructions, care is taken to divide the integration limits within and outside the range $-k < k_z < k$. To implement this algorithm for a finite-sized hologram, a recording of the pressure on a cylindrical surface (hence called the hologram) of length z_0 and radius r_0 is performed at equally spaced axial and circumferential points whose angular spacing is $\Delta\phi$ and axial spacing Δz given by

$$\Delta\phi = \frac{2\pi}{N} \text{ and } \Delta z = \frac{z_0}{M} \tag{27}$$

where the integer numbers $\bar{N} = 2N + 1$ and $\bar{M} = 2M + 1$ represent the total number of measurement points in the circumferential and the axial scans, based on Nyquist rate for the acoustic wavelength of the measurement, respectively. Performing digital FFT on the hologram $\tilde{p}(r_0, \phi, z)$ gives

$$\hat{P}_n(r_0, k_{z,m}) = z_0 \sum_{q=-M}^{M} \sum_{l=-N}^{N} \tilde{p}(r_0, \phi_l, z_q) e^{-i(\frac{2\pi}{M})qm} e^{-i(\frac{2\pi}{N})ln} \tag{28}$$

at the measured point ϕ_l and z_q. The inverse FFT for the reconstruction of the pressure and velocity vector:

$$\begin{Bmatrix} \tilde{p}(r, \phi_l, z_q) \\ \vec{v}(r, \phi_l, z_q) \end{Bmatrix} = \frac{\bar{M}\bar{N}}{z_0} = \sum_{m=-M}^{M} \sum_{n=-N}^{N} Q_n \hat{P}_n(r_o, k_{z,m}) e^{i(\frac{2\pi}{M}qm)} e^{i(\frac{2\pi}{N})ln} \tag{29}$$

where the discrete value of the axial wavenumber $k_{z,m} = 2\pi m / z_0$ and k_r are computed at the same discrete numbers from Eq. (25)

3.2 Evanescent Wave Processing

Discretization errors introduced by measurements at discrete points and in the performance of the inverse FFT are similar to those in planar holography. In the axial direction, the longer the aperture length, the smaller the axial wavenumber increment. In the circumferential direction, the use of a complete 2π circumferential scan creates an infinite aperture, but an incomplete one could cause a significant error. To reduce the axial error due to insufficient k_z-space sampling, the aperture is increased by adding N zeros beyond each aperture edge, doubling z_0 and halving Δk_z. Similarly, by adding a guard band of zeros to the hologram before performing FFT, one can minimize overlap and aliasing.

Due to the essential reconstruction of the nearfield, which is strongly dominated by the evanescent waves, where the spatial

decay is exponential in the radial direction and obeys a power law in the circumferential direction, the limited response of the transducer in frequency and sensitivity means that the evanescent wave information is lost when it falls below the dynamic range of the system. This was rectified in the planar holography system by the introduction of wavenumber windowing, suggested by Veronesi [15] and Veronesi and Maynard [16]:

$$\begin{aligned} W(k_z) &= 1 - \frac{1}{2} e^{(-s|\frac{|k_z|}{k_{zmax}} - 1|)}; |k_z| \le k_{zmax} \\ &= \frac{1}{2} e^{(-s|\frac{|k_z|}{k_{zmax}} - 1|)}; |k_z| > k_{zmax} \end{aligned} \tag{30}$$

where s is a steepness factor used to tune the filter to achieve the optimum balance between discontinuity resolution and underfiltered evanescent noise. The value of k_{zmax} for the propagator is the axial wavenumber threshold determined by the dynamic range, α of the system in decibels and the reconstruction radius r,

$$k_{zmax} = \sqrt{k^2 + (\frac{0.11 + 0.18\alpha}{r_0 - r})^2} \tag{31}$$

A similar window for the circumferential scan, where s' could be different from s, is

$$\begin{aligned} W(n) &= 1 - \frac{1}{2} e^{(-s'|\frac{|n|}{ak_{zmax}} - 1|)}; |n| \le ak_{zmax} \\ &= \frac{1}{2} e^{(-s'|\frac{|n|}{ak_{zmax}} - 1|)}; |n| > k_{zmax} \end{aligned} \tag{32}$$

where α is again optimally chosen as 12 dB and $s' = s = 0.7$. Similar processing is performed on the hologram as in planar holography.

4 • Spherical and General Surface Nearfield Acoustical Holography

For the spherical and general surface nearfield acoustical holography, Veronesi [15] and Veronesi and Maynard [16] have given a thorough treatment for recording and processing acoustical holograms. Williams [17] gives a complete treatment of this subject.

Acknowledgments The author wishes to express his appreciation for the contributions of his colleagues K.E. Eschenberg, T.W. Luce, J.D. Maynard, E.J. Skudrzyk, W.A. Veronesi, and E.G. Williams, all of whom have collectively contributed tremendously to the field of NAH.

References

1. D.L. Van Rooy, "Digital Untrasonic Wavefront Reconstruction in the Near Field", IBM Publication No. 320.2402 (19 May 1971)
2. A.L. Boyer, et al., "Reconstruction of Ultrasonic Images by Backward Propagation", *Proceedings of Acoustical Holography*, Vol. 3, Plenum Press, New York, p. 333 (1971)
3. R.L. Cohen, "Digital Wavefront Reconstruction for Acoustical Applications," PhD Thesis, The Pennsylvania State University, University Park, PA (1979).
4. E.G. Williams, J.D. Maynard, and E. Skudrzyk, "Sound Source Reconstructions Using a Microphone Array," *J. Acoust. Soc. Am.*, **68**, pp. 340–344 (1980).
5. E.G. Williams, H.D. Dardy, and R.G. Fink, "Nearfield Acoustical Holography Using an Underwater, Automated Scanner," *J. Acoust. Soc. Am.*, **78**(2), 789–798. (1985)
6. J.D. Maynard, E.G. Williams, and Y. Lee, "Nearfield Acoustic Holography: 1 Theory of Generalized Holography and the Development of NAH," *J. Acoust. Soc. Am.*, **78**(4), 1395–1413 (1985)
7. E.G. Williams, H.D. Dardy, and R.G. Fink, "A Technique for Measurement of Structure-Borne Intensity in Plates," *J. Acoust. Soc. Am.*, **78**(6), 2061 (1985)
8. H.M. Smith, *Principles of Holography*, Wiley-Interscience, New York (1969)
9. K.E. Eschenberg and S.I. Hayek, "Measurement of Submerged-Plate Intensity Using Nearfield Acoustic Holography," *Proceedings of Inter-Noise '86*, June 1986, pp. 1229–1234.
10. K.E. Eschenberg and S.I. Hayek "Imaging the Nearfield of a Submerged Plate Using Acoustical Holography," *Proceeding of Symposium on Acoustical Imaging*, Vol. 15, held in Halifax, Canada, pp. 659–673 (June 1987).
11. S.I. Hayek and T.W. Luce, "Aperture Effects in Planar Nearfield Acoustical Imaging," *ASME Journal of Vibration, Acoustics, Stress, and Reliability in Design*, **110**, 91–96 (1988).
12. T.W. Luce and S.I. Hayek, "An Examination of Aperture Effects in Cylindrical Nearfield Holography," *Proceedings of Symposium on Acoustical Imaging*, Vol. 16, held in Chicago, IL, pp. 267–276 (June 1988).
13. T.W. Luce "The Development and Implementation of Digital Acoustic Nearfield Holography in a Circular Cylindrical Coordinate System," M.Sc. Thesis, Penn State University, University Park, PA (August 1992).
14. E.G. Williams, H.D. Dardy, and K.B. Washburn, "Generalized Nearfield Acoustical Holography for Cylindrical Geometry: Theory and Experiment," *J. Acoust. Soc. Am.*, **81**(2), 389–407 (1987).
15. W. Veronesi, "Computational Methods in Nearfield Acoustical Holography (NAH)," Ph.D. Thesis, The Pennsylvania State University, University Park, PA (1986).

16. W.A. Veronesi and J.D. Maynard, "Nearfield Acoustical Holography (NAH) II. Holographic Reconstruction Algorithms and Computer Implementations," *J. Acoust. Soc. Am.*, **81**(5), 1307–1322 (1987).
17. E.G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London (1999)

60

Actuator Design and Deployment

Colin H. Hansen

School of Mechanical Engineering, University of Adelaide, Adelaide, SA 5005, Australia, e-mail: chansen@mecheng.adelaide.edu.au

1 • Introduction

In structural acoustics and vibration, actuators are used to excite a structure for the purposes of modal analysis or active noise and vibration control. Here, the design of various actuators is discussed followed by a discussion of how their locations may be optimised for best effect.

There are two main types of actuator: those that require an external power source and those that require no external power source. Those that require no power source act to apply damping to a structure on which they are mounted and are often referred to as vibration absorbers. A vibration absorber essentially consists of a mass connected to a vibrating structure by a spring. The vibration of the added mass on the spring absorbs energy from the structure at frequencies around the resonance frequency of the absorber. A vibration absorber may also take the form of an array of small masses distributed through an acoustic foam blanket so that they act as mini-absorbers with the resilience of the foam providing the spring element (Kim et al. 2005). Embedding the masses at different depths in the foam allows the absorbing blanket to provide vibration attenuation over a range of frequencies. The other generic type of actuator has to be driven by a power source at a frequency or frequencies at which it is

desired to excite the connected structure. Only this latter class of vibration actuators will be discussed here. The design of a range of different actuator types will be discussed in detail in the following sub-sections.

2 • Actuator Design

2.1 Hydraulic Actuator

Hydraulic actuators generally consist of a piston in a hydraulic cylinder which has hydraulic fluid openings at each end. Introduction of high-pressure fluid into one end of the cylinder causes the piston to move to the other end, expelling fluid from the opening at that end. Switching the supply of high-pressure fluid to the other end of the cylinder causes the piston to move in the opposite direction. The most common method to alternate the hydraulic supply from one end of the cylinder to the other is to use a servo-valve, which consists of moveable spool connected to a solenoid.

Applying voltage to the solenoid coil causes the spool to move, which in turn results in the opening and closing of valves that are responsible for directing hydraulic fluid to either end of the hydraulic cylinder (see Figure 1). Note that the servo-valve spool is usually maintained in its inactive position with springs at either end.

Servo-valves can be quite expensive, ranging from $100 to $3000, depending on the upper frequency at which they are required to operate, which can be 3–20 Hz for an active vehicle suspension or 150 Hz for structural vibration excitation.

One advantage of hydraulic actuators is their large displacement and high force-generating capability for a relatively small size. Disadvantages include the need for a hydraulic power supply which can be inherently noisy and non-linearities between the

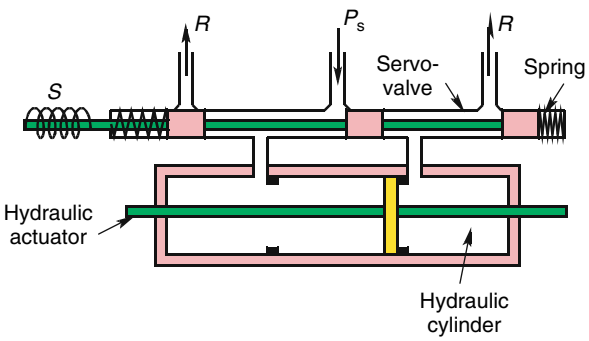


FIGURE 1 Hydraulic actuator schematic.

servo-valve input voltage and the hydraulic actuator force or displacement output. It is extremely important that the servo-valve be mounted as closely as possible to the hydraulic cylinder to minimise loss of performance, especially at high frequencies (above about 20 Hz). For best results at high frequencies, the servo-valve should be incorporated in the actuator assembly as was done by King (1988).

In an attempt to make the actuator motion correspond more closely to the solenoid drive current, some manufacturers have developed two-stage servo-valves, in which a torque motor is used to operate a valve, which in turn directs hydraulic fluid to one side or the other of the servo-valve spool.

2.2 Pneumatic Actuator

Pneumatic actuators are very similar in operation to hydraulic actuators, except that the hydraulic fluid is replaced by air. One advantage of pneumatic actuators is that they can use the same air supply as passive air springs, which may be mounted in parallel with them. Pneumatic actuators may also be the preferred option in cases where an existing compressed air supply is available. The major disadvantage of pneumatic actuators is their relatively low bandwidth (less than 10 Hz) due to the compressibility of the air.

2.3 Inertial Actuator

This type of actuator consists of a mass, which is free to slide horizontally or vertically along a track, which in turn is attached to the structure to be excited. Some detail design effort is necessary to prevent the mass rattling on its track, thus producing vibration at frequencies other than the drive frequency. The mass is usually a magnet and is accelerated using an alternating electromagnetic field (see Figure 2a), so in one sense this type of actuator is similar to an electrodynamic shaker, which will be described in the next section. The mass is also usually constrained by springs as shown in the figure. In some clever designs, these springs may be replaced by a magnetic field generated by using magnets in place of the end caps and polarised so that they are opposite to the polarisation of the nearest surface of the sliding mass.

Another inertial actuator configuration is shown in Figure 2b. For this actuator, the structure is attached to the central rod and the thin steel or fibreglass shim provides stiffness as the outer magnet vibrates in the vertical direction.

An inertial actuator has a resonance frequency associated with the moving mass and the stiffness of its support springs. If excited at its resonance frequency, an inertial actuator can impart quite large excitation forces into the attached structure, and at frequencies

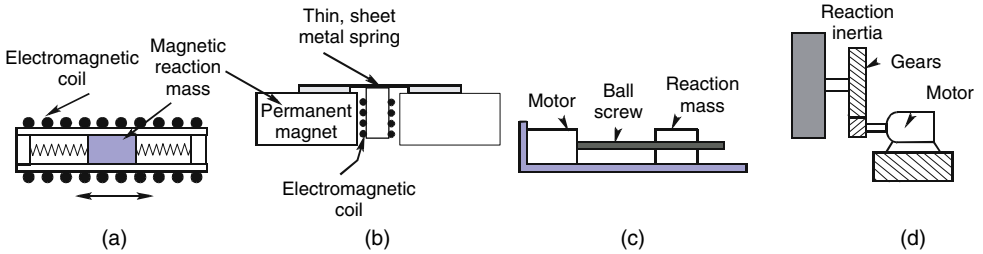


FIGURE 2 Inertial actuator designs. (a) linear sliding mass; (b) linear sprung mass; (c) linear screwed mass; (d) rotary mass.

above and below the resonance frequency, the available force diminishes. For active vibration control, the excitation frequency band should not include the resonance frequency because the rapid change in phase of the generated force relative to the actuator input signal causes stability problems. Inertial actuators are typically used to excite structures over the frequency range from 20 to 1,000 Hz. Their main advantages are their low cost and the absence of a need for a support structure, as the actuators are attached directly to the structure to be excited.

The moving mass of an inertial shaker can be accelerated by means other than an electromagnetic field, for example, by connecting it to a ball screw (see Figure 2c) (Clark et al. 1990). The proof mass results in a reaction in the ball screw or electromagnetic stator, which is attached to the structure to be controlled.

Another type of inertial actuator is one that imparts moments rather than forces to a structure by control of a rotating mass which results in reaction moments on the support structure (Figure 2d).

2.4 Electrodynamic and Electromagnetic Actuators

Electrodynamic actuators or shakers consist of a moveable core (or armature) to which a cylindrical electrical coil is fixed. The core and coil move back and forth in a magnetic field which is generated by a permanent magnet (smaller shakers) or by a d.c. current flowing through a second coil fixed to the stationary portion of the shaker (stator), as shown in Figure 3a. The moveable core is supported on mounts that are very flexible in the axial direction in which it is desired that the core move and rigid in the radial direction to prevent the core contacting the outer armature. To maximise lightness and stiffness, the core is usually constructed of high-strength aluminium alloy and is usually hollow. Mechanical stops are also usually provided to prevent the core from being over driven. In some cases, the driving coil is air cooled. These types of actuator are used to excite structures in the frequency range from 10 to 5,000 Hz and the

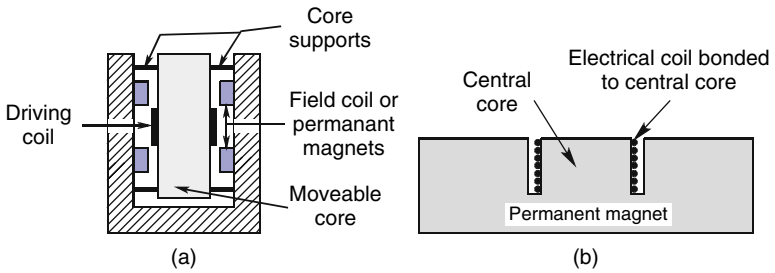


FIGURE 3 (**a**) Electrodynamic shaker, (**b**) Electromagnetic shaker.

amount of excitation force available is directly related to the size of the actuator which can range from 0.5 kg to many tons. The major disadvantages of this type of actuator are their high cost, their requirement for accurate alignment when being attached to a structure and their need for a support structure.

When a sinusoidal voltage is applied to the drive coil, the polarity and strength of the drive coil magnetic field changes in phase with the applied voltage and so then does the force of attraction between the field coil (or permanent magnets) and the driving coil, resulting in axial movement of the core (or armature).

Electromagnetic actuators are similar in construction to electrodynamic shakers except that the inner core as well as the armature is fixed (see Figure 3b). It is constructed by surrounding an electrical coil with a permanent magnet. Supplying a sinusoidal voltage to the coil results in the production of a sinusoidally varying magnetic field which can be used to shake ferro-magnet structures, or structures made from other materials, if a thin piece of shim steel is bonded to them. Clearly, the electromagnetic actuator must be mounted on a rigid fixture, preferably not connected to the structure to be excited.

If the permanent magnet were not used, the attraction force between the coil and the structure being shaken would vary like a rectified sine wave, resulting in most of the excitation energy being at twice the coil driving frequency.

Electromagnetic drivers such as those just described are easily and inexpensively constructed by bonding the coil from a loudspeaker driver into the core of the permanent magnet where it is normally located in its at rest condition.

2.5 Magnetostrictive Actuators

The phenomenon of giant magnetostriction in rare earth, Fe_2 compounds is well known with the most effective alloy being referred to as terfenol, having the composition $Tb_{0.3}$ $Dy_{0.7}$ $Fe_{1.93}$. The subscripts refer to the atomic proportions of each element (terbium, dysprosium and iron). Multiplying these by the atomic

weights gives the weight proportions (18%, 42% and 40%, respectively). The notable properties of this material are its high strain ability (25 times that of nickel, the only other commonly used magnetostrictive material and 10 times that of piezo-electric ceramics – see next section) and high energy densities (100 times that of piezo-ceramics). Thus, terfenol-D can produce high force levels and high strains relative to other expanding materials. The maximum possible theoretical strain is 2,440 μstrain, and 1,400 μstrain is easily achievable in practice. The properties of terfenol are discussed in detail in the literature by Hiller et al. (1989) and Moffett et al. (1991a, b). A typical commercially available terfenol actuator is illustrated in Figure 4 (Anonymous 1990).

Examination of the actuator illustrated in Figure 4 shows the use of high-permeability iron end caps to help reduce fringing of the magnetic flux lines at the rod ends. Flux density is also improved by the coil extending past the ends of the terfenol rod.

Terfenol actuators are usually characterised by a resonance frequency in the few kilohertz range and generally have a reasonably flat frequency response at frequencies below about half the resonance frequency. The actual resonance frequency depends on the mass driven and the stiffness of the terfenol rod. Longer, thinner drives and larger masses result in lower resonance frequencies.

The force output for a clamped actuator may be calculated using $F = EA\Delta L/L$, where E is the elastic modulus ($3.5 \times 1010\,\mathrm{N/m^2}$), A is the cross-sectional area of the terfenol rod ($\mathrm{m^2}$) and $\Delta L/L$ is the strain. Thus, a 6-mm diameter rod operating in a magnetic field capable of producing $\pm 500\,\mu$strain about the bias value of +500 can produce a force of $\pm 500\,\mathrm{N}$.

The main advantage of terfenol is its high force capability for a relatively low cost, weight and size.

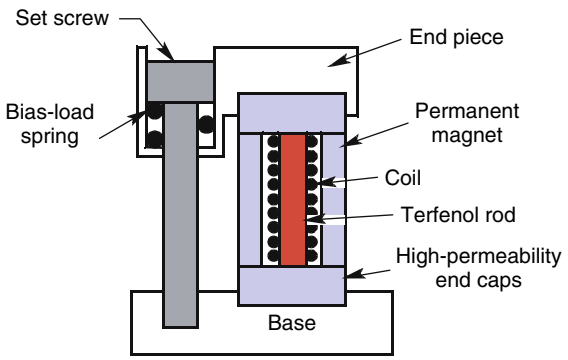


FIGURE 4
Cylindrical terfenol actuator.

The main disadvantage of Terfenol is that it is very brittle and must be carefully handled. Its tensile strength is low (100 MPa) although its compressive strength is reasonably high (780 MPa). Another important disadvantage is the low displacement capability, which would be a problem in low-frequency vibration control (below 100 Hz). One way around this limitation is to use displacement expansion devices, which rely on the lever principle to magnify the displacement at the expense of reduced force output. Another important disadvantage associated with magnetostrictive actuators is the hysteresis inherent in terfenol, which results in a non-linear actuation force for a linear voltage input. This results in excitation frequencies being generated that are not related to the voltage input frequencies. Finally, as for electrodynamic, electromagnetic, hydraulic and pneumatic shakers, a terfenol actuator needs a supporting structure to react against so that it can impart its force to the structure to which it is attached.

2.5.1 Magnetic Bias

As application of a magnetic field of any orientation will cause the terfenol rod to expand, a d.c. magnetic bias is necessary if the terfenol rod expansion and contraction is to follow an a.c. input signal into the coil. Thus, a d.c. magnetic bias supplied either with a d.c. voltage signal applied to the coil or with a permanent magnet is required so that the terfenol rod expands and contracts about a mean strain (usually about 500 μstrain). Use of d.c. bias current is characterised by problems associated with drive amplifier design and transducer overheating; thus, it is more common to use a permanent magnet bias. The cylindrical permanent magnet bias shown in Figure 4 also acts as the flux return path for the exciter coil.

2.5.2 Mechanical Bias or Prestress

To optimise performance (maximise the strain ability of the actuator), it is necessary to apply a longitudinal prestress between 7 and 10 MPa. The prestress ensures that the magnetically biased actuator will contract as much as it expands when a sinusoidal voltage is applied to the drive coil. It is essential that the prestress arrangement also be highly compliant to avoid damping the drive element. The mechanical bias also helps to mate components and allows transfer of forces with low loss.

2.6 Shape Memory Alloy Actuators

A shape memory alloy is a material which, when plastically deformed in its low-temperature (or martensitic) condition, and the external stresses removed, will regain its original (memory) shape when heated. The process of regaining the original shape

is associated with a phase change of the solid alloy from a martensite to an austenite crystalline structure.

The most commonly used shape memory alloy in active control systems is nitinol (an alloy of approximately equi-atomic composition of nickel and tin). Room temperature plastic strains of typically 6%–8% can be completely recovered by heating the material beyond its transition temperature (which ranges from 45 to 55°C). Restraining the material from regaining its memory shape can yield stresses of up to 500 MPa (for an 8% plastic strain and a temperature of 180°C – See Rogers 1990). On being transformed from its martensite to its austenite phase, the elastic modulus of nitonol increases by a factor of three (from 25 to 75 GPa) and its yield strength increases by a factor of eight (from 80 to 600 MPa).

The attractiveness of nitinol over other shape memory alloys for active control is its high resistivity, which makes resistive heating feasible. This means that nitinol can be conveniently heated by passing an electrical current through it.

Nitinol is commercially available in the form of thin wire or film that can be used for the active control of structural vibration by either embedding the material in the structure or fixing it to the surface of the structure. Before being embedded or fixed to the surface of the structure, the nitinol is plastically elongated. When used as fibres in composite structures, the nitinol fibres must be constrained so that they cannot contract to their normal length during the high temperature curing of the composite.

The upper frequency limit for structural excitation with nitinol is about 10 Hz, due to the difficulty in cooling the nitinol between cycles.

2.7 Piezo-Electric (Electrostrictive) Actuators

This type of actuator may be divided into two distinct categories: the thin plate or film type and the stack (shaker) type. The thin type may be further categorised into piezo-ceramic plates (PZT) or piezo-electric film (PVDF). The thin type is usually bonded to the structure it is to excite and generates excitation by imparting a bending moment to the structure. The stack type is used in a similar way to an electrodynamic shaker or magnetostrictive actuator, generally applying to the structure a force distributed over a small area. The thin type and stack type actuators will be considered separately in the following discussion. Both types of actuator generate forces and displacements when subjected to an electric field—hence their alternative description, electrostrictive.

2.7.1 Thin Actuators

A detailed review of work on thin piezo-actuators was done by Smits et al. (1991), and the reader is referred to this for a detailed survey of the uses of this material.

The behaviour of piezo-ceramic plate and PVDF film actuators is similar (although their properties differ significantly) and both actuator types can also act as sensors.

For thin actuators, the extension as a result of an applied voltage is in a direction normal to the direction of polarisation as shown in Figure 5. If the piezo-element is bonded to a structure, the increase in length will be less and a resulting bi-directional longitudinal force will be applied to the structure surface which will result in a moment induced in the structure.

Although PZT actuators can be bonded directly on to structures using epoxy adhesive, their brittle nature makes this a task requiring considerable skill, especially on curved surfaces. Usually, a thin layer of epoxy is first bonded to the surface under a layer of glass which has release agent sprayed on it. When dry, the glass is removed, leaving a smooth surface on to which the PZT is bonded. Some researchers prefer to order the PZT from manufacturers already bonded to a thin brass shim. This shim is then bonded (with a lot less trouble) directly to the structure to be excited.

2.7.2 Thick Actuators

These actuators are generally thick compared to their other dimensions and act by expanding across their thickness in the direction of polarisation. The voltage is applied to the polarised surfaces as for the thin actuators, but in this case, the expansion of the thickness rather than expansion in the other dimensions is used to perform useful work. Stack actuators are made by bonding many thin actuators together as shown in Figure 6. Commercially available stack actuators usually consist of up

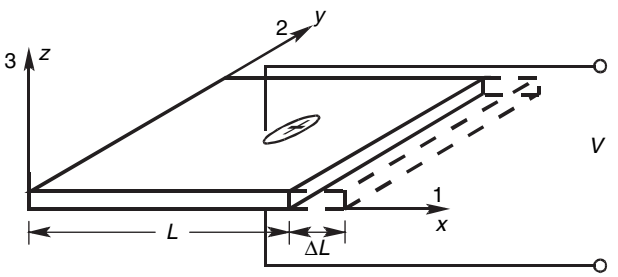


FIGURE 5 Thin PZT actuator showing lateral extension as a result of a voltage applied across its thickness.

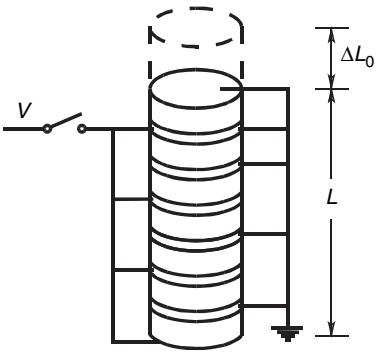


FIGURE 6 PZT stack.

to 100 thin actuators (0.3–1 mm thick) stacked together, with diameters ranging from 15 to 50 mm.

One problem with piezo-electric actuators (both the thin sheet type and the stack type), which is also experienced by magnetostrictive actuators, is their hysteresis property, which means that the expansion as a function of a given applied electric field (or voltage) is dependent on whether the field level is rising or falling. It is also a non-linear function of the applied electric field in either direction. This results in non-linear excitation of a structure attached to one of these actuator types. Thus, if the exciting voltage used to drive the actuator is sinusoidal, the resulting structure motion will contain additional frequencies. In practice, a strong signal at the first harmonic of the exciting frequency is usually found.

2.8 Novel Actuator Configurations

One problem with applying piezo-ceramic stacks, magnetostrictive rods, hydraulic or electrodynamic actuators to a structure in the traditional way is the need for a reaction mass and support for the actuator, which is often inconvenient to implement. However, the need for this reaction mass can be eliminated by using a more imaginative actuator configuration which applies control bending moments rather than control forces to the structure. For stiffened structures such as aircraft fuselages or submarine hulls, implementation of the actuator configurations is even more convenient.

Two possible actuator configurations that need no reaction mass are illustrated in Figures 7a and b. The first figure illustrates the arrangement where stiffeners already existed on the panel, and the second figure shows an alternative arrangement, which might be used if no prior stiffeners existed. The configurations

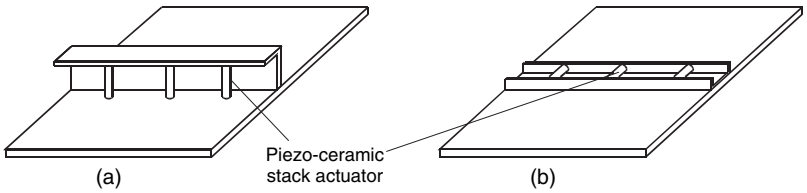


FIGURE 7 Actuator configurations that eliminate the need for a reaction mass or support.

shown for panel vibration control can easily be extended to cylindrical structures such as aircraft fuselages and submarine hulls.

2.9 MEMS Actuators

MEMS actuators rely on electrostatic force to provide their actuation. Electrostatic MEMS actuators may have their electrodes configured either as comb-drive or as parallel-plates. The comb-drive configuration is suitable for applications that require a relatively large actuation stroke and small actuation force, whereas the parallel-plates configuration is suitable for applications that require a large force output and small actuation stroke. In addition, the comb-drive actuation configuration has a linear force output and is stable under all actuation conditions, whereas the parallel-plates configuration has a non-linear force output and may become unstable when large actuation voltages are applied. Electrical isolation is another challenge for designing a MEMS actuator consisting of parallel-plates. The usual way of achieving electrical isolation is to structurally separate the parts not at the same potential and to mount everything to a non-conducting substrate. The disadvantages of having to include a non-conducting substrate and mechanical separation in the actuator design have been overcome by Chen et al. (2002), who have developed a parallel-plate micro-actuator suitable for disk drive suspension active vibration isolation, and they used an electrical isolation MEMS fabrication technique to maximise the actuation force.

3 • Actuator Deployment Using Genetic Algorithms

When undertaking modal analysis, it is important that the set of actuators used to excite the structure is capable of exciting all modes of interest. That means that at least one actuator needs to be located well away from a modal node (and preferably close to an antinode) for each mode that is of interest. One way of ensuring this is to use many actuators. Another way is to have

some idea of the basic mode shapes to begin with and locate the actuators strategically.

When applying active vibration control or active sound radiation control by applying actuators to a structure, the location of the actuators is of great importance, unless great numbers are used. With active control, it is not so easy to determine the best actuator locations if noise or vibration is to be reduced over a reasonably wide frequency band (one to two octaves), especially if multiple actuators are to be used. The complex multi-modal nature of most practical active sound and vibration control applications involving large structures and multiple control sources mean that an exhaustive search of all possible control source configurations is usually impractical. Even gradient-based search methods, which move from point to point in the direction of maximum sound reduction (commonly referred to as a hill-climb search), are ineffective for searching for a global optimum arrangement of control source locations across many local optima. One way to optimise the placement of actuators is to use genetic algorithm (GA) optimisation and this is the subject of the following discussion.

The GA may be regarded as a guided random search that begins with the arbitrary specification of a population, typically consisting of between 40 and 100 solutions. Finding the optimal solutions involves "breeding" new solutions from the original population which involves three processes: fitness evaluation, selection of appropriate "parents" and use of the "parents" to breed new population members ("children"). Fitness evaluation involves the assessment of the particular individual solution in terms of its ability to minimise the cost function in the active noise control system. For optimisation of control source locations, a particular solution is one configuration of all the control sources, and for the control filter weight optimisation, one solution is a particular set of filter weight values.

Selection (used only for the filter weight optimisation process) involves removing a proportion of the population based on their lack of "fitness" (usually about 30% of the total). Breeding is the use of the remainder of the population to generate a new population for fitness testing. The choice of parents for breeding is based on probability considerations with fitter individuals more likely to be chosen; this will be discussed in more detail later. The characteristics of the two individuals are combined randomly to produce a third individual ("child") that becomes a member of the next generation. The breeding process continues until there are enough new individuals created to completely replace all members of the breeding population (except the fittest

one when an "elitist" model is used). The breeding cycle then repeats until it is clear that further new high-fitness solutions are unlikely to be found.

To achieve optimal convergence behaviour, each GA must be "tuned" to suit the particular problem at hand although the approach described here is generally applicable to most actuator placement problems. The search behaviour can be altered significantly by varying the parameters controlling the algorithm such as crossover probability, mutation probability, selection probability distribution and population size. The influence that these parameters have may be described in terms of their effect on selective pressure and population diversity (Whitley 1989). Selective pressure is defined as the bias that exists towards the high-fitness members of the breeding population during parent selection. Population diversity is a measure of the degree to which strings in the breeding population are distributed throughout the solution space. Selective pressure and population diversity are inversely related. For example, as selective pressure increases, the focus of the search shifts towards the high-fitness members of the breeding population, often at the expense of excluding low-fitness solutions from the breeding process. This represents a "hill-climb" procedure in which population diversity is commonly lost resulting in a narrow band of similar string types. A low selective pressure means that more members of the breeding population are included in the search process, resulting in a wider search with more diverse string combinations produced. However, in this case, the search slows down due to lack of focus on the high-fitness individuals.

The key to a successful global optimisation search is maintaining the correct balance between selective pressure and population diversity. In the following sections, various methods for controlling selection pressure and population diversity are described, with the aim of improving search behaviour such that optimum performance is achieved. Questions such as the use of an individual more than once in generating the next generation of "children" will also be addressed. The breeding cycle is usually stopped at a point when it becomes clear that no significant improvement in the best solution is likely.

3.1 Coding Selection for Search Variables

GAs are sufficiently robust to exhibit favourable convergence characteristics for a range of different coding alternatives. However, there may be significant benefits in selecting one coding type over others in a given problem configuration due to improved convergence and performance characteristics. Here, various coding alternatives are reviewed with the appropriateness

of each discussed with a view to obtaining best empirical search performance in a generalised multi-modal optimisation problem.

The first coding type considered is that of a single number (integer or floating point) coding scheme (Katsikas et al. 1993; Tsahalis et al. 1993; Yao et al. 1993) so that each control source location is represented by a single number proportional to its coordinate values or weight value, respectively. In this case, each position at which actuators may be located (in the problem) is assigned a unique integer location number. This approach lends itself directly to finite element modelling with each node in the model represented by a unique integer. The coding is then formed by concatenating the location numbers representing each actuator position into a single numerical string, of length equal to the number of control sources (or number of filter weights). Although maintaining a constant number of control sources is inherent to this coding method, duplicate actuator positions within individual strings are possible, requiring additional constraints to ensure that all actuator locations (in a single string) are distinct.

The application of the single number (numerical) coding method is not constrained to locations represented by integer numbers. The possibility arises for the real floating point values of coordinate locations (representing each control source position) to be used directly in the string structure. However, to ensure that an even distribution of control source locations is achieved (throughout the problem space), the continuous variable range defining the control source coordinate locations must be discretised into a finite number of equidistant values suitable for finite difference models. The real values of these discretised variables could then be used directly in the string structure in a similar manner to that used for the integer values.

A second coding scheme, which has been used in the past for actuator placement optimisation, consists of using a binary string to trace the status of each possible actuator placement location in the problem space. With each location assigned a unique position in the binary string, an actuator present is assigned a bit value of "1" in the string, with the remaining (empty) positions assigned a bit value of "0". This scheme is not really practical for the problems of interest here because large binary string lengths would be required to trace a relatively low number of actuators among a large number of potential actuator locations in the placement problems to be considered. However, the approach may be useful for the actuator placement problem when the potential number of actuator locations is relatively small.

A third coding scheme has been proposed in which a multivariable binary string (MVBS) is used to represent the positions of each actuator configuration. In this coding scheme, each individual actuator location is mapped to a n-bit binary string. To represent M actuator positions, the length of the binary string n must satisfy the requirement: $2n \geq \mathrm{M}$. Once actuators are assigned a unique binary number representing their location, the binary numbers themselves are concatenated into a single (combined) binary string. To represent continuous variables in this manner, the variable range is discretised into $2n$ separate equidistant values, enabling each variable value to be mapped to a distinct n-bit binary number.

3.2 Parent Selection

Two selection processes are carried out during implementation of the GA: the choice of individuals to be removed or kept and the choice of parents to be used for breeding the next generation. Both processes are usually implemented using a simulated (and biassed) roulette wheel where each segment on the roulette wheel corresponds to an individual with the size of the segment proportional to the probability of the individual being chosen (selection probability). Selection probabilities are assigned such that low–fitness individuals are more likely to be removed and such that high-fitness individuals (those that result in greatest reduction of the cost function – noise or vibration) are more likely to be chosen as parents for breeding. Selection without replacement is used for removal, where once an individual is chosen, it is removed from the roulette wheel. For parent selection, replacement is used so that the selected parent is not removed from the roulette wheel after being selected. This allows the entire population to be available for breeding of each new individual in the next generation. In addition, when an "elitist" model is used, the implication is that the parent with the best fitness is carried into the next generation population unchanged.

The selection probability of member i from a total population of n_{pop} individuals for removal (or killing) is

$$P_{\mathrm{r}}(i) = \frac{1.0 - f_i}{n_{\mathrm{pop}} \sum_{j=1} f_j}, \tag{1}$$

where f_i is the assigned string fitness. The selection probability of member i for breeding is

$$P_s(i) = \frac{f_i}{\sum_{j=1} f_j}. \qquad (2)$$

Concerns over stochastic errors involved with the roulette wheel selection procedure have led to proposals of many alternative parent selection schemes (Goldberg 1989). The parent selection scheme known as "stochastic remainder selection without replacement" has been widely accepted as providing an adequate, more accurate alternative to roulette wheel selection. This method ensures that the expected number of breeding opportunities allocated to strings in each generation correspond as closely as possible to the breeding probability distribution. All parent strings required for one generation are preselected, in proportion to the breeding probability distribution, into a separate parent selection population (of size $2n_{\text{pop}}$). To do this, an expected value parameter e_i, for member i, is defined as

$$e_i = \frac{f_i}{\bar{f}}, \qquad (3)$$

where $\bar{f}$ is the mean fitness value of the breeding population. The truncated integer value of e_i (if $e_i \geq 1$) indicates the number of whole copies that each string is allocated in the selection population.

Once completed, the fractional value of e_i (regardless of integer value) is used as a success probability in a weighted coin toss procedure. If successful, a string is allocated a whole copy in the selection population and removed from further contention. This process continues until all available places in the selection population have been filled. Then, when choosing two parent strings for breeding, all that is required is the random selection of two parent candidates from the selection population (without replacement).

It is clear from the preceding equations that the numerical value of each probability and hence the difference between the fitness of various individuals will depend on the fitness measure. If there is not much numerical difference between fitness values in the population, the search may stagnate in its ability to find new solutions, due to a general lack of direction in allocating reproductive opportunities. Secondly, an unusually high-fitness individual may be found early in a search, causing a disproportionate number of reproductive opportunities to be allocated to

it. This usually results in that individual dominating the breeding population within a few generations, leading to search stagnation due to the lack of population diversity, which in turn could result in the search terminating in a local optimum.

To overcome these problems, more control is required over the way in which the selection probability distribution is allocated. A common method used for this purpose is linear scaling (Goldberg 1989). The principle of this scaling procedure is to scale each string fitness value such that the differences in all fitness values are enhanced (or reduced) to maintain an appropriate bias towards high-fitness individuals in a population (selective pressure). Therefore, the relatively minor differences in fitness values of a near-homogeneous population may be scaled up such that adequate selective pressure is maintained throughout a search.

Even with fitness scaling procedures, the selection probability distribution may not always be appropriate for the desired search behaviour. A more direct way of controlling selective pressure has been proposed (Whitley 1989), where the selection probability is assigned to members of the breeding population according to the rank position of their fitness values rather than in proportion to their absolute fitness values. Of course, different probability distributions are used for removal selection and breeding selection as shown in Figure 8. The use of rank selection has the effect of removing problems associated with either high-fitness string domination or search stagnation as a result of using fitness proportional schemes.

Altering the distribution of the selection probability determines the nature of the search performed. A linear probability distribution ensures that the median assigned fitness corresponds to the median rank position of the population. A non-linear probability distribution creates a bias towards higher fitness

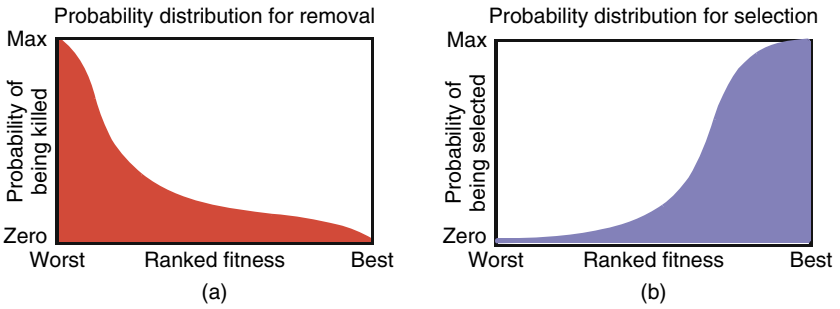


FIGURE 8 Typical probability distributions for (**a**) killing and (**b**) breeding.

individuals as shown in Figure 8 and will tend to display hill-climbing characteristics. It is generally advantageous to use moderate selective pressure at the start of the search to enable a wide search for high-fitness solutions. However, when the search is converging on a final solution, a hill-climb solution is usually more appropriate to speed up the process by allocating a higher selective pressure.

One of the problems encountered with the standard implementation of the GA is that individuals with higher fitness values are not automatically guaranteed survival into the next generation. As the composition of the breeding population dynamically changes throughout the duration of the search, it is likely that some good solutions may not be reproduced. Therefore, the potential exists for best-fitness solutions found in a search to be lost during the generational transition, requiring several more generations to recover. To overcome this problem, an elitist model (Goldberg 1989) was implemented (in both applications discussed here) to ensure survival of the previous best string into the next generation, if that string was not already present. In this manner, the focus of the search was maintained on the best performance individual throughout the search process. One drawback, however, is that occasionally the search may focus exclusively on the best performance solution found, at the expense of undertaking a wider search for other potentially better solutions.

The steady state genetic algorithm (SSGA) (Zimmerman 1993; Whitley 1989; Syswerda 1989; Whitley and Hanson 1989) is a variation of the standard GA and has been developed in an attempt to minimise the potential loss of vital search information from generation to generation. The main difference with this type of algorithm is that the breeding population is sorted in order of fitness magnitude and retains the highest fitness solutions found from the entire search undertaken so far. Reproduction of new child strings (i.e. crossover and mutation) is the same as in the previous GA case. However, after each child string is created and assigned a fitness value (by the objective function), its relative performance is immediately compared with those in the existing breeding population. If the fitness exceeds that of the lowest fitness string in the population, then the newer child string is inserted at the correct rank position in the population (displacing the low-fitness solution from the bottom) and becomes an immediate active member of the breeding process. This contrasts significantly with the original GA concept, where each child produced must wait until the next generation before it is used as a parent. Also, the SSGA eliminates the need to apply the elitist model (discussed above) as the retention

of high-fitness solutions is inherent. The steady state GA using integer coding has been found to be the best approach for control source location optimisation problems. The starting population of individuals is selected at random using the available actuator locations and the specified number of control sources.

3.3 Crossover

Crossover is a process where a new individual (child) string is created by the random copying of information from two parent strings, with either parent string equally likely to be the information source for each element of the child string. A biased coin toss (with a success probability, P_c) is used in determining whether the crossover will occur or whether one parent string (with highest fitness value) is copied directly to the child population. In most control source location optimisation examples, P_c is set equal to one.

The random uniform crossover method (Syswerda 1989) simply involves randomly copying string information from either parent (with equal probability) to fill each string position in the child string as shown in Figure 9a and c. The exact form in

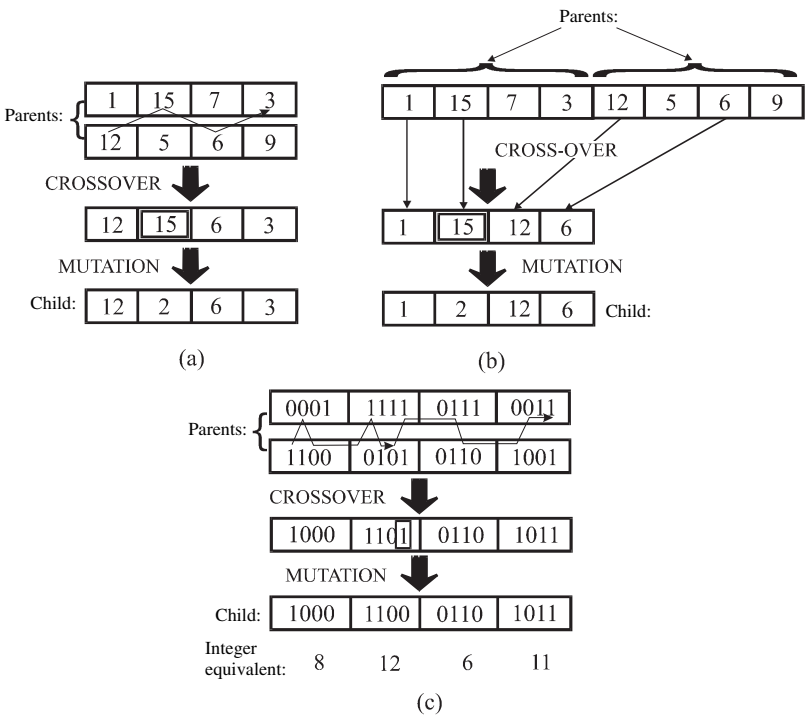


FIGURE 9 Illustration of crossover and mutation for (**a**) and (**b**) numerical string coding and (**c**) multi-variable binary string coding.

which crossover is performed remains fundamentally unchanged despite the fact that the string may be made up of either separate integer numbers or "bit" values from a binary string. However, the effect that crossover has on the variable values represented in the child string varies significantly between coding types.

In the case of numerical strings, each position in the string (consisting of either an integer value or a discretised real number) represents the whole value of an individual search variable. When crossover occurs, the newly created child string will contain only a random mix of values that exist in either parent string (Figure 9a). This contrasts with the crossover of MVBSs, where partial exchanges of the binary information representing each search variable can occur (Figure 9c). The result of partial exchanges of binary information is that the values represented by the child string may not necessarily be the same as those in the corresponding positions of either parent (in contrast to the case of the numerical string). Thus, an inherent diversity is observed when two variable values represented as binary numbers undergo crossover.

The net effect of this increased diversity during MVBS crossover remains unclear. Although increased diversity in the search population is desirable in sustaining a meaningful GA search for longer periods of time (Whitley 1989), it may also cause the search to lose focus and slow down the convergence rate. Therefore, as suggested above, it remains necessary to investigate the benefits of either coding method empirically.

In a traditional crossover operation, the new child is generated by randomly copying string information from one or other of the parents as shown in Figure 9a or c. However, it can be seen that the method illustrated in Figure 9a eliminates the possibility of the two values corresponding to any particular string location for the two parents ever appearing together in the child. Thus, it would seem that a better technique would be not to use crossover in the strict sense but instead allow any of the string values for both parents to appear in any of the string locations for the child; that is, use a "meta-string" representation as shown in Figure 9b. When an element is selected from the meta-string, it is removed to avoid duplication through reselection. If a duplicate is selected, which can happen when there are multiple entries in the meta-string, particularly near convergence, it is rejected and the selection is made again. This ensures uniqueness of all actuator locations in the final solution. In practice, this method has been found to give the best performance.

The technique of Figure 9b differs subtly from the traditional crossover technique used for breeding of integer-coded strings in

the number of unique combinations of actuator locations that can be produced. The simplest crossover scheme involves randomly copying string information from either parent to fill each string of the child and for a string length of n in each parent there are $2n$ possible unique children. The technique represented in Figure 9b produces $(2n)!/(n!)2$ possible unique children. As an example, for a 5-actuator system, the crossover method can result in 32 possible different children, whereas the meta-string method can result in 252 different children; thus the latter method is a more exhaustive search of the solution space.

3.4 Mutation

It is important to have a mechanism by which the population can be prevented from becoming too uniform; that is, population diversity is essential to ensure that local optima do not become the final solution. The maintenance of population diversity is achieved using a process known as mutation, where one of the string values is occasionally changed on a random basis in an attempt to prevent the irreversible loss of string information, which may be vital to the progress of the search. In addition to mutation, diversity can be encouraged by penalising solutions that are similar to other solutions in the population.

The mutation operation itself is an occasional (with small probability) random alteration of string values. For integer strings, this involves replacing selected string values with a randomly selected integer value (Figure 9a or b). For binary strings, this simply means flipping the value of a single bit position (Figure 9c). By itself, mutation represents a random walk through the parameter space. However, in the context of GAs, mutation plays an important secondary role in reintroducing lost information from overzealous exploitation by the parent selection and crossover operations. In practice, mutation minimises the risk of the search terminating at a local optimum, which is far removed from the global optimum.

For the optimisation of control source location when binary strings are used, mutation involves the random flipping of bits (e.g. 5 or 6 in one thousand bit transfers) during the crossover operation. When numerical strings were used, on average, 1 in 50 strings were affected and one number in that string would be replaced by a number selected randomly within the search space.

Another way of encouraging population diversity is to prevent the existence of duplicate solutions in the breeding population. To do this, forced mutation is used, which consists of a simple mutation procedure with a high probability, such that duplicate strings are significantly altered from their original state, by introducing randomly selected values at selected string locations.

By allowing some values of that string to undergo a random mutation and be replaced by a different (valid) string values, a new (distinct) search solution is created and included back in the search in the normal way. The forced mutation concept may be used in the control source location optimisation problem to take care of solutions in which more that one control source occupies the same location or if two control sources do not satisfy minimum clearance requirements. The forced mutation operator is also used to deal with circumstances where a newly created string already exists in the breeding population to be used next.

In addition to preventing duplicate strings and string values, forced mutation allows the reintroduction of new genetic material into the search. This occurs at a much faster rate than natural mutation, without significant detriment to the overall search progress. The result is potentially significantly higher diversity levels being sustained throughout the search procedure, meaning that for multi-modal search problems, the GA search is much less prone to focussing exclusively on a local optimum. Thus, forced mutation enables a more diverse range of strings to be reproduced, allowing the search to span a wider proportion of the solution space when searching for the global optimum solution.

3.5 Sharing

In natural genetics, the concepts of niche and species are fundamental to the simultaneous coexistence of many different living species. In each case, stable sub-populations of a particular species are formed when there exists increased competition for finite resources (which are vital for survival) limiting population growth and size. A similar concept to this has been developed in GAs (Goldberg and Richardson 1987). In this method, a practical sharing scheme was developed so that similar string solutions (in the breeding population) competed for breeding opportunities and continued survival. Sharing is a method that achieved this by applying a penalty (based on the value returned by a summary measurement function) to the parent selection probabilities of breeding population members that are similar in the string values they contain. In this way, the uncontrolled growth of a single string type amongst all population members is limited according to the number of similar solutions that coexist around it, thus preventing the search from stagnating due to homogeneity. In multi-modal problems, it was reported previously that as a result of sharing, stable sub-populations were formed around each optimum, where the number of strings around each occurred in relative proportion to the magnitude of the locally optimum fitness value.

Sharing is an effective mechanism for maintaining adequate levels of diversity amongst members of the breeding population. To do this, a sharing distance d_{ij} is defined, which determines the neighbourhood and degree of sharing that exists between strings i and j of the breeding population. A sharing function F_s, is defined such that $F_s = F_s(d_{ij})$ and is characterised by a linear inverse relationship with the sharing distance d_{ij}, illustrated in Figure 10.

A constant d_s is defined, which represents the upper bound for sharing distance values, below which sharing is considered significant. The sharing function values range from 0 (strings have no similarity) to 1 (strings are exactly the same). The niche count M_i is defined (for string i) as the sum of the sharing function values for each string comparisons such that

$$M_i = \sum_{j=1}^{n_{pop}} F_s(d_{ij}). \tag{4}$$

The fitness value f_i (of string i) is then derated in direct proportion to its niche count M_i such that

$$f_i' = \frac{f_i}{M_i}. \tag{5}$$

Altering the fitness values in this manner has the effect of increasing the focus of the search towards high-fitness solutions

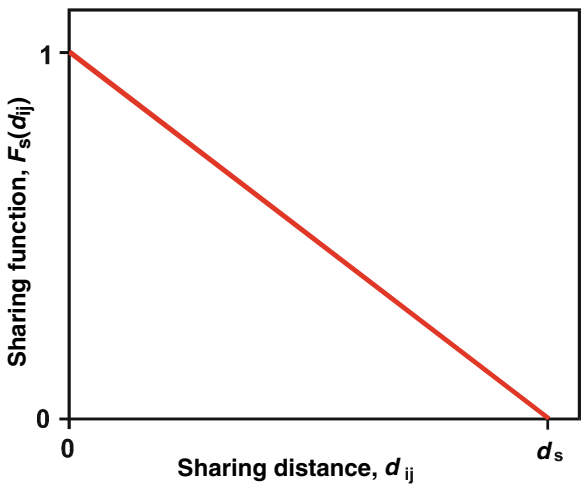


FIGURE 10 Sharing function versus distance between adjacent locations.

which are under-represented within the breeding population, thus maintaining sustainable levels of diversity.

Examples of the results obtained with the various approaches described above are provided for a longitudinally stiffened cylinder by Hansen et al. (1999). It is interesting to note that as the number of actuators increases the importance of where they are located for best performance decreases.

References

Chen, T-L., Oldham, K., Li, Y. and Horowitz, R. (2002). Suspension vibration compensation using a MEMS microactuator in hard disk drives. Proceedings of 2nd IFAC Conference on Mechatronic Systems, Berkeley, California, pp. 757–762, December 9–11.

Clark, W.W., Robertshaw, H.H. and Warrington, T.J. (1990). A comparison of actuators for vibration control of the planar vibrations of a flexible cantilevered beam. Journal of Intelligent Material Systems and Structures, 1, 289–308.

Anonymous (1990). RA101 Actuator Handbook. Edge Technologies, Amers.

Hansen, C.H., Simpson, M.T. and Cazzolato, B.S. (1999). Genetic algorithms for active sound and vibration control. Proceedings of IEE Inter-Active'99.

Hiller, M.W., Bryant, M.D. and Umega, J. (1989). Attenuation and transformation of vibration through active control of magnetostrictive terfenol. Journal of Sound and Vibration, 134, 507–519.

Katsikas, S.K., Tsahalis, D., Manolas, D., and Xanthakis, S., (1993). Genetic algorithms for active noise control. Proceedings of NOISE-93, St. Petersburg, Russia, pp. 167–171.

Goldberg, D.E., (1989). Genetic algorithms in search, optimisation and machine learning. Addison-Wesley Publishing Company, Inc., New York.

Goldberg, D.E. and Richardson, J. (1987). Genetic algorithms with sharing for multi-modal function optimisation. Proceedings of the Second International Conference on Genetic Algorithms, pp. 41–49.

Kim, Y.S., Kim, G.M., Ha-Roh, C. and Fuller, C.R. (2005). A hybrid control development to suppress the noise in rectangular enclosure using an active/passive smart foam actuator. International Journal of Precision Engineering and Manufacturing, 6, 37–43.

King, S.P. (1988). The minimisation of helicopter vibration through blade design and active control. Aeronautical Journal, 92, 247–263.

Moffett, M.B., Clark, A.E., Wun-Fogle, M., Linberg, J., Teter, J.P. and McLaughlin, E.A. (1991a). Characterisation of terfenol-D for magnetostrictive transducers. Journal of the Acoustical Society of America, 89, 1448–1455.

Moffett, M.B., Powers, J.M. and Clark, A.E. (1991b). Comparison of terfenol-D and PZT-4 power limitations. Journal of the Acoustical Society of America, 90, 1184–1185.

Rogers, C.A. (1990). Active vibration and structural acoustic control of shape memory alloy hybrid composites: experimental results. Journal of the Acoustical Society of America, 88, 2803–2811.

Smits, J.G., Dalke, S.I. and Cooney, T.K. (1991). The constituent equations of piezoelectric bimorphs. Sensors and Actuators A, 28, 41–61.

Syswerda, G. (1989). Uniform crossover in genetic algorithms. Proceedings of the Third International Conference on Genetic Algorithms, George Mason University, Washington, DC, pp. 2–9.

Tsahalis, D.T., Katsikas, S.K. and Manolas, D.A. (1993). A genetic algorithm for optimal positioning of actuators in active noise control results from the ASANCA project. Proceedings of Inter-Noise 93, Vol. 1, pp. 83–88.

Whitley, D., (1989). The genitor algorithm and selection pressure: why rank-based allocation of reproductive trials is best. Proceedings of the Third International Conference on Genetic Algorithms, George Mason University, Washington, DC, pp. 116–121.

Whitley, D. and Hanson, T. (1989). Optimising neural networks using faster, more accurate genetic search. Proceedings of the Third International Conference on Genetic Algorithms, George Mason University, Washington, DC, pp. 391–396.

Yao, L., Sethares, W.A. and Kammer, D.C., (1993). Sensor placement for on-orbit modal identification via a genetic algorithm. AIAA Journal, 31, 1922–1927.

Zimmerman, D.C. (1993). A Darwinian approach to the actuator number and placement problem with non-negligible actuator mass. Journal of Mechanical Systems and Signal Processing, 7, 363–374.

61

Materials Testing

Walter Lauriks[1,3] *and Philippe Leclaire*[2,3]

[1]*Katholieke Universiteit Leuven, Heverlee, Belgium,*
e-mail: Walter.Lauriks@fys.kuleuven.ac.be
[2]*Université de Bourgogne, Nevers, France.*
[3]*Laboratorium voor Akoestiek en Thermische Fysica, K.U. Leuven, Leuven, Belgium*

1 • Introduction

In recent years, a number of models to calculate the acoustical behaviour of porous materials (sound absorption, sound insulation and vibration damping) have been developed [1]. Although these models are based on physical sound theories, they require a number of material parameters and the output of a calculation will depend on the accuracy of the input parameters. Depending on the complexity of the porous material and the configuration to be modelled, up to seven parameters may be needed.

A porous material, as used in noise control applications, is made of a solid elastic matrix, saturated by air. Between these two phases, several interactions exist, making the acoustical behaviour of a porous material totally different from that of nonporous elastic solids. Although the materials are often highly porous (porosities higher than 95 % can occur), viscous friction of the air in the pores and heat exchange between the air and the frame have to be accounted for.

2 • Sound Propagation in Rigid Porous Materials

In many cases for air-saturated materials, the porous frame can be considered to be much more rigid and heavier than air and a simplified model can be used in the rigid-frame approximation. The interest in theory of sound propagation in rigid porous materials dates from the time of Lord Rayleigh, with substantial contributions of Zwikker and Kosten [2] in the 1940s. The importance of the flow resistivity or permeability was quickly recognised and this resulted in some simple empirical formulas that predict the acoustic behaviour on the basis of this one single parameter [3,4,5]. New life was breathed into this research with the work of Attenborough [6,7], who stressed the importance of the tortuosity and pointed out that the complexity of the pore shape should be accounted for by some pore shape factor. A substantial refinement in the physical study of wave propagation in porous media was performed by Johnson et al. [8], who introduced the concept of dynamic tortuosity to account for the frequency-dependent behaviour of the viscous effects in the pores. The concept of the viscous characteristic length Λ was introduced to integrate the complexity of the pore shape. For sound-absorbing materials saturated by air, the deformation of the air is isothermal at low frequencies due to the very good thermal exchange between the air in the pores and the frame that has an almost infinite thermal heat capacity and hence remains at constant temperature. At high frequencies, the compression and dilatation of the air become adiabatic, resulting in a frequency-dependent compressibility to cover the entire frequency range to describe this effect. The most significant work in this area has been done by Champoux and Allard [9] and Lafarge [10], who introduced the concept of thermal characteristic length Λ'.

2.1 Porosity

One of the basic material parameters for sound-absorbing and sound-insulating materials is the porosity ϕ. It is defined as the volumic ratio of the amount of air in the sample that is accessible from outside to the total volume of the sample. In acoustic terminology, closed cells are not considered as part of the porosity but constitute part of the elastic frame. The main influence of the presence of closed cells is that they change the effective elasticity and density of the frame. If the material contains no closed cells and if the density of the frame material is known, the porosity can easily be determined from the apparent density of the sample. This is the case, for instance, for mineral wool materials or materials made from sintered glass beads. If the material may contain closed cells or if the density of the frame material is not exactly known, as is often the case for plastic foams, evaluation



FIGURE 1 Schematical representation of the porosimeter. The sample is contained in a closed cell and volume is changed using a moveable piston. In practice, the volume changes are extremely small.

of the porosity is slightly more complicated. The basic technique has been proposed by Zwikker and Kosten, [2] but it has been developed into a precise and modern method by Champoux et al. [11] and Henry [12]. The method is shown schematically in Figure 1 and is based on the ideal gas law at constant temperature. The pressure change in a closed volume (with and without sample) is measured as a function of small volume changes, the temperature being kept constant. This method can easily be automated using an electronic pressure transducer.

It is our experience that for low-density plastic foams, the porosity is almost always very close to 1. For granular materials, the porosity is often below 0.3. For compressed thermoplastic foams, the porosity can deviate substantially from 1.

Recently, a new method based on the analysis of a reflected ultrasonic pulse at the surface of the sample has been proposed by Fellah et al. [13]. This method is particularly interesting in combination with the ultrasonic method for evaluating the tortuosity described in Section 2.3.

2.2 Flow Resistivity

The flow resistivity σ (or alternatively the permeability) has been recognised as a crucial parameters for the attenuation of sound in a porous material from the time of Zwikker and Kosten. For simple sound-absorbing materials, like open-cell foams or mineral wool, some empirical formulas have been designed that predict the sound velocity in the material and the characteristic impedance of the material as a function of frequency using only the flow resistivity as a material parameters [3]. These equations have later been refined by Mechel [4] and Miki, [5] and within their range of applicability, these equations give reasonably accurate results. The flow resistivity can be determined with a simple method described in ISO 9053 [14] and is shown in Figure 2. A static airflow Q is passed through the

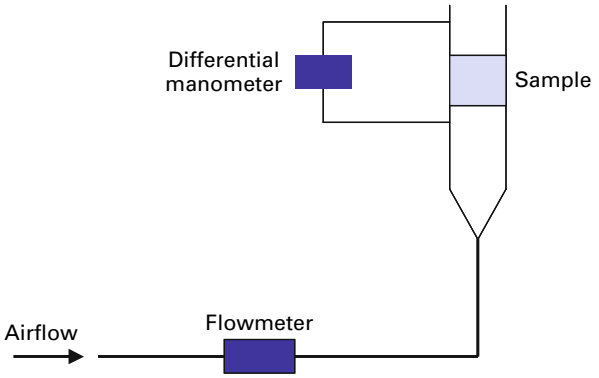


FIGURE 2
Schematic representation of the flow resistance meter.

sample and the pressure drop Δp that is caused by the viscous friction is measured across the sample. The flow resistivity is the ratio of these two quantities, normalised per unit thickness and unit area of the sample:

$$\sigma = \frac{\Delta p}{Q}\frac{S}{d},$$

with S the area and d the thickness of the sample. An alternative method, using an AC fluid flow generated by an oscillating (at low frequency) piston is also used and the results agree reasonably well with the static method [15]. The flow resistivity describes the viscous friction in the sample at DC or low frequency. At very high frequency, the influence of viscous friction becomes negligible. The transition between the "low" and "high" frequency regime is determined by the viscous characteristic length (see Section 2.4). The experimental setup of the flow resistivity meter is extremely simple, the main sources of errors being leaks between the sample and the tube wall and possible nonlinear effects due to a volume flow that is too high.

2.3 Tortuosity

The significance of the tortuosity (also called sinuosity or "shape factor") was already recognised in early publications. A precise definition was given by Biot (see, for instance, Ref. [1]). Basically, the tortuosity is needed because the actual particle velocity at a microscopic level in the pores deviates from the average particle velocity at a macroscopic scale. This leads to a difference in kinetic energy, which is corrected for by multiplying the density with a factor that is equal to the ratio of these kinetic energies:

$$\alpha_\infty = \frac{\int_V v^2 \, dV}{\left(\int_V \vec{v} \, dV \right)^2},$$

with V a volume much larger than the pore size, but much smaller than the wavelength.

It has been shown by Brown [16] that the tortuosity can be evaluated from an electrical conduction experiment. The sample is saturated with an electrical conducting liquid and the electrical resistivity of the saturated sample Ω_{sample} is compared with the resistivity of the liquid Ω_{liquid}. The tortuosity is then given by

$$\alpha_\infty = \phi \frac{\Omega_{\text{sample}}}{\Omega_{\text{liquid}}}.$$

The method can obviously only be applied when the frame of the porous material is a nonconducting solid. The method has been optimised by Champoux [17].

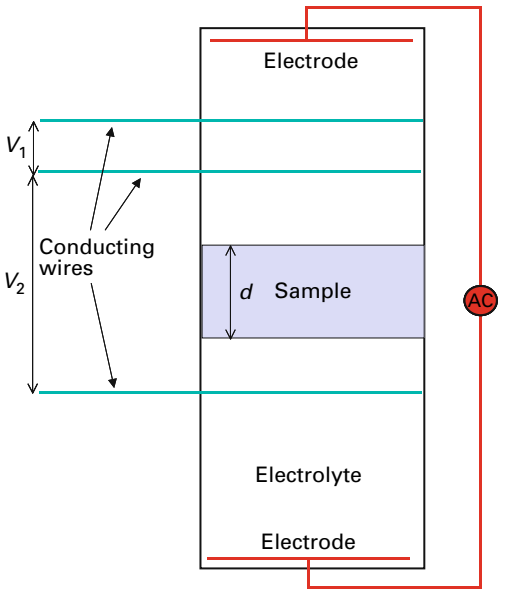


FIGURE 3 Schematic representation of the tortuosimeter proposed by Champoux. The sample is mounted in a plastic cylinder tube that is terminated by two electrodes and filled with an electrolyte. An AC source generates a plane electric field in the tube. The voltage difference V_1 leads to the resistivity of the liquid, and the resistivity of the sample can be deduced from V_2 [17].

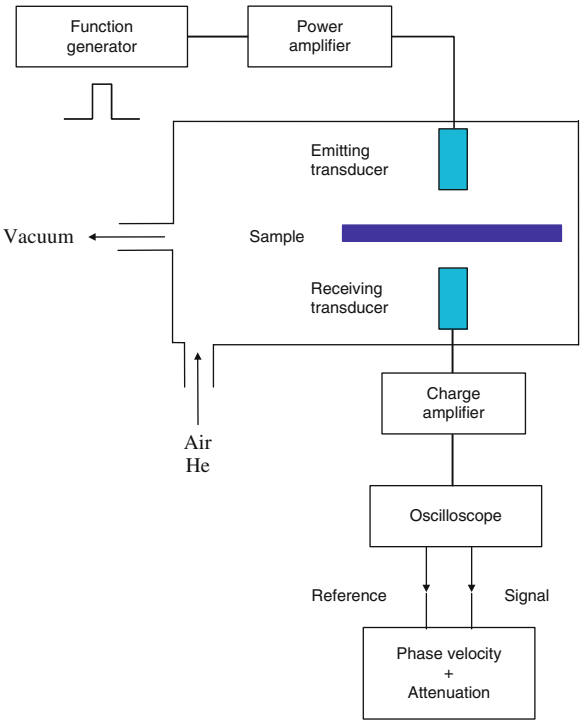


FIGURE 4
Schematic representation of the ultrasonic tortuosimeter [20].

In 1996 Allard et al. [18] and Leclaire et al. [19] have proposed an alternative, quicker method. The method is based on the fact that the phase velocity $c(\omega)$ at very high frequencies in the porous sample depends solely on the tortuosity:

$$\lim_{\omega \to \infty} \frac{c(\omega)}{c_0} = \frac{1}{\sqrt{\alpha_\infty}}$$

with c_0 the sound velocity in free air. An ultrasonic pulse (usually with a frequency content between 100 kHz and 1 MHz) is emitted using an air-coupled ultrasonic transducer (ULTRAN NTC or MICROACOUSTIC capacitive transducer) and a high-voltage pulser. The signal is captured by a second (identical) transducer a few centimetres away from the first and visualised on an oscilloscope. When the sample is placed between the two transducers, the pulse will be delayed because the phase velocity in the sample will always be lower than the velocity in free air (due to the tortuosity effect). The phase velocity in the sample and hence the tortuosity can directly be calculated from this delay and the thickness of the sample. In practice, the phase velocity in the sample is

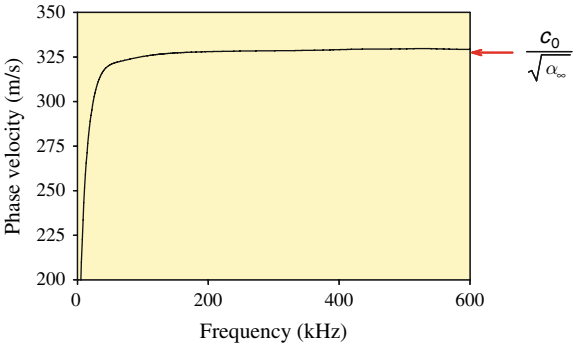


FIGURE 5 Phase velocity in the foam as a function of frequency.

measured as a function of frequency to verify that the high frequency limit has been obtained. This is done by calculating the phase of the Fourier transform of the received pulse after passing through the sample relative to the phase of the Fourier transform of the reference pulse. Figure 5 shows the phase velocity in a porous foam as a function of frequency. The high frequency limit is characterised by a frequency-independent phase velocity and the value of the phase velocity in this regime gives directly the tortuosity.

Both methods agree within the experimental uncertainties.

2.4 Viscous and Thermal Characteristic length

2.4.1 Viscous Characteristic Length

The viscous interaction between the oscillating air and the pore walls is known exactly at low and high frequencies. In the intermediate-frequency regime, this interaction is approximated by a simple analytical function that links the low-and the high-frequency regime [8]. The average pore radius determines the transition between "low" and "high" frequencies. For a material with straight cylindrical pores, this would simply be the radius of the pore, but for a material with pores with a variable cross section, narrow sections will have to contribute more to this average pore size than wide section. Johnson [8] gave the exact definition of the viscous characteristic pore size

$$\Lambda = 2\frac{\int_V v^2 \mathrm{d}V}{\int_S v^2 \mathrm{d}S},$$

where the integral in the numerator is taken over a volume larger than the pore size and the integral in the denominator is taken over the surface of the pores and v is the microscopic velocity

of a nonviscous liquid in the pores. The viscous characteristic length is hence the ratio of the volume to the surface area in the pores, weighted by the square of this velocity (which makes small section contribute more than large sections).

2.4.2 Thermal Characteristic Length

The thermal interaction between compressed and dilated air and the pore walls is known exactly at low (isothermal) and at high (adiabatic) frequencies. In a similar way as for the case of the viscous interaction, this thermal effect is approximated in the intermediate-frequency regime by an analytical function that links the low-to the high-frequency regime [9,10]. The transition between "low" and "high" frequency regime depends on the average distance the heat has to travel to the pore walls and is characterised by the distance:

$$\Lambda' = 2\frac{\int_V \mathrm{d}V}{\int_S \mathrm{d}S}.$$

The thermal characteristic length is hence the ratio of the pore volume to pore surface. It is obvious that since no weighting is applied in this case, the thermal characteristic length will be larger or equal to the viscous characteristic length. In case of a porous material with straight cylindrical pores, these two characteristic lengths are equal and equal to the radius of the pores. In case of fibrous materials, it can be shown [1] that the thermal characteristic length is twice the viscous characteristic length.

2.4.3 Evaluating Viscous and Thermal Characteristic Lengths

The method for evaluating the characteristic lengths has been established by Leclaire et al. [19,20]. These are extracted from the attenuation of an ultrasonic pulse propagating in the porous material. It has been shown that in the high frequency limit, the attenuation per cycle Q is given by

$$\lim_{\omega\to\infty} Q\delta = \left(\frac{1}{\Lambda} + \frac{\gamma-1}{B\Lambda'}\right)^{-1},$$

where δ is the viscous skin depth equal to

$$\delta = \sqrt{\frac{2\eta}{\omega\rho_\mathrm{f}}},$$

with η the viscosity, ρ_f the density of the fluid, γ the ratio of the specific heats and B the square root of the Prandtl number of the gas saturating the pores.

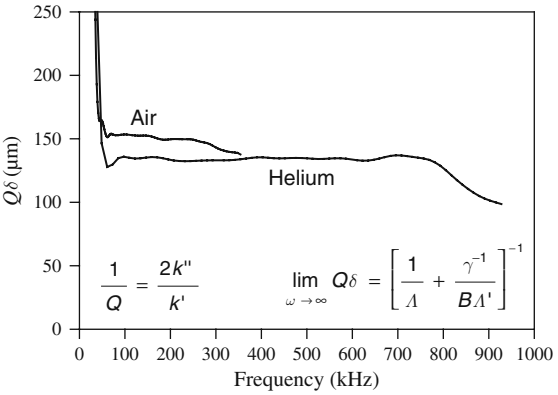


FIGURE 6 Damping in the material at high frequencies can be used to measure the characteristic lengths.

Evaluating the characteristic lengths is particularly difficult since distinguishing between damping due to viscous effects and that due to thermal effects requires two independent measurements in different gases with different thermal and viscous properties. In practice, helium and air or helium and Forane 132a are chosen. Figure 6 shows the $Q\delta$ factor of a foam, measured in helium and air. The horizontal plateau between 100 and 300 kHz for air and between 100 and 800 kHz for helium in the curve characterises the high frequency limit. The characteristic lengths can be obtained by solving these equations for the two unknown quantities Λ and Λ'. The drop-off at 300 kHz in air and 800 kHz in helium is due to scattering of the sound wave in the porous material.

The thermal characteristic length can also be linked with the internal surface area of the porous material and can be determined from a BET adsorption measurement, well known in powder technology [21]. These measurements agree well with the ultrasonic results.

3 • Sound Propagation in Elastic Porous Materials

As mentioned in the previous chapter, approximating the material as a rigid porous material gives good results in many applications for many sound-absorbing materials. However, this approximation will not always hold, and in a number of cases, the acoustic material will have to be described with the full theory for sound propagation in poroelastic media [1]. For instance, if the porous layer is fixed to a vibrating plate, the (complex) elastic coefficients of the material will be of crucial importance [22]. If the flow resistivity and the tortuosity of the material are high (as is often the case for good sound-absorbing materials), the movement of the frame cannot be neglected and the elasticity will

have to be included in the model. Evaluating the elastic coefficients of poroelastic materials is less trivial than it appears. First of all, the elastic coefficients of the "dry" (unsaturated) frame are needed in the models. This leads to two options: either designing an experiment in vacuum where the numerical inversion to extract the elastic coefficients will be easy or designing an experiment in air that is more easy to perform but where the numerical inversion may require the full model. Pioneering work on this subject has been done by Pritz [23]. Moreover, many sound-absorbing or sound-damping materials are viscoelastic and their elastic coefficients may depend on frequency and temperature.

The simplest experiments to evaluate the elastic coefficients are based on the measurement of a vibrational resonance frequency in a sample (rod-shaped, bar-shaped or a plate) of small dimensions compared to the wavelength involved. If the material is isotropic, two complex elastic coefficients are needed. In case the material is anisotropic, more parameters are needed [24,25].

3.1 Evaluating the Complex Shear Modulus

The shear modulus of a poroelastic sample can be measured with the device shown in Figure 7.

Two identical pieces of the sample (typical surface area S is 10 by 10 cm^2 and thickness ℓ about 1 or 2 cm) are glued to a rigid backing (1) and to a thin, moveable plate (2). This plate is connected via an impedance head (4) to a shaker (5). The shaker is excited with a broadband white noise and generates shear waves in the sample. The transfer function between the force F and the acceleration a at the output of the impedance head shows the resonance frequencies in the sample. Neglecting the effect of the air in the pores, the transfer function in the frequency domain is given by

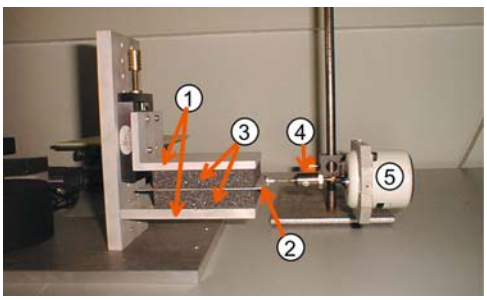


FIGURE 7 Overview of a setup to measure the shear modulus of the frame. (1) rigid plate; (2) moveable plate; (3) sample, (4) B&K 8001 impedance head; (5) shaker.

$$\frac{F}{a} = -\frac{2S}{\omega}\sqrt{N\rho}\,\tan\left(\sqrt{\frac{\rho}{N}}\omega\ell\right) - m,$$

where m is the mass of the moveable plate, the connection rod and the moveable part in the impedance head.

From these resonance frequencies, the density ρ and the thickness ℓ of the sample, the magnitude of the shear N modulus can be determined. From the width at half height of the resonance peaks, the loss angle can de determined. By performing the experiment with samples with different thickness, information about the frequency dependence of the shear modulus can be obtained. Figure 8 shows a typical transfer function. Up to three resonance frequencies can be observed, decreasing in amplitude due to the attenuation in the sample. In any case, this figure shows that the method is limited to the lower part of the audible frequency range.

3.2 Evaluating the Complex Young Modulus

The Young modulus of the sample can be measured with a similar experiment, shown in Figure 9. A rod-shaped specimen of the sample, with length ℓ and a cross section much smaller than the wavelengths involved, is attached to a shaker via an impedance head. Feeding white noise to the shaker will generate longitudinal waves in the sample. The transfer function between the force F and the acceleration a at the output of the impedance head is given by

$$\frac{F}{a} = -\frac{S}{\omega}\sqrt{E\rho}\,\tan\left(\frac{\rho}{E}\omega\ell\right),$$

with S the surface area of the cross section of the bar.

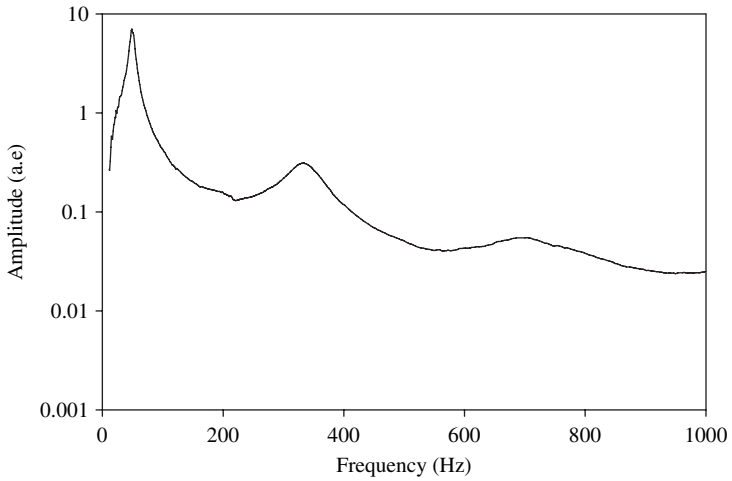


FIGURE 8 Typical shape of the transfer function between force and acceleration of the experiment of Figure 7.

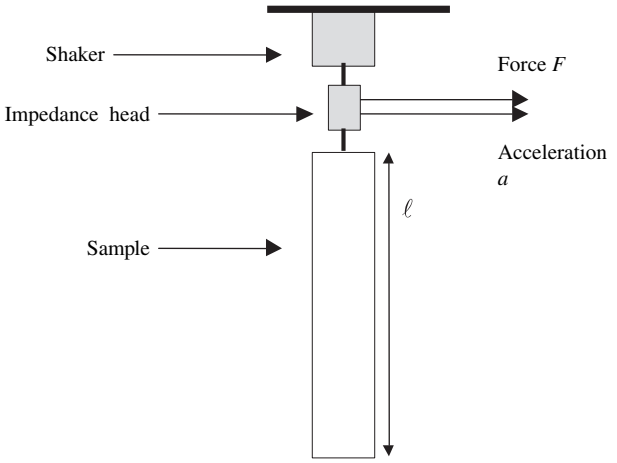


FIGURE 9 Schematic representation of a setup to measure the Young modulus.

From the resonance frequencies in the transfer function, the magnitude of the Young modulus can easily be extracted. The width of the resonance peaks is determined by the damping in the sample.

Typical values of the elastic coefficients of poroelastic foams used for noise control applications range from 10^4 to 10^7 Pa. Several variations of these setups can be designed depending on the shape and size of the available sample. For instance, rod-like samples can be excited in bending mode or plate-like samples can de-excited in a longitudinal thickness resonance.

3.3 Frequency Dependence of the Elastic Coefficients

The methods for evaluating the shear and Young's moduli presented in Sections 3.1 and 3.2 allow a study of these parameters as a function of frequency. However, these methods are based on the vibration of a sample of finite size resulting in a transfer function with resonance peaks. The effect of resonances in the transfer function is that the moduli are determined with a precision that varies and depends on the frequency area examined. The precision is best around the resonance frequencies and poorer elsewhere. In addition, in the experiment for measuring Young's modulus, the wavelengths involved are greater than or in the order of the length of the rod, they must be greater than the lateral size and method is limited to low frequencies (typically 400 Hz). Since most the porous materials are also designed to attenuate structure-borne sound, their structural damping is high, and as a consequence, the elastic coefficients are frequency dependent. Evaluating the elastic

coefficients at low frequencies can have a poor predictive value for higher frequencies.

A new higher frequency method for measuring the shear modulus was proposed by Allard et al. [26] based on the propagation of a structure-borne Rayleigh wave at the interface between a porous solid and a fluid. It was applied to a highly porous foam saturated by air and provided values of the shear modulus at very high frequencies above 3 kHz. Following this work, an experimental and theoretical analysis of the dispersion of guided acoustic wave in a porous layer has been proposed by Boeckx et al. [27] Together with a complete description of a soft porous layer lying on a rigid substrate, they proposed a novel experimental technique for measuring the phase velocities of all the possible modes. The method is based on the generation and detection of standing waves in the layer and on the spatial Fourier transform of the standing-wave pattern. It was inspired from the well-known impedance tube measurement method. The principle is depicted in Figure 10.

The porous layer is excited at one end with a thin aluminium strip used as a line source attached to a shaker at one end and to the sample at the other end. The shaker is fed with a sinusoidally swept signal to maximise the precision of the measurements in a wide frequency range. The other end of the layer is stuck on a rigid end so that incident waves can be reflected. The displacement of the surface is measured with a laser Doppler vibrometer. The laser beam at the output of the laser is collimated and a mirror/lens arrangement ensures that the beam is always focused on the surface any position of the beam. The measurement point can be moved by moving the mirror/lens arrangement. A strip of reflective tape is used to reflect the laser beam in the path of the scanning beam. For each frequency, the path is scanned with a typical step of 1–5 mm and the amplitude and phase of the signal are recorded at each position.

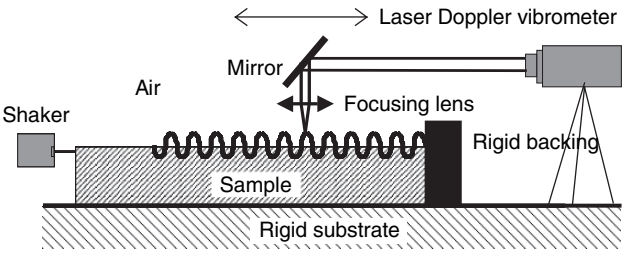


FIGURE 10 Experimental principle for measuring the phase velocities of guided acoustic waves in soft porous layers.

Once the data are recorded, the spatial Fourier transform of the displacement profile in the vicinity of the rigid end can be calculated. The different amplitude peaks in the spatial spectrum provide the wavenumbers of all the guided modes propagating in the layer and the phase velocities are obtained from dividing the frequency by these wavenumbers. The fitting of at least two theoretical dispersion curves to the experimental data allows to investigate both the shear and Young's moduli in the frequency range between the classical vibration method and the Rayleigh wave propagation method. Recent results show a clear frequency dependence of the shear modulus of several foams.

References

1. J.F. Allard, 1993 *Sound propagation in porous media. Modeling sound absorbing materials.* Elsevier Applied Science, London.
2. C. Zwikker and C.W. Kosten, 1949 *Sound absorbing materials*. Elsevier, New York.
3. M.E. Delany and E.N.Bazley "Acoustical characteristics of fibrous absorbent materials" Applied Acoustics **3** 1970 pp. 105–116.
4. F.P. Mechel "Ausweitung der Absorberformel von Delany and Bazley zu tiefen Frequenzen" Acustica **35** (1976) pp. 210–213.
5. Y. Miki "Acoustical models of porous materials – modification of Delany-Bazley model" Journal of the Acoustical Society of Japan **11** (1990) pp. 19–24
6. K. Attenborough "Acoustical characteristics of porous materials" Physics Reports **82** (1982) pp. 179–227
7. K. Attenborough "Acoustical characteristics of rigid fibrous absorbents and granular materials" Journal of the Acoustical Society of America **73** (1983) pp. 785–799
8. D.L. Johnson, J. Koplik, and R.Dashen "Theory of dynamic permeability and tortuosity in fluid saturated porous media" Journal of Fluid Mechanics **176** (1987) pp. 379–402
9. J.F. Allard and Y. Champoux "New empirical equations for sound propagation in rigid frame fibrous materials" Journal of the Acoustical Society of America **91** (1992) pp. 3346–3353
10. D. Lafarge, P. Lemarinier, J. F. Allard, and V.Tarnow "Dynamic compressibility of air in porous structures at audible frequencies" Journal of the Acoustical Society of America **102** (1997) pp. 1995–2007
11. M.R. Stinson and G.A. Daigle "Electronic system for the measurement of flow resistance" Journal of the Acoustical Society of America **83** (1988) pp. 2422–2428.
12. M. Henry Mesures des paramètres caractérisant un milieux poreux. Etude expérimentale du comportement acoustique des mousses aux basses fréquences, PhD thesis Université du Maine, France (1997)
13. Z.E. Fellah, S. Berger, W. Lauriks, C. Depollier, C. Aristegui, and J. Chapelon "Measuring the porosity and the tortuosity of porous

materials via reflected waves at oblique incidence" Journal of the Acoustical Society of America **113** (2004) pp. 2424–2433
14. ISO 9053:1991 Acoustics – Materials for acoustical applications – Determination of airflow resistance
15. M. Garai, and F. Pompoli "A European Inter-Laboratory Test of Airflow Resistivity Measurements" Acta Acustica untited with Acustica **89** (2003) pp. 471–478.
16. R. Brown "Connection between the formation factor for electrical resistivity and fluid-solid coupling factors in Biot's equations for acoustic waves in fluid-filled porous media" Geophysics **45** (1980) pp. 1269–1275.
17. Y. Champoux, *Etude expérimentale du comportement acoustique des matériaux poreux à structure rigide*, PhD thesis, Carlton University, Canada, 1991.
18. J.F. Allard, B. Castagnède, M. Henry, and W. Lauriks "Evaluation of the tortuosity in acoustic porous materials saturated by air" Review of Scientific Instruments **65** (1994) pp. 7654–755
19. P. Leclaire, L. Kelders, W. Lauriks, C. Glorieux, and J. Thoen "Determination of the viscous characteristic length in air-filled porous materials by ultrasonic attenuation measurements" Journal of the Acoustical Society of America **99** (1996) pp. 1944–1948
20. P. Leclaire, L. Kelders, W. Lauriks, M. Melon, N. Brown, and B. Castagnède "Determination of the viscous and the thermal characteristic length of plastic foams by ultrasonic measurements in helium and air" Journal of Applied Physics **80** (1996) pp. 2009–2012
21. P. Lemarinier, M. Henry, J.-F. Allard, J.L. Bonardet, and A. Gedeon "Connection between the dynamic bulk modulus of air in a porous medium and the specific surface" The Journal of the Acoustical Society of America **97** (1995) pp. 3478–3482
22. V. Litwinczik, W. Lauriks, G. Geentjens, G. Vermeir, and A. Lennzi "Experimental acoustic radiation from an elastic plate covered with a porous layer" Proceedings of ISMA2002 conference, Leuven, Belgium pp. 2159–2164
23. T. Pritz "Transfer function method for investigating the complex modulus of acoustic materials" Journal of Sound and Vibration **77** (1981) pp. 93–100
24. S. Sahraoui, E. Mariez, and M. Etchessahar "Linear elastic properties of anisotropic open-cell foams" The Journal of the Acoustical Society of America **110** (2001) pp. 635–637
25. L. Jaouen, "Contribution à la caractérisation mécanique de matériaux poro-viscoélastiques en vibro-acoustique," Ph. D. thesis, Université de Sherbrooke, Canada, 2003.
26. J.F. Allard, G. Jansens, G. Vermeir, and W. Lauriks "Frame-borne surface waves in air-saturated porous media," Journal of the Acoustical Society of America **111** (2002) pp. 690–696.
27. L. Boeckx, P. Leclaire, P. Khurana, C. Glorieux, W. Lauriks, and J. F. Allard "Investigation of the phase velocities of guided acoustic waves in soft porous layers" Submitted to Journal of the Acoustical Society of America **117**(2), (2005) pp. 545–554.

PART XI
BIORESPONSE TO VIBRATION

Ron Verrillo

Syracuse University, Syracuse, USA

62 Tactile Responses to Vibration

Ronald T. Verrillo and Stanley J. Bolanowski 1185

Introduction • Neural Structures • Thresholds of Detection in the Temporal and Spatial Domains • Suprathreshold Responses • Subject Variables • The Four-Channel Model of Cutaneous Mechanoreception • Roughness and Texture • Biophysics

63 Uses and Applications of Somatosensation

Ronald T. Verrillo 1215

Introduction • Intermodality Interactions • Sensory Substitution • Health • Environmental Well-Being • Whole Body Vibration

64 Sensory Substitution

Ronald T. Verrillo 1231

Introduction • Temporal Gap Detection • User Variables • Practical Design Considerations • Mechanical Stimulators • Electrical Stimulators • Coding • The Systems • Training

62

Tactile Responses to Vibration

Ronald T. Verrillo and Stanley J. Bolanowski

Institute for Sensory Research, Syracuse University, NY 13244, USA

1 • Introduction

The skin is by far the largest organ of the human body having an area of approximately 1.85 m^2 in the adult male and 1.65 m^2 in the adult female with an average weight of about 9 lb., comprising about 6 % of the total body weight. In thickness it ranges from an average of 4.4 mm in non-hairy skin to 2.1 mm in hairy skin. It is thin over the eyelids and thick at the soles of the feet. The human body contains four types of skin. The orifices and canals of the body that are connected to the outside are lined with *mucous* membrane. These are the alimentary tract, the urogenital tract, and the respiratory system. This type of tissue secretes digestive and lubricating fluids and interchange nutritive and excretory substances. The intermediate regions of skin between the outer membrane and the mucous linings are composed of *mucocutaneous* tissues which cover the lips, the genital regions, the tongue, the inner margins of the eyelids, and the perianal area.

The portions of the body that do not contain hair, the soles of the feet and palms of the hand, are called *glabrous* skin. This is the most complex type of skin, having many layers and intricate anatomical structures. Most of the body, however, is covered in *hairy* skin. In some places, such as the scalp, the hair is long and tough whereas in others, such as the forehead and volar forearm (and the face of women), the hair is typically short and downy. It

is called *vellus* hair. The rest of the body has hair of intermediate size. We limit our discussion here to hairy and glabrous skin.

Hairy and glabrous skin consists of two layers, the outer surface, the *epidermis* which is visible, and an inner, thicker layer, the *dermis*. In this chapter we shall focus principally on the dermis since it is in this layer that the organs of sensation are located. Each of the peripheral layers of skin is comprised of multiple layers, each with its own function (Figure 1). The epidermis is the more complex layer. It is cellular in nature constantly generating new cells at its base (*stratum germinatium*) which migrate in an orderly fashion toward the surface where they gradually change in shape and ultimately flatten into a pliable but tough protective covering. This process is called terminal differentiation. The top layer gradually is scraped off by friction or by simply sloughing off. We see this as dandruff or flakes of skin on the surface. The entire cycle from generation at the epidermal base to casting off at the surface takes about 27 days in humans.

The skin has a variety of functions. Perhaps its most important function is that of protection; it is the first line of defense between the body and the external environment. It keeps things out with a layer of intricate chemical compounds spread across its surface (the acid mantle) which form a barrier to intrusions by bacteria, fungi, and parasites. The topmost layer contains keratin which is a stable protein that is not affected by most chemicals and therefore prevents intrusion by many of the common chemical agents found in the environment. Just below the surface is a layer of clear cells that is impervious to noxious agents that are not stopped by the keratin layer. The skin is also host to specialized pigment cells (*melanocytes*) that determine the color of the visible surface and form a deep barrier to penetration by destructive ultraviolet rays. Melanin is a product of the pituitary gland. Whatever the color of the skin, whether Negroid, Oriental Amerind, or Caucasian, the number of melanocytes is the same; the skin color is determined not by the number of cells, but rather by the cell's chemical composition. The entire structure of the epidermis is tough enough to prevent mechanical injury to underlying tissues by all but the sharpest objects. The pain resulting from the slightest touch to a region where the skin is scraped off will attest to the protection offered by this tough outer layer (*stratum corneum*). Of equal importance is its role in keeping things within the body. The skin forms a barrier to prevent the escape of vital body fluids from the inside-out. It also has a carefully regulated mechanism, sweating, for the specific purpose of maintaining the core body temperature

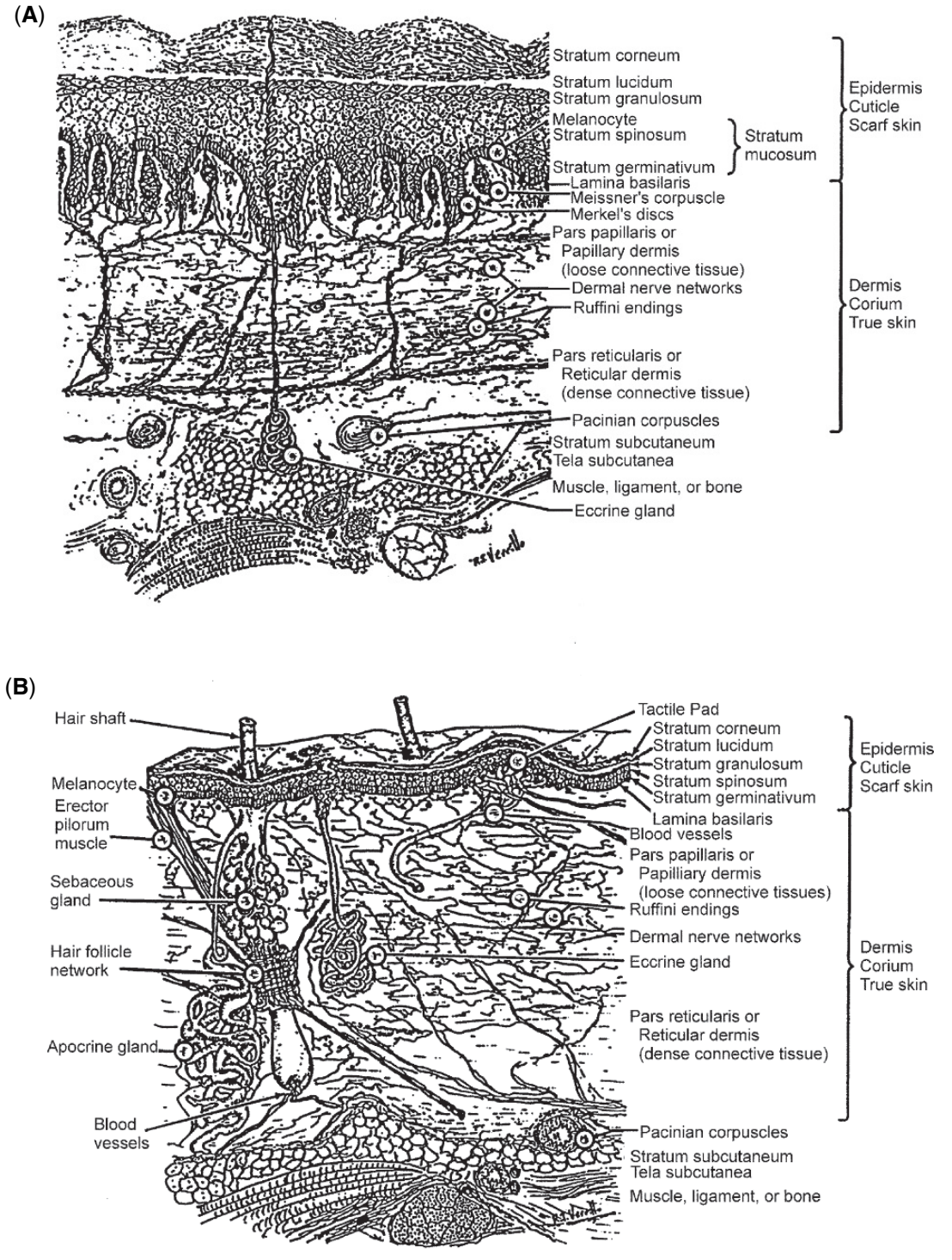


FIGURE 1 Schematic representation of glabrous (**A**) and hairy (**B**) skin showing the multiple layers of cells that undergo a constant cycle of regeneration at the base (stratum germinatirum) and migrate upward to the surface where they slough off. The mechanoreceptors are located in the deeper, thicker layer (dermis) (Verrillo, 1975).

within a narrow range. Any deviation beyond these limits for long periods of time is deadly. The skin's third function, and the one that we are concerned with in this chapter, is that of sensory communication. The skin houses the receptor end-organs and neural elements that enable us to sense and perceive what is happening in the environment and within the body. These sensory organs enable us to respond appropriately to physical changes in the environment and make possible communication between individuals. Indeed, there is research addressing the role of cutaneous sensation in mental-health and developmental issues. The sensory systems located within cutaneous tissues are not simple affairs, and unlike other sensory systems they do not respond to a single physical stimulus. The visual system responds optimally to contrasts in light, the auditory system to sound waves, taste and smell to chemical configurations. The cutaneous systems respond to a wide variety of environmental energies, including mechanical (touch pressure, pain), thermal (hot and cold), noxious chemicals (pain), and a variety of stimulus blends such as tickle and itch.

In this chapter we will confine ourselves to the responses to mechanical stimuli physically applied to the surface of the body excluding pain and thermal sensation. These characteristics should be considered for the generation of appropriate stimulus signal, the analysis of response signals, or the prediction or simulation of tactile responses to vibration. Practical uses and applications of these characteristics are presented in the Chapter 63 and Chapter 64 discusses sensory substitution techniques.

2 • Neural Structures

Glabrous skin contains at least four types of neural structures specifically designed to respond to the mechanical deformations of its surface. The sensations that result from deformations of the surface are not simple, but are multimodal light and deep touch, pressure, vibration, itch, tickle, and other complex sensory blends. For an historical context see Katz (1925, 1989). First, and without a doubt the most thoroughly studied receptor system, is the *Pacinian corpuscle* (PC). The corpuscle consists of non-neural, concentric layers of tissue (*lamellae*) alternating with spaces filled with fluid. The inner core of the capsule has a more complex configuration. Into the end of this structure is inserted a conventional myelinated axon, which loses its myelin after entering the capsular structure. (*Myelin* is the nerve covering that is essential for the transmission of neural impulses). This is the

neural element that is activated by mechanical deformation and transmits its electrical impulses to the central nervous system. The PCs are located in the deep layers of the dermis and along major nerve trunks and arteries. Human skin contains approximately 2,000 PCs, one third of which are located within the fingers and toes (see Bell et al., 1994 for review).

A second major glabrous receptor is the *Merkel cell-neurite complex*. These are leaf-like, flattened ends of neurons that, in glabrous skin, are clustered at the base of the epidermal layer of skin. In hairy skin they are concentrated into small, dome-like structures located on the leeward side of hairs (sometimes referred to as "touch domes" or *Haarscheiben*). The role of these endings in hairy skin is unclear and there is disagreement that they have a sensory function at all. Some have suggested that they play a role in proprioception, sending information to the cerebellum.

A third structure, located only in glabrous skin, is the *Meissner corpuscle*, which is more highly organized than the Merkel cell-neurite configuration. It is located at the tops of finger-like extensions (dermal papillae) of the dermis that protrude upward into the epidermis. It consists of a connective-tissue capsule within which are nerve fibers that overlap upon themselves from base to apex. The entire structure is suspended, hammock-like, from the sides of the dermal papillae by collagen fibers.

The fourth type of mechanoreceptor found in cutaneous tissues is the *Ruffini end organ*. These consist of nerve fibers, not highly organized, contained within a cylindrical capsule fused with the connective tissues of the dermis (Chambers et al., 1972). There still exists some dispute about this receptor. Although the early publications and standard text books, as well as review articles, accepted that the Ruffini endings are found in both glabrous and hairy skin of mammals, a growing number of anatomist contend that they are located primarily, and most probably, exclusively in hairy skin (Dellon, 1981; Winkelmann, 1960; Paré et al., 2003; Paré et al., 2002; and others).

In addition to these highly organized neural structures, there is the ubiquitous *free-nerve endings* that are located in all types of cutaneous tissue and many other places throughout the body. They have no formal accessory structures but are organized into loosely formed, dense, and irregular networks (rete). They have been claimed by investigators to be responsible for the appreciation of cold, heat, touch, and pain. Most investigators

TABLE 1 Location and function of receptors

Receptor	Location	Function
Free-nerve endings	Ubiquitous at surface and interior of body	Pain, thermal reception, some response to mechanical stimulus. Possible affective aspects of touch
Hair-follicle network	Hairy skin only. Around shaft of hair	Responds to hair movements
Meissner corpuscle	Tops of dermal papillae, glabrous skin only	Maximum sensitivity, 10–30 Hz
Merkel-cell neurite	Base of epidermis, near major nerves and arteries	Broad response to vibration frequencies
Pacinian corpuscle	Deep in dermal layer of skin, near major nerves and arteries	Appreciation of vibratory stimulation, high sensitivity at 250–300 Hz
Ruffini ending	Hairy skin, deep in dermis fused with connective tissue	Relatively insensitive.

associate free-nerve endings with the sensations of pain and itch, although in hairy skin it has been suggested that they mediate the pleasurable aspects of touch. We shall not consider the free-nerve endings as mechanoreceptors in this discourse.

Of course, and in hairy skin only, there is the *hair-follicle network*. This is a highly organized system of nerve fibers surrounding the shaft of the hair follicle. The neural elements are organized into two layers, an inner layer that follows parallel to the hair shaft and an outer layer that circles concentrically around the inner fibers. This arrangement forms a basket-like structure that is ideally constructed to respond to any movement of the hair in any direction. It has been suggested that the sensory function of Meissner corpuscles, which are not found in hairy skin, has been assumed by the hair-follicle network.

3 • Thresholds of Detection in the Temporal and Spatial Domains

3.1 Temporal Domain

One of the most fundamental sensory characteristics in both the auditory and cutaneous sensory systems is the absolute threshold of detection measured as a function of stimulus frequency. Figure 2 illustrates the psychophysical curves of absolute

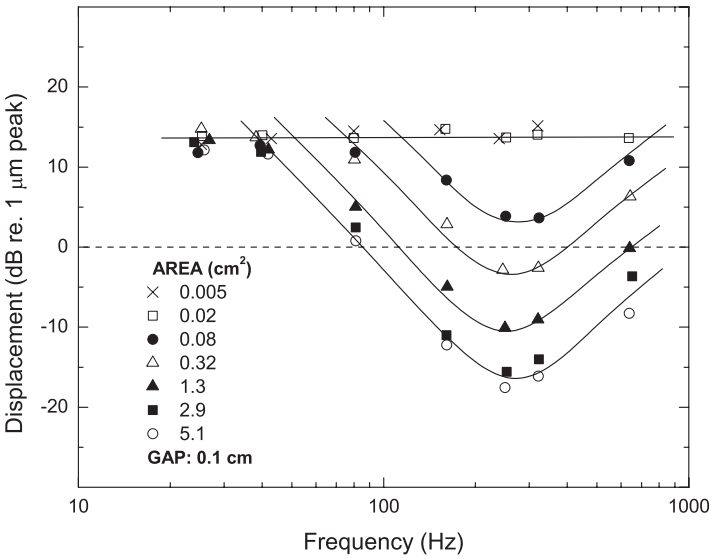


FIGURE 2 Absolute thresholds of detection in dB re 1.0 μm of peak displacement plotted as a function of sinusoidal frequency in hertz. Details of the effects of frequency and contactor size are discussed in the text (Verrillo, 1963).

threshold as a function of sinusoidal frequency measured at the thenar eminence (base of the thumb) of the hand (Verrillo, 1963). Note that the figure contains data obtained with seven different sized contactors. Only the results using the 2.9 cm^2 contactor area will be considered here; the effects of the different sizes will be discussed in Section 3.2.

The data are plotted as displacement in decibels (dB re 1.0 μm peak) as a function of sinusoidal frequency (Hz). It can be seen that the threshold curve (solid squares), after a relatively flat portion in the low frequencies (between 0.4 and 30 Hz), is U-shaped, reaching a maximum of sensitivity in the region of 250–300 Hz before ascending again at higher frequencies. Similar curves were obtained on hairy-skin sites and other regions of the body including the forearm, fingertip, thigh, face, top and bottom of the tongue, and the sole of the foot. On hairy skin the nest of curves is very similar but displaced upward (less sensitive) by 10–20 dB depending upon frequency. By comparing the measured psychophysical responses to the presence of the various receptors innervating different parts of the body and by comparing the data to electrophysiological measurements recorded directly from PCs, it was concluded that the receptors responsible for the U-shaped portion of the psychophysical threshold curve was

indeed the PC. Only those body sites that contain PCs yield the U-shaped curve. The portion of the curve at lower frequencies was ascribed to non-Pacinian receptors that are not dependent upon the frequency of the stimulus. These receptors are not responsive to changes in the temporal domain and are thought to be Meissner corpuscles (Bolanowski, Gescheider, Verrillo and Checkosky, 1988). This dichotomy of receptors became known as the Duplex Model of Mechanoreception (see Verrillo (1968), Section 6, and Figure 11).

An important aspect of auditory (and visual) function is the relationship between the intensity and the duration of the stimulus. It is known that as stimulus duration increases the auditory threshold decreases in an orderly fashion. In fact, in 1960 Zwislocki formulated his theory of temporal summation (Zwislocki, 1960), which gave in precise mathematical terms the relationship between stimulus duration and intensity at the auditory threshold of detection. In order to test Zwislocki's theory within the tactile system, bursts of vibration at several frequencies and at varying durations were delivered to the glabrous skin of the palm (Verrillo, 1965). Figure 3 shows the shift in threshold in decibel as a function of the burst duration in milliseconds.

The solid curve represents the change in threshold as the stimulus duration changes that is predicted by Zwislocki's theory of temporal summation. It clearly demonstrates that at 100, 250, and 500 Hz, the frequencies at which Pacinian corpuscles are responsive to stimulation, the psychophysical threshold data closely approximate the prediction of the theory. Threshold

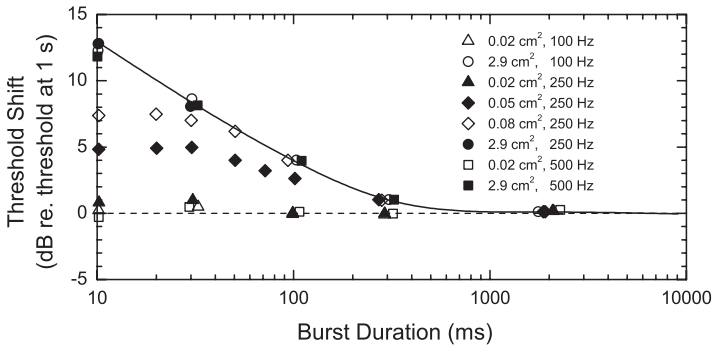


FIGURE 3 Threshold shift in dB re threshold at 1,000 ms plotted as a function of sinusoidal burst duration in milliseconds. The *solid curve* represents the threshold shift predicted by Zwislocki's theory of temporal summation. At frequencies that optimally activate the Pacinian receptors (100,250 and 500 Hz), the data closely obey the predictions of the theory (Verrillo, 1965).

shift decreases as an inverse function of stimulus duration in an orderly fashion reaching an asymptote in the region of 1,000 milliseconds of burst duration. This strongly suggests that the PC system of receptors is able to integrate energy over time. (The data on stimulus size shown in Figure 3 will be considered below.)

Similar experiments were performed using variable pulse-repetition rates of rectangular pulses instead of pure sinusoids. In every case the vibrotactile data supported the predictions made by Zwislocki's theory and were consistent with similar results found in auditory experiments (Verrillo, 1965). In each experiment two modes of response were clearly present. The response characteristics that obeyed the Zwislocki prediction of temporal summation were generated by using stimuli that optimally excited the PC system. The response functions that failed to produce any temporal summation were those in which the stimuli activated the non-Pacinian receptor systems.

3.2 Spatial Domain

Variations in the excitation of receptors over a given area are less complicated in the tactile than in the auditory system. On the skin this is accomplished simply by varying the size of the stimulating contactor. In the auditory system this can be accomplished by varying the band-width of the stimulus and the actual area of stimulation along the basilar membrane is inferred, but never directly observed. In the visual system the scattering of light by the ocular structures and fluids within the eyeball makes difficult the accurate measurement of receptor excitation except by the use of a Maxwellian optics system. Experiments were performed on the skin in which absolute thresholds were measured at different frequencies using a range of contactor areas.

The data in Figure 4 show that at frequencies within the sensitivity range of the PC corpuscle, there is spatial summation; the slopes of the curves are -3 dB per doubling of the area of stimulation indicating a complete integration of energy over space. However, at 25 Hz, a frequency at which there was no temporal summation (see Figure 2), there is also no spatial summation. This is the frequency at which the non-Pacinian receptors are most sensitive. Referring again to Figure 3, the dependence of threshold sensitivity upon the duration of the stimulus is interdependent between frequency and the size of the stimulating contactor. At all frequencies, as the size of the contactor is reduced the dependence of threshold upon the duration of the stimulus is accordingly reduced. The departures from Zwislocki's predictions occur in an orderly fashion. As the principal neural activity shifts from the PC channel to the non-Pacinian systems of receptors, the amount of temporal summation declines accordingly.

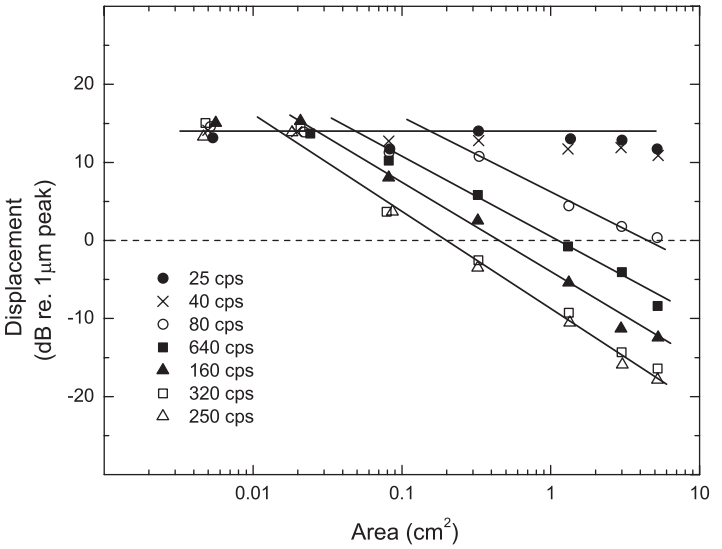


FIGURE 4 Absolute thresholds of detection in dB re 1.0 μm of peak displacement plotted as a function of contactor area in square centimeters. Thresholds at frequencies that optimally excite the Pacinian receptors (80–320 Hz) decrease with increased contactor size at the rate of −3 dB per doubling of contactor area indicating complete summation of energy over space. At frequencies that activate only the non-Pacinian receptors (25 and 40 Hz), there is no spatial summation (Verrillo, 1963).

Thus, it was concluded that there are at least two receptor systems contained in cutaneous tissue; one that is capable of integrating energy over time and space (PCs) and another (non-Pacinian receptors) that is unresponsive to changes in the temporal and spatial domains, although it is noted that other channels may be contributing to somatosensation.

4 • Suprathreshold Responses

The experimental data shown thus far all have been the results of measurements made at the threshold of detection. But, for the most part, we live in a world that is filled with stimulating energies that occur above threshold, i.e., at suprathreshold levels of stimulation. This section concerns the sensory experience when suprathreshold stimuli are applied to the skin. Auditory scientists and practitioners are familiar with the classical equal-loudness contours of Fletcher and Munson (1933) and later modifications made by Robinson and Dadson (1956). Until 1969 these theoretically relevant and practically useful curves were not available to somatosensory scientists. Psychophysical

functions relating sensory magnitude to the physical magnitude of the stimulus were studied for a number of sensory continua by Stevens (1961) who determined that these two variables were related according to a power function, and thus the Stevens *power law* (Stevens, 1955, 1956, 1957) became a benchmark in the history of psychophysics. The method of measurement employed by Stevens and later extended and refined by Hellman and Zwislocki (1963) is known as magnitude estimation and its counterpart, magnitude production (Zwislocki and Goodman, 1980). In a sense the subject is matching the psychological value of numbers to the subjective magnitude of the stimulus. These assigned numbers are plotted as a function of the vibratory amplitude in decibels. The method of magnitude production requires that the subject adjust the physical intensity of the vibration by a variable attenuator to numbers given by the experimenter. The results from the two procedures are then averaged to produce the curve of *numerical magnitude balance*. Figure 5

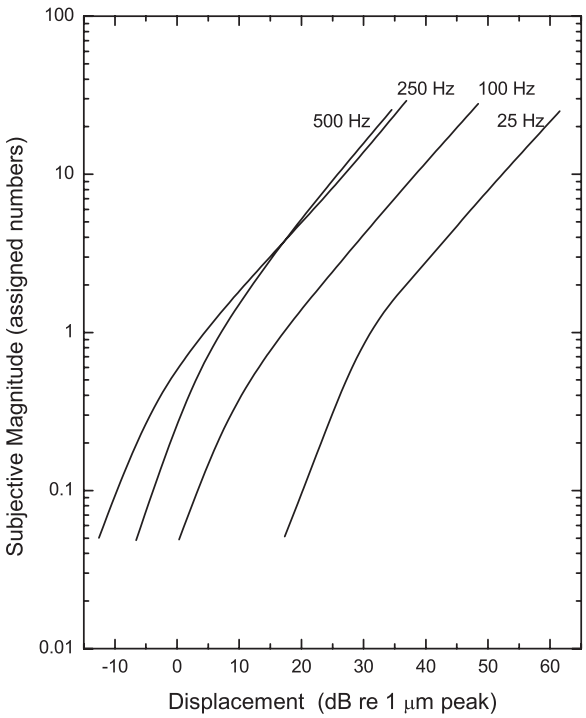


FIGURE 5 Subjective magnitude in assigned numbers plotted as a function of absolute displacement in dB re 1.0 μm peak. Data of just four frequencies are shown for illustrative purposes. All of the data describe power functions (Verrillo et al., 1969).

is an example of results measured at 250 Hz using 10 stimulus amplitudes (Verrillo et al., 1969).

By making appropriate cuts through these curves, a family of equal sensation contours for vibrotaction was generated, which were equivalent to the equal loudness contours in audition.

This family of curves is shown in Figure 6. All of the points along any single curve will result in a sensation of equal subjective magnitude when applied to the skin. A similar set of curves was generated by the older, more time-consuming method, of direct intensity matching. In this procedure the subject is required to adjust a variable attenuator in order to match the subjective intensity of a test stimulus to that of a standard. Thus, all of the frequencies were matched to two standards. When the magnitude balance results were compared to those of the direct intensity matching procedure the two sets of data were practically identical and were combined to produce the curves shown in Figure 6.

In its original form proposed by Stevens (1957) the power function that relates the physical intensity of the stimulus to the subjective stimulus magnitude is stated as:

$$\Psi = k\phi^a$$

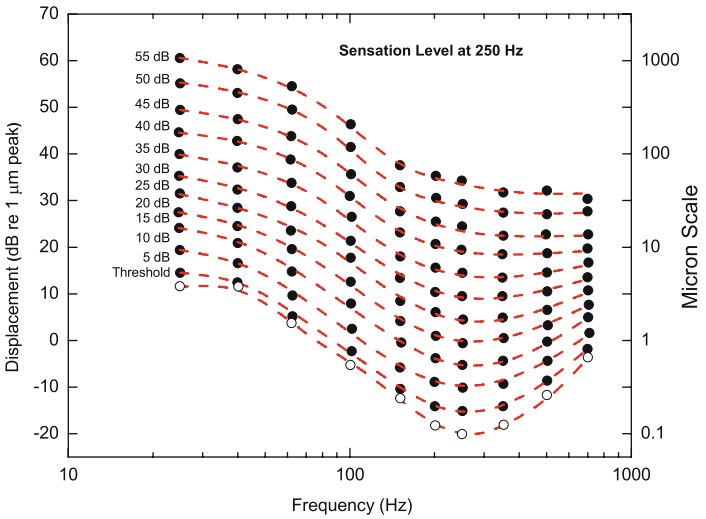


FIGURE 6 Contours of equal subjective magnitude for vibrotaction. Vibratory stimuli applied to the skin at any combination of frequency and displacement along any given curve will produce the same sensory experience (Verrillo, Fraioli and Smith, 1969).

where Ψ is the sensation magnitude, ϕ is stimulus intensity, k is an arbitrary constant, and a is the power exponent that varies with the sensory modality. This formulation underwent a number of modifications by different investigators, but its most general formulation was proposed by Zwislocki (1965). It is stated as:

$$L_s = k(E_s + E_N)^{\theta} - E_N^{\theta}$$

where L_s is the loudness of the signal equal to a constant k times the sum of the energy of the signal E_s and the energy of the noise E_N raised to a power θ, minus the energy of the noise E_N raised to a power θ. This formulation permits the calculation of subjective magnitude by subtracting the magnitude of noise from the magnitude of the signal plus noise. (For a fuller discussion of this formulation the reader is referred to Zwislocki (1965) and Gescheider (1997), Chapter 13.)

5 • Subject Variables

In the design of technologies to assist humans in their reception of vibration, factors such as age, sex, and hand preference can be important. Signal stimulus, analysis, or simulation should take such factors into account because they influence the normal individual responses.

5.1 Age

A topic of enormous importance in this era of lengthening life expectancy concerns the changes in sensory capacity that occur as a function of the normal process of aging. All of the sensory systems suffer deficits with advancing age. Reductions in the auditory and visual systems have been well documented for many years. Eyeglasses have been used for centuries to aid those of advancing years to see as well as they did in their younger years and millions of people today are provided with hearing aids to overcome the hearing deficits that so often accompany advancing years. However, changes in the appreciation of tactile stimuli have been noted, but not studied systematically until relatively recently (Gescheider et al., 1994; Gescheider et al., 1994; Gescheider et al., 1996; Gescheider et al., 1992; Verrillo, 1977; 1979a; 1980; 1982; 1993).

It can be seen in Figure 7 that over a span of between 10 and 70 years of age there is a pronounced loss of sensitivity, but this loss is confined principally at those frequencies that are the primary activators of the PC system of receptors (80–250 Hz). Frequencies at which non-Pacinian receptors are most sensitive show little or no loss of sensitivity as a function of aging. These data were

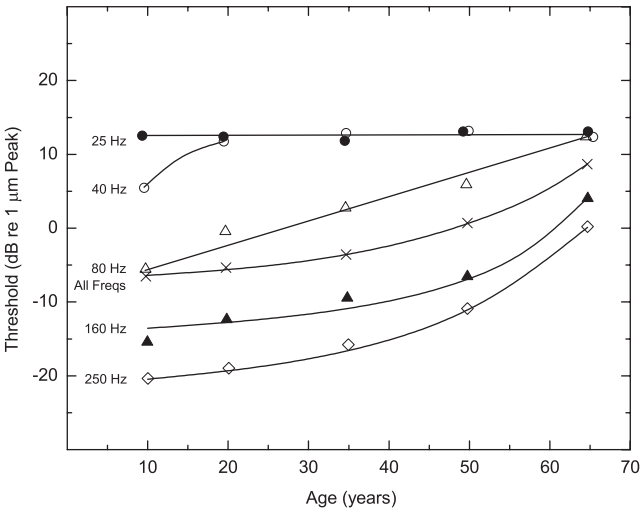


FIGURE 7 Absolute vibrotactile thresholds of detection plotted as a function of age in years. Sensory loss occurs at frequencies mediated by the Pacinian channel; the non-Pacinian receptors do not appear to be affected by age (Verrillo, 1980).

based not only on cross-sectional measurements, but also on a longitudinal study in which the same subjects were tested and then retested after a lapse of years from childhood to adulthood.

Similar experiments were conducted at suprathreshold levels of stimulation (Verrillo et al., 2002). Two groups of subjects were tested, a younger group with a mean age of 23 years and an older group having a mean age of 68 years. The method of absolute magnitude estimation was used. It is apparent in the data shown in Figure 8 that there is a considerable difference between the two groups; for the same intensity of the stimulus (dB re 1 μm peak) the younger group assigned higher numbers to perceived subjective magnitude at every physical level of stimulus intensity. The figure shows data from different subjects tested at times separated by 11 years. It must be concluded that at both threshold and suprathreshold levels of stimulation, advancing age is accompanied by a deficit in vibrotactile function.

Other studies of the effects of the aging on tactile responses report similar deficits for gap detection (Van Doren et al., 1990), masking of one stimulus by another and temporal summation (Gescheider et al., 1992), as well as enhancement of sensation by prior stimulation (Verrillo, 1982). There can be little doubt that the aging process has a negative effect on the cutaneous sensitivity to vibration, especially at frequencies that optimally

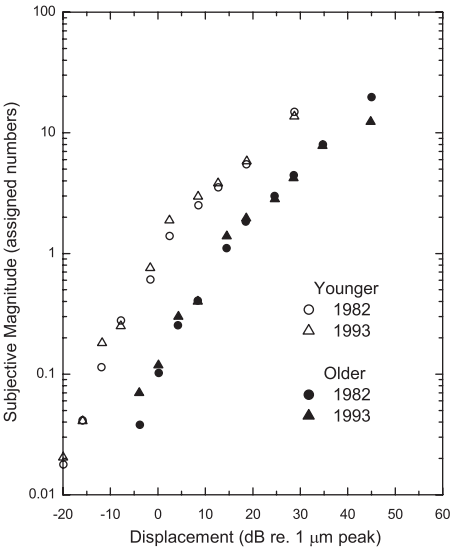


FIGURE 8 Subjective magnitude of vibration at 250 Hz plotted as a function of displacement. The results obtained from two younger groups (mean age 23 years: *open symbols*) are compared to those of two older groups (mean age 68 years; *closed symbols*). The groups were tested 11 years apart (Verrillo et al., 2002).

activate the Pacinian receptor system. Sensitivity to vibration at frequencies that activate non-Pacinian receptors appear not to be affected by the aging process.

5.2 Sex

The scientific and popular literature has for centuries contained a multitude of treatises expounding on the differences between men and women. For instance, in the scientific literature auditory tests have shown that men and women have approximately the same acuity up to age 50, but then decline is more rapid in men than in women (Spoor, 1967) and that the onset of presbycusis is more gradual in women but becomes more rapid than men once started (Corso, 1963).

With regard to tactile function, one study reports a gradual loss of sensitivity in women after the late teens which does not occur in men until after age 45 (Goff et al., 1965). Most studies, however, have not found differences between men and women in vibrotactile thresholds (Plumb and Meigs, 1961; Steinberg and Graber, 1963; Verrillo and Ecker, 1977). Figure 9 shows the results of testing a group of men and women of comparable age and with careful controls of the stimulus conditions (Verrillo, 1979b). It is apparent that with regard to the absolute threshold of detectability there exists no difference between the sexes.

With respect to suprathreshold responses there is a significant difference between men and women (Verrillo, 1979b). Women perceive the same intensity of the physical stimulus as being more intense subjectively than do men (Figure 10). It was suggested that this result may not have been due to a difference

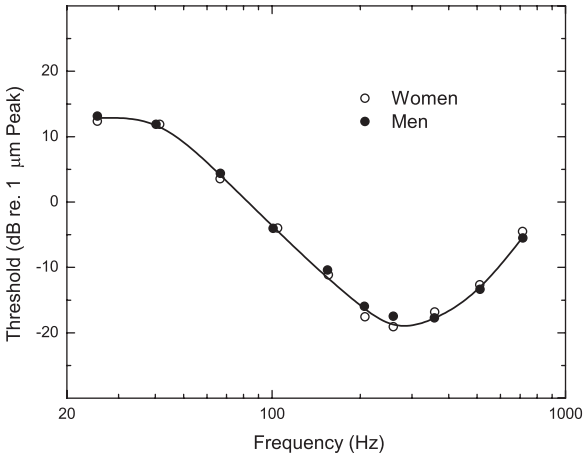


FIGURE 9 Absolute thresholds of detectability of a group of men and women of comparable age. There does not appear to be any difference between the sexes with regard to their sensitivity to vibration (Verrillo, 1979a).

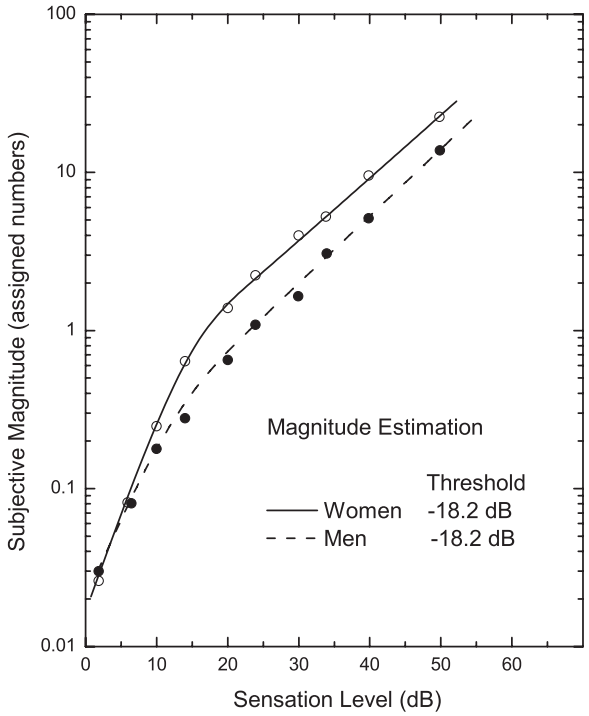


FIGURE 10 Comparison of men and women in the subjective magnitude of vibration. At all levels of stimulation above 6 dB SL, women perceive the same stimulus amplitude as more intense than do men (Verrillo, 1979b).

in sensory function but rather to a difference between sexes in the evaluation or use of numbers. However, when the same experiment was repeated using the judgment of the length of lines (a more "neutral" stimulus than touching the skin), there was no difference in the performances of men and women. The use of numbers was the same in both men and women.

Another sex-related difference that deserves mention is the effect of the menstrual cycle upon sensitivity to vibration. It had long been noticed that the performance of women in threshold experiments was more variable than that of men. One of the reasons for this difference suggested that the variations in the menstrual cycles between women might account for this phenomenon. Two groups of women of comparable age were selected, one that was taking birth control pills and another that was not. The subjects monitored their resting oral temperature each morning to provide an index of the stage of the cycle. Threshold measurements were obtained at the same time each morning to monitor their vibrotactile sensitivity. Two sinusoidal frequencies were tested, 250 Hz to monitor activity within the Pacinian system and 15 Hz to assess non-Pacinian effects. The threshold values at 250 Hz gradually lowered (more sensitive) up to the onset of menstruation and then increased (less sensitive) to a maximum value in 12–13 days later (Gescheider et al., 1984). No such changes were observed in women taking birth control pills. The threshold responses to a 15-Hz signal remained unchained over the entire cycle. It is apparent that changes in the internal environment (probably hormonal) of women over the course of the menstrual cycle has a marked effect upon the response capabilities of the PC receptor and has no effect on the NP receptor systems.

5.3 Hand Preference

A final subject variable worthy of note is that of hand preference. The preferred use of one hand over the other, and the accompanying myths concerning this preference has a long, usually misleading and mostly non-scientific legacy. Because vibrotactile testing is usually performed on a single hand (or side of the body—usually the right), it is important to know if any contralateral differences are present and if they are related to hand preference. Testing revealed that there are no significant differences in absolute sensitivity between the two sides of the body regardless of hand preference. However, when a suprathreshold procedure was performed an unexpected and profound difference was revealed when the task involved judgments of simultaneous contralateral stimuli, i.e. judgments of sensation produced by the stimulation of both sides of the body. There was no difference between right- and left-handed

subjects in this task (Verrillo, 1983). The sensory effects transferred completely from one side of the body to the other. However, in ambi-handed subjects (subjects who used either hand with equal ease) there was no transfer of sensory effects from one side of the body to the other. It was concluded that in right- and left-handed persons there is a transfer of neural activity within the central nervous system from one side of the brain to the other. The results suggest that the two cerebral hemispheres were acting quite independently in ambi-handed individuals. No anatomical or physiological basis for this phenomenon has been discovered.

In addition to subject variables, there are a number of variables of the physical stimulus that can affect the sensations and therefore the responses of the subject in a psychophysical experiment. Thus, these experiments must be carefully controlled for optimal validity and reliability of the data. These factors include the temperature at the surface of the skin (Bolanowski and Verrillo, 1982; Gescheider et al., 1963; Verrillo and Bolanowski 1986), duration of the stimulus (Verrillo, 1965; Verrillo and Smith, 1976), size of the stimulator (e.g., see Figure 2, Verrillo, 1963, 1966a), simultaneous auditory stimulation (Gescheider et al., 1974; Verrillo and Capraro, 1974), and the state of the tissue surrounding the site of stimulation (Gescheider et al., 1978; Verrillo, 1979c).

6 • The Four-Channel Model of Cutaneous Mechanoreception

The history of research on the sensory processes of skin is marked by an ongoing controversy concerning the manner in which stimuli applied to the surface activate the underlying receptors, how these neural impulses are transmitted to the central nervous system, and the fate of these impulses upon their arrival in the brain. Two contrasting theories were proposed. One, called the *specificity theory*, claimed that for each sensation there is a specific receptor end organ uniquely designed to be activated by a specific type of stimulation, heat, cold, touch, pain, etc. The neural impulses generated by this stimulation are carried over neural-transmission lines, equally specific, to sites within the cerebral cortex that are specific in the sensations aroused by their excitation of incoming signals. In contrast, the *pattern* theory denies the existence of specific receptors, transmission lines and areas of the brain. Different kinds of stimulation at the surface of the skin excite non-specific nerve endings producing different patterns of nerve impulses, the pattern being determined by the

nature of the physical stimulus. Transmission of these patterns of excitation is carried over non-specific neural elements to a site in the brain where a spatio-temporal pattern of excitation elicits the conscious sensation. Thus, for every form of physical energy and for each nuance of every form, there is a pattern of neural excitation upon some area of the cortex that interprets these patterns and converts them to consciousness. Unfortunately, the proponents of the pattern theory are vague about the definition of a pattern and their assertion that the specific end organs were mere artifacts of cutting and staining by the anatomist could not stand up in the face of compelling empirical evidence. Few, if any, sensory scientists today would defend the pattern theory of sensation as originally proposed although there is evidence of spatio-temporal patterns of excitation in the brain. The exact origin, nature, and purpose of these patterns is unknown.

In 1966 evidence was reported supporting the notion that the U-shaped portion of the psychophysical threshold function was the result of the activation of PCs (Verrillo, 1966b, c). This idea was later expanded and further experiments resulted in the formulation of a duplex model of cutaneous mechanoreception (Verrillo, 1968). The duplex model proposed that the channels of tactile sensation from glabrous cutaneous tissues are mediated by two receptor systems: the PCs corpuscles at frequencies between approximately 40 and 700 Hz and non-Pacinian receptors at the lower frequencies. This model was later expanded and refined to include three (Capraro et al., 1979; Gescheider et al, 1985) and then four channels, which has become known as the quadruplex model (Bolanowski et al., 1988). The model links physiological and anatomical evidence to identifiable portions of the psychophysically measured absolute threshold of detectability as a function of sinusoidal frequency. The four types of nerve fibers that previously had been clearly identified by electrophysiological measurement are the PC, one rapidly adapting (RA) and two slowly adapting units called SAI and SA II. The four portions of the psychophysical curve were tentatively named the Pacinian (P), non-Pacinian I (NPI), non-Pacinian II (NP II), and non-Pacinian III (NP III) channels. Based on a wide range of experimental results from many different laboratories, it was concluded that the non-painful and non-thermal tactile sensations experienced by humans are the consequence of the activation by mechanical distortion of glabrous skin of four distinct cutaneous receptors. The range of frequencies in these experiments was between 0.4 and 500 Hz. Figure 11 is a compilation of these results, the four-channel model of vibrotaction. The psychophysical nomenclature (NPI, NP II, NP III, and P), the

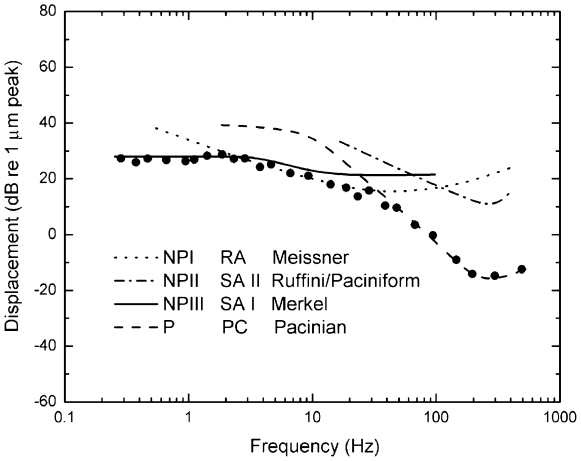


FIGURE 11 Graphic representation of the four-channel model of cutaneous mechanoreception. Shown are the theoretical threshold curves of the psychophysically named channels, the fiber types that have been recorded from, and the receptor end organs that are associated. The *solid data points* are measured results from psychophysical experiments (Verrillo et al., 2000).

type of nerve fiber (RA, SA II, SA I, and PC) and the associated receptor end-organs (Meissner, Ruffini/Paciniform, Merkel and Pacinian) are identified in this figure.

The receptors whose response characteristics most closely approximate these curves, and hence those that have been identified as being responsible for the sensations experienced by humans are (1) the Pacinian corpuscle (P channel) for responses in the U-shaped portion between 30 and 500 Hz (dashed line), (2) Meissner's corpuscle for responses up to approximately 100 Hz with a maximum sensitivity at 30 Hz (fine dotted line), and (3) Merkel cell-neurite complex at the lowest frequencies (solid line). The fourth curve (fine dashed line) is described by a U-shaped function similar to that of the PC, but considerably elevated. This curve was formerly thought to be mediated by the Ruffini end organ, but as noted above no contemporary investigator has been able to find Ruffini endings in glabrous skin. However, recent studies have shown that a receptor resembling the PC exists in glabrous skin. This receptor has been variously referred to as a small Pacinian corpuscle, a paciniform receptor (Malinowsky, 1966) and in other literature as a Golgi-Mazzoni body (Chauchkov, 1973; Dogiel, 1903). It is a layered receptor, as is the PC, but it lacks the inner-core structure that is typical of the PC.

To summarize, the four information bearing channels found in glabrous skin and their purported neural substrates are as follows: (1) The PC channel, mediated by the Pacinian corpuscle with rapidly adapting nerve fibers. This is called the P channel in psychophysical terms. (2) The NP I channel mediated by Meissner corpuscles with RA nerve fibers. (3) The psychophysical NP II channel which is probably mediated by paciniform receptors with slowly adapting fibers (SA II). (4) The NP III channel having slowly adapting (SA I) fibers terminating in the Merkel-cell neurite complex.

7 • Roughness and Texture

The fingers are the primary organs used for exploring the environment in the tactile mode (with the exception of infants who delight in exploring most objects with their mouths). As a stimulus, texture is complex not only in its physical characteristics, but also as a sensory experience because it is comprised of a number of qualities such as smoothness, roughness, softness, hardness, temperature, etc. Each of these qualities is different as a sensation, but they are all perceptually related. Katz (1925, 1989) and Krueger (1970) was the first to seriously consider the physical details of stimuli imposed upon the cutaneous surface as well as the activity or passivity of the subject in judging textural experiences (See Lederman, 1982; Loomas and Lederman, 1987; Greenspan and Bolanowski, 1996 for historical reviews).

The quality of tactual roughness has been studied using a wide variety of stimuli. Sandpaper (Stevens and Harris, 1962; Stone, 1967) etched gratings (Lederman, 1974), and dots raised on a plastic surface (Johnson and Lamb, 1981), all have been used in the effort to provide a precise definition of the stimulus. Judgments of roughness do not appear to be influenced whether the subject actively explores the object or is passive as the object is imposed upon the skin (Lamb, 1983; Lederman, 1981, 1983; Verrillo et al., 1999). The perceived roughness of objects decreases if the temperature at the surface of the skin is lowered below its normal temperature of 32°C, but raising the temperature has little effect (Green et al., 1979). With increased frictional force and wetness, fabrics are judged to be not only rougher but also to be more unpleasant (Gwosdow et al., 1986).

In a study that used a detailed and precise definition of the stimulus, Verrillo et al. (1999) showed that the growth of roughness perception is a power function when assigned numbers (Absolute Magnitude Estimation) are plotted as a function of the particle diameter (in μm) of 11 grades of

sandpaper. Both active and passive modes of stimulation were used at the finger pads of the index finger and the thumb. There is no difference in the perceived roughness of the stimuli whether the subject actively stroked the stimuli or remained passive as the stimuli were drawn across the skin, in agreement with the observations of Lederman (1981, 1983) and Lamb (1983). There were also no differences between the index finger and the thumb alone. However, when the two sites received the stimulation simultaneously, the judgments of roughness increased above those of single-site stimulation. This is a clear indication that there is spatial summation for the sensation of roughness, at least at these two sites.

Subjects were also asked to judge the pleasantness of the sensation that they experienced. The function for perceived pleasantness was inversely related to the perceived roughness of the stimuli. Clearly, pleasant sensations appear to be related to the absence of roughness, in agreement with the finding of Gwosdow et al. (1986). When the stratum corneum was infused with water (hydration) or with a surfactant solution (substances commonly used in soap products), the sensations of roughness were decreased below those determined in normally hydrated skin. This is in agreement with an earlier finding that subjective roughness is directly a function of the degree of hydration; as hydration increases the estimates of roughness also decreases (Verrillo et al., 1998). It was suggested that hydration alters the mechanical properties of the cutaneous tissues by increasing the friction coefficient at the surface. This change in turn has an effect on perceptual judgments. Rather than acting as a lubricant, moisture appears to increase the drag on the skin.

8 • Biophysics

In recent years modern techniques of greater sophistication, such as the video-microscope, have been employed in the investigation of the mechanical properties of cutaneous mechanoreceptors. It had long been assumed that the lamellar structure of the Pacinian corpuscle acted as a band-pass filter to stimuli applied to its surface (Loewenstein, 1971). Thus, the frequency-dependent "tuning" curve exhibited by the PC channel in both physiological and psychophysical experiments was attributed solely to the mechanical properties of the capsular structure surrounding the central neural element. Theoretical models formulated by Hubbard (1958) and Loewenstein and Skalak (1966) concentrated exclusively on the accessory capsule. In contrast, the model proposed by other investigators (Holmes and Bell, 1990;

Bell and Homes, 1992; Bell et al., 1994) suggests that the capsule contributes nothing to mechanical filtering. It is interesting that in every model and in all of the empirical experiments, the skin itself appears to be transparent; its mechanical properties appear to contribute nothing at all to mechanical filtering of the displacements imposed on the surface. Recent experiments confirm that although the amplitude of the displacements is attenuated between the capsule's surface and the neurite, the effect is not significantly dependent upon frequency in any way. It is suggested that the role of the capsule is to provide an environment that promotes the mechano-electric transduction at the neural element within the inner core.

In another aspect of contemporary research, the power of the computer has been employed with increasing effectiveness in the analysis of the behavior of not single, but multiple arrays of receptors acting in concert. Although information about what is happening at the surface of the skin is transmitted to the brain over parallel lines coming from an array of receptors, our techniques for accessing neural activity are limited to recording from a single nerve fiber at a time. This is a serious constraint on our understanding of cutaneous mechanoreceptor systems since most, if not all, environmental stimuli will activate more than a single receptor. Mountcastle (1957) proposed that the population response could be determined from the activity of a single fiber by moving the stimulus across the receptive field of that fiber. He called the method "reciprocal interpretation." A host of electrophysiologists have used this method of investigation (see reviews by Johnson and Hsiao (1992) and Greenspan and Bolanowski (1996)). Two problems accompany attempts to reconstruct the population response of receptors from single-unit recordings. One is the assumption that the receptors are uniformly distributed beneath the surface of the skin and the second is that the receptors are equidistant from each other. This makes it difficult to evaluate models such as the "spatial event plots" of Phillips et al. (1992). Glüçlü and Bolanowski (2002, 2003) attempted to overcome these difficulties by simulating three hypothetical distributions, namely rectangular, uniformly random, and proximo-distally Gaussian. They studied the effects of stimulus amplitude, average innervation density, and contractor location upon the response measures of the number of active fibers, summated firing rate, and the average firing rate. The results indicated that in addition to the stimulus parameters there was a significant effect of the anatomical organization of the receptors. The authors suggest that the organization of cutaneous receptors must be studied more

intensively. The same authors applied their model to data from cat and monkey (Johnson, 1974) paws and in both cases found the receptor distribution to be lognormal rather than normally distributed as assumed by models based on the reciprocal interpretation principle. It is important to note that problems still exist in producing a definitive model that characterizes the response of receptor populations because the innervation density of nerve fibers (Johansson and Vallbo, 1979) differs from that of receptors (Bolton et al., 1966; Bollanowski and Pawson, 2003), which strongly suggests that one nerve fiber innervates multiple receptors. It may be concluded that the assumption of a normal distribution of receptors is unfounded, that models based on this assumption (such as reciprocal innervation) are not adequate to characterize the responses of receptor populations, that results of experimentation are complicated by the multiple innervation of receptors by a single nerve fiber, and finally that our limited knowledge concerning the anatomical distribution of cutaneous mechanoreceptors must be expanded before a definitive model of population responses may be accepted.

Acknowledgements The authors express their deep appreciation to Lorraine Pawson for her invaluable assistance with the figures and to Sarina Burkett for her prompt and patient work in typing the manuscript. David Havelock deserves special mention for his timely responses to the many inquiries of the authors and his extensive help with the figures that went far beyond the call of duty.

References

Bell, J., Bolanowski, S.J., Holmes, M.H.: The structure and function of Pacinian corpuscles: A review. Prog. Neurobiol., 4, 79–128, (1994).

Bell, J., Holmes, M.: Model of the dynamics of receptor potential in a mechanoreceptor. Math. Biosci., 110, 139–174, (1992).

Bolanowski, S.J., Gescheider, G.A., Verrillo, R.T., Checkosky, C.M.: Four channels mediate the mechanical aspects of touch. J. Acoust. Soc. Am., 84, 1680–1694, (1988).

Bolanowski, S.J., Pawson, L.: Organization of Meissner corpuscles in the glabrous skin of monkey and cat. Somatosens. Mot. Res., 20, 223–231, (2003).

Bolanowski, S.J., Jr., Verrillo, R.T.: Temperature and criterion effects in the somatosensory system: A neurophysiological and psychophysical study. J. Neurophysiol., 48, 837–856, (1982).

Bolton, C.F., Winkelmann, R.K., Dyck, P.J.: A Quantitative study of Meissners corpuscles in man. Neurologica, 16, 1–9, (1966).

Capraro, A.J., Verrillo, R.T., Zwislocki, J.J.: Psychophysical evidence for a triplex system of cutaneous mechanoreception. Sens. Processes, 3, 252–334, (1979).

Chambers, M.R., Andres, K.H., Duering, M.von, Iggo, A: The structure and function of the slowly adapting Type II mechanoreceptor in hairy skin. Q. J. Exp. Physiol., 57, 417–445, (1972).
Chouchkov, Ch.N.: The fine structure of small encapsulated receptors in human digital glabrous skin. J. Anat., 114, 25–33, (1973).
Corso, J.F.: Age and sex differences in pure-tone thresholds. Arch. Otolaryngol., 77, 385–405, (1963).
Dellon, A.L.: Evaluation of Sensibility and Re-education of Sensation in the Hand. Baltimore: Williams and Wilkins, (1981).
Dogiel, A.S.: Uber die Nervendapparate. in: der Haut des Menschen. Zeit. Wissenschaft. Zoo., 75, 46–110, (1903).
Fletcher, H., Munson, W.A.: Loudness, its definition, measurement, and calculation. J. Acoust. Soc. Am, 5, 82–108, (1933).
Gescheider, G.A.: Psychophysics: The Fundamentals. Mahwah, N.J., Lawrence Erlbaum, (1997).
Gescheider, G.A., Beils, E.J., Bolanowski, S.J., Checkosky, C.M., Verrillo, R.T.: Effects of aging on information-processing channels in the sense of touch: II. Temporal summation in the P channel. Somatosens. Mot. Res, 11, 359–365, (1994).
Gescheider, G.A., Bolanowski, S.J., Hall, K.L., Hoffman, K.E., Verrillo, R.T.: Effects of aging on the information–processing channels in the sense of touch: I. Absolute sensitivity. Somatosens. Mot. Res., 11, 345–357, (1994).
Gescheider, G.A., Capraro, A.J., Frisina, R.D., Hamer, R.D., Verrillo, R.T.: The effects of a surround on vibrotactile thresholds. Sens. Processes, 2, 99–115. (1978).
Gescheider, G.A., Edwards, R.R., Lackner, E.A., Bolanowski, S.J., Verrillo, R.T.: The effects of aging in the information-processing channels in the sense of touch: III. Differential sensitivity to changes in stimulus intensity. Somatosens. Mot. Res., 13, 73–80, (1996).
Gescheider, G.A., Kane, M.J., Sager, L.C., Ruffolo, L.J.: The effect of auditory stimulation on responses to tactile stimuli. Bull. Psychon. Soc., 3, 204–206, (1974).
Gescheider, G.A., Sklar, B.F., Van Doren, C.L., Verrillo, R.T.: Vibrotactile forward masking: Psychophysical evidence for a triplex theory of cutaneous mechanoreception. J. Acoust. Soc. Am., 78, 534–543, (1985).
Gescheider, G.A., Thorpe, J.M., Goodarz, J., Bolanowski, S.J.: The effects of skin temperature on the detection and discrimination of tactile stimulation. Somatosens. Mot. Res., 14, 181–188, (1963).
Gescheider, G.A Valetutti, A.A., Jr., Padula, M., Verrillo, R.T.: Vibrotactile forward masking as a function of age. J. Acoust. Soc. Am., 91, 1690–1696, (1992).
Gescheider, G.A., Verrillo, R.T., McCann, J.T., Aldrich, E.M.: Effects of the menstrual cycle on vibrotactile sensitivity. Percep. Psychophys., 36, 586–592, (1984).
Goff, G.D., Rosner, B.S., Detre, T., Kennard, D.: Vibration perception in normal man and medical patients. J. Neurol. Neurosurg. Psychiatr., 18, 503–509, (1965).

Green, B.G., Lederman, S.J., Stevens, J.C.: The effect of skin temperature on the perception of roughness. Sens. Processes, 3, 327–333, (1979).
Greenspan, J.D., Bolanowski, S.J.: The psychophysics of tactile perception and its peripheral physiological basis. in: Krueger, L. (ed.) Handbook of Perception and Cognition Vol. 7: Pain and Touch, Chicago: Academic Press, (1996).
Güçlü, B., Bolanowski, S.J.: Modeling population responses of rapidly-adapting mechanoreceptive fibers. J. Comput. Neurosci., 12, 201–218, (2002).
Güçlü, B., Bolanowski, S.J.: Distribution of the intensity-characteristic parameters of cat rapidly adapting mechanoreceptive fibers. Somatosens. Mot. Res., 20, 149–155, (2003).
Gwosdow, A.R., Stevens, J.C., Berglund, L.G., Stolwijk, J.A.J.: Skin friction and fabric sensation in neutral and warm environments. Textile Res. J., 56, 574–580, (1986).
Hellman, R.P., Zwislocki, J.J.: Monaural loudness function of a 1,000-cps tone and interaural summation. J. Acoust. Soc. Am., 36, 1618–1627, (1963).
Holmes, M., Bell, J.: A model of a sensory mechanoreceptor derived from homogenization. J. Appl. Math., 50, 147–166, (1990).
Hubbard, S.J.: The mechanical properties of Pacinian corpuscles. J. Physiol., (London), 141, 198–218, (1958).
Johnson, K.O.: Reconstruction of population response to a vibratory stimulus in quickly adapting mechanoreceptive afferent fiber population innervating glabrous skin of the monkey. J. Neurophysiol, 37, 48–72, (1974).
Johnson, K.O., Hsiao, S.S.: Neural mechanisms of tactual form and texture perception. Ann. Rev. Neurosci., 15, 227–250, (1992).
Johnson, K.O., Lamb, G.D.: Neural mechanisms of spatial tactile discrimination: Neural patterns evoked by Braille-like dot patterns in the monkey. J. Physiol., 310, 117–144, (1981).
Johansson, J.S., Vallbo, A.B.: Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. J. Physiol., 280, 283–300, (1979).
Katz, D.: Der Aufbau derTastwelt. Zeitshrift Psychologia, 11, 270, (1925).
Katz, D.: The World of Touch. Translation by L.E. Krueger, Hillsdale, NJ: Lawrence Erlbaum Associates, (1989).
Krueger, L.: David Katz's Der Aufabu der Tastwelt (The world of touch): A synopsis. Percep. Psychophys., 7, 337–341, (1970).
Lamb, G.D.: Tactile discrimination of textured surfaces: Psychophysical performance measurements in humans. J. Physiol. (London), 338, 551–565, (1983).
Lederman, S.J.: Tactile roughness of grooved surfaces: the touch process and effects of macro and microsurface structures. Percep. Psychophys., 16, 385–395, (1974).
Lederman, S.J.: The perception of surface roughness by active and passive touch. Bull. Psychon. Soc., 18, 253–255, (1981).
Lederman, S.J.: Tactual roughness perception: spatial and temporal determinants. Can. J. Psychol., 37, 498–511, (1983).

Lederman, S.J.: The perception of texture by touch. in: Schiff, W., Foulke, E. (eds.) Tactile Perception: A Sourcebook, 130–167 New York: Cambridge University Press, (1982).

Loewenstein, W.R.: Mechano-electric transduction in the Pacinian corpuscle. in: Loewenstein, W.R. (ed.), Handbook of Sensory Physiology. Vol. I. Principles of Receptor Physiology, 269–290. Berlin, Springer, (1971).

Loewenstein, W.R., Skalak, R.: Mechanical transmission in the Pacinian corpuscle. An analysis and a theory. J. Physiol, (London), 182, 346–378, (1966).

Loomas, J.M., Lederman, S.J.: Tactual perception. in: Boff, K., Kaufman, L., Thomas, J. (eds.) A Handbook of Human Perception and Performance, Vol. III, 31–41, New York: John Wiley and Sons, New York, 1987.

Malinowsky, L.: Variablility of sensory nerve endings in foot pad of a domestic cat (Felis Ocreata L., F. Domestica) Acta Anatom., 64, 82–106, (1966).

Mouncastle, V.B.: Modality and topographic properties of single neurons of cat's somatic sensory cortex. J. Neurophysiol., 20, 408–438, (1957).

Paré, M, Behets, C., Cornu, O.: Paucity of presumptive Ruffini corpuscles in the index finger pad of humans. J. Comp. Neurol., 456, 260–266, (2003).

Paré, M., Smith, A.M., Rice, F.L.: Distribution and terminal arborizations of cutaneous mechanoreceptions in the glabrous finger pads of the monkey. J. Comp. Neurol., 445, 347–359, (2002).

Phillips, J.R., Johansson, R.S., Johnson, K.O.: Responses of human mechanoreceptive afferents to embossed dot arrays scanned across fingerpad skin. J. Neurosci., 12, 827–839, (1992).

Plumb, C.S., Meigs, J.W.: Human vibration perception: Part I. Vibration perception at different ages. Arch. Gen. Psychiatry, 4, 611–614, (1961).

Robinson, D.W., Dadson, R.S.: A redetermination of the equal-loudness relations for pure tones. Brit. J. App. Phys., 7, 166–181, (1956).

Spoor, A.: Presbyacusis values in relation to noise induced hearing loss. Int. Audiol., 6, 48–57, (1967).

Steinberg, F.U., Graber, A.L.: The effect of age and peripheral circulation on the perception of touch. Arch. Phys. Med. Rehab., 44, 645–650, (1963).

Stevens, S.S.: The measurement of loudness. J. Acoust. Soc. Am., 27, 815–829, (1955).

Stevens, S.S.: The direct estimate of sensory magnitude-loudness. Am. J. Psychol., 69, 1–25, (1956).

Stevens, S.S.: On the psychophysical law. Psychol. Rev., 64, 153–181, (1957).

Stevens, S.S.: The psychophysics of sensory function. in Rosenblith, W.A. (ed.) Sensory Communication, 1–33, New York: John Wiley and Sons; Cambridge: M.I.T. Press, (1961).

Stevens, S.S., Harris, J.R.: The scaling of subjective roughness and smoothness. J. Exp. Psychol., 46, 489–494, (1962).

Stone, L.A.: Subjective roughness and smoothness for individual judges. Psychon. Sci., 9, 347–348, (1967).
Van Doren, C.L., Gescheider, G.A., Verrillo, R.T.: Vibrotactile temporal gap detection as a function of age. J. Acoust. Soc. Am., 87, 2201–2206, (1990).
Verrillo, R.T.: Effect of contactor area on the vibrotactile threshold. J. Acoust. Soc. Am, 35, 1962–1966, (1963).
Verrillo, R.T.: Temporal summation in vibrotactile sensitivity. J. Acoust. Soc. Am., 37, 843–846, (1965).
Verrillo, R.T.: Effect of spatial parameters on the vibrotactile threshold. J. Exp. Psychol., 71, 570–575, (1966a).
Verrillo, R.T.: Vibrotactile sensitivity and the frequency response of the Pacinian corpuscle. Psychon. Sci., 4, 135–136, (1966b).
Verrillo, R.T.: Specificity of a cutaneous receptor. Percept. Psychophys. 1, 149–153, (1966c).
Verrillo, R.T.: A duplex mechanism of mechanoreception, in: Kenshalo, D.R. (ed.) The Skin Senses, 139–159, Springfield, Ill: Charles C. Thomas., (1968).
Verrillo, R.T.: Chapter 3 – "General Biology of Sensory Systems – The Somatosensory System". In: Scharf, B. (ed.) Experimental Sensory Psychology, Scott, Foresman, Glenview, (IL), pp. 101–107, (1975).
Verrillo, R.T.: Comparison of child and adult vibrotactile thresholds. Bull. Psychon. Soc., 9, 197–200, (1977).
Verrillo, R.T.: Change in vibrotactile thresholds as a function of age. Sens. Processes, 3, 49–59, (1979a).
Verrillo, R.T.: Comparison of vibrotactile threshold and suprathreshold responses in men and women. Percept. Psychophys., 26, 20–24, (1979b).
Verrillo, R.T.: The effect of surface gradients on vibrotactile thresholds. Sens. Processes, 3, 27–36, (1979c).
Verrillo, R.T.: Age-related changes in the sensitivity to vibration. J. Gerontol., 35, 185–193, (1980).
Verrillo, R.T.: Effects of aging in the suprathreshold responses to vibration. Percept. Psychophys., 3, 61–68, (1982).
Verrillo, R.T.: Vibrotactile subjective magnitude as a function of hand preference. Neuropsychologia, 21, 383–395, (1983).
Verrillo, R.T.: The effects of aging on the sense of touch. In Verrillo, R.T. (ed.) Sensory Research: Multimodal Perspectives. 260–275. Hillsdale, NJ: Lawrence Erlbaum, (1993).
Verrillo, R.T., Bolanowski, S.J., Jr.: The effects of skin temperature on the psychophysical responses to vibration on glabrous and hairy skin. J. Acoust. Soc. Am., 80, 528–532, (1986).
Verrillo, R.T., Bolanowski, S.J., Checkosky, C.M., and McGlone, F.B.: Effects of hydration on tactile sensation. Somatosens. Mot. Res., 15, 93–108, (1998).
Verrillo, R.T., Bolanowski, S.J., Gescheider, G.A.: Effects of aging on the subjective magnitude of vibration. Somatosens. Mot. Res., 19, 238–244, (2002).

Verrillo, R.T., Bolanowski, S.J., McGlone, F.P.: Subjective magnitude of tactile roughness. Somatos. Mot. Res., 16, 352–360, (1999).
Verrillo, R.T., Bolanowski, S.J., McGlone, F.P: Cutaneous spatial perception. In Bonnet C. (Ed.) Fechner Day 2000, International Society for Psychophysics, Strasbourg, pp. 351–355, (2000).
Verrillo, R.T., Capraro, A.J.: Effect of simultaneous auditory stimulation on vibrotactile thresholds. Percept. Psychophys., 16, 597–600, (1974).
Verrillo, R.T., Ecker, A.: Effects of root or nerve destruction on vibrotactile sensitivity in trigeminal neuralgia. Pain, 3, 239–255, (1977).
Verrillo, R.T., Fraioli, A.J., Smith, R.L.: Sensation magnitude of vibrotactile stimuli. Percept. Psychophys., 6, 366–372, (1969).
Verrillo, R.T., Smith, R.L.: Effect of stimulus duration on vibration sensation magnitude. Bull. Psychon. Soc., 8, 112–114, (1976).
Winkelmann, R.K.: Nerve Endings in Normal and Pathological skin. Springfield: Charles C. Thomas, (1960).
Zwislocki, J.: Theory of temporal auditory summation. J. Acoust. Soc. Am., 32, 1046–1060, (1960).
Zwislocki, J.J.: An analysis of some auditory characteristics in: Luce, R.D., Galanter, E. (eds.) Handbook of Mathematical Psychology (Vol. III) 1–97 New York, Wiley, (1965).
Zwislocki, J.J., Goodman, D.A.: Absolute scaling of sensory magnitude. Percept. Psychophys., 28, 28–30, (1980).

63

Uses and Applications of Somatosensation

Ronald T. Verrillo

Institute for Sensory Research, Syracuse University, NY 13244, USA

1 • Introduction

The previous chapter presented details of the fundamental characteristics of somatosensation. In this chapter, the focus will be on the practical uses and applications of these characteristics. Of what use is all of the information about basic somatosensory characteristics and how does it apply to real-life situations that humans are exposed to? Among the issues to be addressed are the uses of the cutaneous surface as a substitute channel of communication, the effects of multiple sensory stimulation and the effects of stimulation upon the health and well being of humans.

The use of somatosensation as a surrogate channel of input from the environment may well go back to the days when humans lived in caves and a person with a visual deficit picked up a stick and tapped his or her way to avoid the perils that he could not see. The white cane of blind person is an affective tool that is still in common use. A myriad of devices, including the use of echolocation, have been developed to signal environmental dangers to those who cannot see. Some of these devices provide feedback to the auditory system with high and low frequencies coupled with intensity to locate and signal the proximity of unseen objects. Other devices couple this information to the vibrotactile modality with stimulation to the hand or other parts of the body.

Negotiating the environment is not the only problem confronting the blind. Of equal importance is the ability to read the written word. The system of raised dots developed by Louis Braille in 1829 was a giant step towards a solution to this problem giving the blind access to current events through the written word although a day or perhaps a week late. However, this method was slower than reading by sighted individuals and required an intermediate step of translating from a visual stimulus to the Braille system. Another important aspect of this problem is the issue of privacy. The blind may have access to the news and to literature via the Braille system, but the utility bills, etc. do not have a Braille translation, a sighted person must intervene and privacy is forfeit.

2 • Intermodality Interactions

The affects of stimulation in one modality upon the responses of another sensory system have become increasingly important. Consider the myriad of instruments that confront the pilots of our commercial airlines, military aircraft and space vehicles. It is essential that the input to one system upon another be known. Most of the studies concerning intermodality effects have involved vision and hearing. Although a rather large literature exists concerning these interactions, the findings are far from consistent. Depending on a variety of parameters, sensory deficit or improvement may be produced in one modality by the simultaneous stimulation of another (Loveless et al., 1970). For obvious reasons, the military have been particularly active in pursuing this line of research.

The concerns of many investigators have been directed to the effects of redundant information in vigilance tasks (Buchner, 1963; Osborn et al., 1963). The basic question that has been addressed is whether the effects of multiple sensory inputs can be accounted for by specific neural interactions or merely by statistical outcomes (Brown and Hopkins, 1967). The literature on the effects of interactions between vibrotaction and other sense modalities is sparse. However, the effects of temperature on cutaneous sensations have been well documented, including tongue sensitivity (Green, 1987), tactile acuity (Stevens, 1982), perception of roughness (Green et al., 1979), two-point threshold (Stevens, 1989) and object recognition by touch (Stevens and Hooper, 1982). Decrements of vibrotactile sensitivity in cooled skin have been reported (Weitz, 1941; Békézy, 1957; Green, 1977). It remains uncertain whether the results were due to the direct action of temperature upon the sensory end organs

and their surrounding tissues or to more central neural interactions. Sato (1961) and Loewenstein (1961) also reported changes in the electrical activity in Pacinian corpuscles responding to mechanical stimulation as a result of cooling.

Psychophysical studies show that the Pacinian system is affected detrimentally by a reduction in skin temperature but that the non-Pacinian systems (frequencies below 40 Hz) are only mildly affected (Verrillo and Bolanowski, 1986). Consistent with this result is the finding that at suprathreshold levels of stimulation, the Pacinian system suffers a significant loss of sensitivity as a result of lowered skin temperature whereas the non-Pacinian systems do not (Verrillo and Bolanowski, 2003). As the temperature was decreased from 40°C to 15°C, the perceived magnitude of the sensation in the Pacinian system decreased in an orderly fashion.

Multiple tactile stimuli applied to different parts of the body also has an effect. Verrillo et al. (1999) found that there was no difference in perceived roughness of surfaces when estimates were made by the thumb and forefinger alone. However, when the two sites were stimulated simultaneously, the estimates of perceived roughness increased indicating that some sort of spatial summation was occurring at these sites. Also, a stimulus applied to one part of the body can effectively reduce (mask) the intensity of another stimulus applied at a remote site (Verrillo et al., 1983; Ferrington et al., 1977). However, this effect was only observed when both stimuli were within the same receptor channel, Pacinian or non-Pacinian. Cross-channel effects were never observed.

Of a more serious nature is the fact that military and civilian divers when exposed to low-frequency, water-borne vibrations report adverse reactions such as conflicting and confusing tactile sensations that induce distractions from underwater tasks. This can easily lead to a decrease in work performance and may constitute a hazard to the health of the diver. This problem led to a series of experiments in which vibrotactile thresholds were measured under the environmental conditions encountered by Navy and civilian submariners. These conditions included altered atmosphere pressure from 1.0 atmospheric absolute (ATA) to 10.1 ATA (the equivalent of a 300 foot dive) and altered breathing mixtures normal for dive depths. Measurements were made before, during and after 6- and 8-day saturation dives in a dry hyperbaric chamber. The results indicated that neither ATA nor breathing mixture had any effect upon vibrotactile thresholds in either the Pacinian or non-Pacinian receptor systems (Verrillo et al., 1996). Measurements were also made to assess the effect

of skin hydration on tactile sensitivity. It was determined that submersion in distilled water and in a sea water solution for 20 min had no effect on the threshold of detection in the Pacinian or non-Pacinian channels on either the glabrous skin of the hand or hairy skin of the forearm. In an extension of these studies, measurements of vibrotactile thresholds as well as suprathreshold magnitude estimates were made under various controlled levels of hydration (Verrillo et al., 1998). In addition, suprathreshold estimates of roughness were determined under the same conditions. The results showed that neither vibrotactile thresholds nor estimates of vibrotactile sensation magnitude were influenced by skin hydration. However, the perceptions of roughness were substantially altered; the sensations of roughness were dramatically less intense in hydrated skin when compared to the results obtained in air (unhydrated). It must be concluded that the mechanics of the skin as well as hydration play a major role in perception of textured (spatio-temporal) surfaces but no such relationship exists in the predominantly temporal (vibration) domain.

Thermal (heat) pain induced upon the surface of the skin can substantially alter its sensitivity to vibratory stimuli (Apkarian et al., 1994). When a tonic, pain-producing heat was presented to the hand, diminished sensitivity to vibration occurred just below and just above the temperature that produced pain. Furthermore, there was decreased sensitivity to suprathreshold stimuli at the same temperatures. It may be concluded that pain induced by heat can significantly diminish the sense of touch at threshold and at suprathreshold levels.

It also has been shown that simultaneous vibrotactile and auditory stimulation can have a detrimental effect on the vibrotactile response (Gescheider and Niblette, 1967; Gescheider et al., 1969). Verrillo and Capraro (1974) found that continuous auditory stimulation had little effect on vibrotactile thresholds, whereas auditory signals pulsed to coincide temporally with vibratory signals produced a marked deficit in the detection thresholds for vibration.

To our knowledge, interactions between vibrotactile and visual stimuli have not been systematically investigated.

3 • Sensory Substitution

Although the sense of touch has been used as a surrogate for both the visual and the auditory modalities, the primary focus of investigators has been on the efforts to develop a tactile substitute for

hearing, especially for that of human speech. Speech is a complex, sophisticated and subtle form of communication without parallel in the lower animals. A severe handicap is imposed upon those whose ability to use it is limited. Speech acquisition is normally predicated upon a feedback system between the ear and the apparatus for producing speech sounds. Although hand signing alleviated the problem between deaf people, there still exists a considerable gap in communication between the world of the deaf and the hearing. The question of tactile communication will be considered in detail in Chapter 64.

4 • Health

This section includes the uses of vibrotaction in a medical setting and the effects of high-intensity vibratory stimulation on the health of an individual. Testing for the sensitivity to vibration is a part of the standard repertoire of the general practitioner's physical examination, along with the Snellen eye chart and the ticking watch, all of them cursory tests at best. The test for vibration sensitivity in many medical settings consists of applying a turning fork to the anklebone of the patient until the patient no longer feels it and then applying it to the physician's anklebone. If the physician cannot feel it (equality of sensation), the patient is deemed "normal", but if the physician can feel it, the patient is said to have a deficit in peripheral nerve endings or conduction pathways. This is predicated on the assumption that the physician is "normal". Crude, but still practiced!

There are a number of physical conditions and disease entities that can affect the functioning of the tactile systems including diabetes mellitus, carpal tunnel syndrome and focal dystonia. In my opinion, some of these have not been systematically investigated, with adequate controls and under carefully controlled laboratory conditions.

One of the sensory consequences of diabetes is the loss of sensitivity in the peripheral extremities, particularly in the toes and feet. The loss can be precursor to the gradual loss of vision and eventual blindness. To my knowledge, there has been no study of this phenomenon to date in which the age of onset and longevity of the condition by group were controlled and the vibrotactile functioning at threshold and suprathreshold levels was assessed under controlled laboratory conditions. These studies have not been performed in spite of the fact that over 30 % of those seeking aid for assistance from organizations for the blind are suffering from the effects of diabetes.

Two disease entities have shown alterations in vibrotactile functioning. In one of them, trigeminal neuralgia ("tic douloureux"), the patient suffers from a paroxysmal, excruciating pain on the face in response to a slight brushing of the skin in a "trigger" zone, typically on the cheek near the nose (Verrillo and Ecker, 1977; Lindblom and Verrillo, 1979). The condition is usually unilateral. There are a number of medical procedures that appear to alleviate the affliction, including pharmaceuticals, surgery and injection into cerebral ganglia that serve the face, all of which achieve approximately 80%–85% success. In one study, a group of patients was tested on the normal and afflicted sides of the face before and after a procedure in which a portion of the Gasserian ("semi-lunar") ganglion within the brain was destroyed to alleviate the pain by alcohol injection (alcoholic Gasserian rhizolysis). The result showed that there was essentially no difference in vibrotactile thresholds between the normal and the afflicted sides of the face prior to surgery. Following the rhizolysis, there was no sensation on the operated side. A number of these patients underwent a procedure (neurectomy) in which a section 4 cm in length of the infraorbital nerve was removed in order to determine whether they wished to have the irreversible rhizolysis performed. Because peripheral nerves regenerate, these patients were tested before the neurectomy and at prescribed intervals up to 1 year beyond. The contralateral side of the face was the control condition. Soon after the neurectomy, the surgical side was totally insensitive, but there was gradual return of sensitivity up to 1 year beyond surgery when thresholds were again at preoperative levels; the removed section of nerve had fully regenerated. An interesting and unexpected side effect was that during the period of recovery, the unoperated side of the face also showed loss of sensitivity. Standard anatomy texts state unequivocally that innervation of the face is strictly unilateral, that nerves from one side never crossover to the other side, yet the destruction of nerve on one side of the face had marked effect on contralateral site, an unexpected and intriguing result that belies an accepted dogma.

There is a neurological condition first fully described during the Civil War in the United States, then called "causalgia" and now known as chronic neuralgia or reflex sympathetic dystrophy. It is the result of traumatic tearing of nerves by gunshot, shrapnel or similar means. (A clean cut by scalpel will not result in this condition.) Thus, most of these cases are recorded during and after armed conflict. The major symptom occurs after full recovery from the wound and normal motor function is restored. In spite of this recovery, a severe pain persists in the region such

that the slightest brushing or touch to the afflicted region will result in severe, paroxysmal pain, usually described as a burning, throbbing or an aching pain.

Patients were tested at threshold and suprathreshold levels of stimulation on the afflicted and at a homologous site on the contralateral normal side (self-control). Vibrotactile thresholds were elevated (less sensitive) in the area around the spots of intense pain (trigger zones). This confirmed the patient's report of numbness in these areas. Also, estimates of perceived stimulus magnitude produced power functions with steeper curves on the afflicted sides. This might be considered as analogous to the phenomenon knows as "recruitment" observed in certain types of hearing loss usually associated with neural pathologies within the auditory system.

Carpal tunnel syndrome is a nerve-compression condition resulting in severe pain in the hand and wrist brought on by repetitive or prolonged mechanical strains to the region. It often afflicts typists and servers in restaurants who carry heavy trays with the carrying hand held backwards. Treatments include immobilization of the wrist and hand for varying periods of time, injection and in severe cases, surgery. One of the commonly accepted indicators of the condition (or precondition) is the loss of vibrotactile sensation in the hand. Elaborate, expensive but too often faulty equipment is on the market so that physicians can make these measurements during an office visit. In a well-controlled experiment, measurements of vibrotactile thresholds in carpal tunnel patients showed that there were no significant differences between these patients and a control group of unafflicted subjects (Checkosky et al., 1996). The results also showed that the presence of pain related to carpal tunnel syndrome does not affect the sensitivity to vibration.

Today, men and women are exposed to a variety of environments in which the entire body is subject to severe vibrations in heavy-duty vehicles, earth-moving equipment and space ships, to name a few. What are the levels of vibration that these individuals are exposed to and what are the consequences and limits of such exposure? Investigations have been made in private and university laboratories, in government laboratories (OSHA), military and NASA.

5 • Environmental Well-Being

In addition to the many benefits brought about by the industrial revolution, it was accompanied by a multitude of negative consequences imposed upon humans, including pollution of the

air, water and sound environments. Not least amongst these was high-intensity vibration exposure to the whole or to limited parts of the body in the form of vibrating work places, vehicles and hand-held tools.

Over 1.5 million workers use vibrating tools in the United States alone. There has been an extensive amount of research into the occupational health problems associated with the use of vibrating tools including powered hammers, chain saws, sanders, grinders, drills, etc. The health problems commonly accompanying the use of vibrating tools show signs and symptoms of numbness, pain and blanching of fingers, thus the name "white finger" or primary Raynaud's disease. The symptoms are the result of a combination of peripheral vascular and neural disorders. The severity of the disease depends upon the vibratory energy produced by the tool, the length of time the tool is used each day, the cumulative period (months, years) the tool has been used, temperature and the ergonomics of tool design. Severe cases of the disease can result in ulcerations and the loss of fingers. The average incidence of the disease among populations that routinely use vibrating tools is about 50%. Those nations most actively engaged in research in this area are the United States, Canada, Japan, Sweden and Great Britain.

The vibration acceleration levels associated with the occurrence of these disturbances range from 10 to 2,014 (m/s^2 rms) and the incidence of symptoms is very closely related to magnitude of acceleration level. There is clear evidence that the development of hand–arm vibration syndrome is linked to the use of vibration-producing, hand-held tools. Coutu-Wakulczk and her co-workers (Coutu-Wakulczk et al., 1997) proposed to develop a self-report questionnaire to assess the ability of workers to perform simple hand/finger tasks, such as buttoning clothing or manipulating small objects. They correlated self-reports with measures of tactile acuity at the fingertips in power-tool operators and found a high degree of correspondence. The authors suggest that the questionnaire may be developed into a screening device for persons working in high-risk environments.

Shipyard workers use a variety of hand-held tools that expose them to high levels of vibration (chippers, grinders, etc.). Cherniack et al. (2004b) studied a group of 214 shipyard workers in 1988 to document the neurological symptoms they displayed and to assess the duration and levels of vibration to which they were exposed. The workers were put on a work routine that reduced their exposure time. In 2001, the workers were examined again to assess the effects of reduced exposure time.

The follow-up study revealed that the workers showed a reduction in the numbers of neurological symptoms; however, at lower levels of exposure, the number of symptoms increased. The authors conclude that reduction of exposure time alone is not sufficient to reduce neurological symptoms if not accompanied by a reduction of vibration amplitude.

In a straightforward study of the physiological consequence of exposure to vibration, Cherniack et al. (2004a) measured sequential nerve-conduction velocities from the wrist to the hands and digits. They found that in shipyard workers exposed to industrial vibrations, there was a reduction in conduction velocity in these nerves. This change has been associated with a number of neurological pathologies of the fingers and hands.

There have been many measurements of the vibration characteristics in the handles of power tools. Only two will be reported here. Keith and Brammer (1994) measured the vibration in the handles of pneumatic rock drills during normal use. They found that handle vibration increased as a function of percussion frequency with the largest amplitude peak at about 3 kHz. They go on to suggest that the frequency weighting recommended by the International Standards Organizaiton (ISO) overestimates the hazard by two times and that the National Institute of Occupational Safety and Health overestimates the hazard by a factor of 6. Gurram et al. (1995) have made an analysis of the driving-point mechanical impedance values of tool handles at frequencies from 10 to 100 Hz, derived from a large number of reports in the literature. From this data, they have constructed lumped-parameter models of the hand/arm system in all three orthogonal axes.

Treatment of the hand–arm vibration syndrome usually consists of water baths (most effective), drugs, vasodilators and calcium channel blockers. It should be noted that these treatments may be effective only in the early stages of the disorder where there is bleaching of one or more finger tips with no tingling or numbness. After tingling and numbness persist and blanching becomes extensive, there is no reversal of symptoms and treatment becomes palliative. Preventative measures include the improvement in the design of handles, protective gloves that contain heating elements and modifications of work habits that include adequate rest periods. Because of the extensive amount of research in this area, the reader is referred to summaries and conference proceedings such as Wasserman and Taylor (1977) and N10SH (1989).

6 • Whole Body Vibration

Vibration of the body is not singular in nature but rather a complex action without a single consequence: "Vibration may be a nuisance or nauseating, exhilarating or excruciating, a source of pleasure or the cause of pain" (Griffin, 1990, p. 1). Vibration of the entire body is experienced when the body is supported on a vibrating surface in one of three positions: seated, standing or lying. It is also possible that the surface may transmit the disturbance to a local region such as the head (head rest), hands (steering wheel) or feet (floor).

There are several ways that vibrations can affect the body: comfort, activities, health and motion sickness.

6.1 Comfort

The nature of vibration is one or a combination of a single sinusoid, a combination of frequencies or random noise. Griffin (1976) found that random vibrations are more uncomfortable than less complex sinusoidal frequencies. There are many methods that investigators use to assess the comfort index of the many types of motion to which the body may be subjected; these methods include multiple axis, translational, rotational, multiple input and others. One factor that can influence the degree of discomfort from vibration is posture (Messenger, 1987; Messenger and Griffin, 1989). In a seated position, maximum discomfort for frequencies between 10 and 40 Hz is felt in the upper part of the body (Whitman and Griffin, 1978). The discomfort can be reduced by assuming a relaxed position. In the standing position, discomfort can be reduced by bending the legs. Body size can also influence the discomfort as larger bodies were found to be less sensitive to low-frequency vibration.

6.2 Activity

Environmental vibration can have an adverse effect upon both sensory and motor activities of the individual. In vision, vibrations can interfere with the stability of the retinal image and thus provide the brain with misinformation. Prolonged exposure to intense vibrations or sudden intense exposure can cause changes in the supply of blood to the retina resulting in permanent damage to vision (Andreyuk, 1966; Jazov et al. 1975). Vibrations can also cause retinal detachment and the consequent serious visual impairment (Dupuis et al., 1975). (Of the many comprehensive reviews of the effects of vibration on vision, the reader is referred to Griffin and Lewis (1978) and Griffin et al. (1986)).

6.3 Control Activity

These effects are primarily in the domain of motor activities that are disturbed by vibration of the whole body and are subsumed under the rubric of "vibration-correlated errors". An excellent

review of these errors may be found in McLeod and Griffin (1989). The degree of errors caused by vibration depends in part on the complexity of the task, being most evident in simple manual tasks that require precise hand movements such as writing or pushing buttons. Vibrations can also inhibit reflex muscle activity (Martin et al., 1984) and produce illusions of movements (phantom movement) that can lead to performance errors (Goodwin et al., 1972). Also, Johnston (1979) showed that adverse vibrations will interfere with and distort speech production.

6.4 Health

Environmental vibrations have been found to have adverse effects upon health, some of which have been described above. In this section, the emphasis will not be on disease entities but rather on the effects of vibration to specific organ systems.

Guignard (1985) found that moderate to high vibration in the frequency range of 2–20 Hz can have effects on the cardiovascular system that are similar to those of moderate exercise. Changes occurred in heart rate, cardiac output, blood pressure and oxygen uptake. These and other changes occur in varying amounts depending on the frequency, intensity and duration of the vibration. Effects upon the activity of fetal hearts when the mother is exposed to vibration have also been reported (Gagnon et al., 1986a, b).

Transitory hyperventilation has been noted to result from short exposure to intense vibrations and that the effect increases with increased amplitude and duration of exposure (Duffner et al., 1962). Many studies have reported adverse effects of vibration on normal muscle function dating from 1860 (Rood, 1860) to the present. On the other hand, vibration can be beneficial when used therapeutically in assisting the motor control of spastic muscles (Hagbarth and Eklund, 1968), in Parkinsonism, spastic disorders and in cerebellar dysfunction (Young, 1968; Krestler, 1963).

The use of high-vibration equipment (power shovels, high cranes, dump trucks, etc.) can lead to illusory movements and balance disorders that constitute a serious danger to the operator (Kublanora, 1960; Lerashov et al., 1985). Riopellel and colleagues (Riopellel et al., 1958) also found that these operators showed a higher incidence of damage to the vestibular system. It should be noted that much of our information comes from non-laboratory field studies. In the laboratory, subjects can only be exposed to "tolerance" levels because of danger to the subjects. Also, the effects of laboratory stimuli typically are temporary.

Specific effects have been found in women who work in environments with 40–55 Hz background vibrations. Disorders

of menstruation, internal inflammation and abnormal child births were reported by Gratsianskaya et al. (1974). Low back pain is perhaps the most common complaint that brings a person into the physician's office. Exposure to vibration of high intensity and/or long duration can increase the incidence of premature degeneration of spinal vertebrae (Heide and Seidel, 1978; Sandover, 1983; Seidel et al., 1986). However, it should be noted that low back pain may also be traced to poor posture, heavy lifting, heredity and a host of other causes.

6.5 Motion Sickness

Many say and believe that motion sickness is "all in the mind". It has many different names depending on the mode of transport: car sickness, sea sickness, air sickness, etc. The stimuli may be different, but the results are always the same – nausea, vomiting and a wretched feeling. Although the susceptibility varies widely, a suitable combination of stimuli can produce motion sickness in any normal person.

The cause of motion sickness is a complex combination of sensory inputs, but the most common explanation is that it is due to a disparity of sensory inputs from the visual, vestibular and somatosensory systems. A common example occurs in the cabins of ships. The visual system tells us that the cabin is stable and non-moving. However, the vestibular system informs that the environment is moving and the somatosensory system is uncertain where gravitational "down" is. Studies have shown that although vibrations at 2–20 Hz may cause discomfort and even injury, they never lead to nausea or vomiting; two symptoms that can be caused by oscillations of less than 0.5 Hz (Johnson and Taylor, 1961; Guignard and McCauley, 1982). There were clear increases in motion sickness when the magnitude of the stimulation was increased (O'Hanlon and MaCauley, 1974; Lawther and Griffin, 1987). Although vibration cannot be considered the sole cause of motion sickness, experimental evidence suggests that its effects are certainly a contributing factor.

References

Andreyuk, V.I.: Some changes in the organ of vision secondary to vibration disease. Vest. Oftal'mologli, 79, 76–69, (1966).

Apkarian, A.V., Stea, R.A., Bolanowski, S.J.: Heat induced pain diminishes vibrotactile perception: A touch gate. Somatosens. Mot. Res., 11, 259–267, (1994).

Békésy, G. von: Neural volleys and the similarity between some interactions produced by tones and by skin vibrations. J. Acoust. Soc. Am., 29, 1059–1069, (1957).

Brown, A.E., Hopkins, H.K.: Interaction of the auditory and visual sensory modalities. J. Acoust. Soc. Am., 4, 1–6, (1967).

Buchner, D.N. (ed.): Vigilance–A Symposium. New York: McGraw Hill, 1963.

Checkosky, C.M., Bolanowski, S.J., Cohen, J.C.: Assessment of vibrotactile sensitivity in patients with carpal tunnel syndrome. J. Occup. Environ. Med., 38, 593–601, (1996).

Cherniack, M., Brammer, A.J., Lundstrom, R., Meyer, J., Morse, T.F., Nealy, G., Milsson, T., Peterson, D., Toppilla, E., Warren, N., Fu, R.W., Bruneau, H.: Segmental nerve conduction velocity in vibration-exposed shipyard workers. Int. Arch. Occup. Environ. Med., 77, 159–176, (2004a).

Cherniack, M., Morse, T.F., Brammer, A.J., Lundstrom, R., Meyer, J., Nilsson, T., Peterson, D., Toppilla, E., Warren, N., Fu, R.W., Brumeau, H., Croteau, M: Vibration exposure, workplace modification and disease in a shipyard: A 13 year revisit. Am. J. Indust. Med., 45, 500–512, (2004b).

Coutu-Wakulczk, G., Brammer, A.J., Piercy, J.E.: Association between a quantitative measure of tactile acuity and hand symptoms reported by operators of power tools. J. Hand Surg., 22A, 873–881, (1997).

Duffner, I.R., Hamilton, L.H., Schmitz, M.A.: Effect of whole-body vertical vibration or respiration in human subjects. J. Appl. Psychol., 17, 913–916, (1962).

Dupuis, H., Draegen, J., Hartung, E.: Vibration transmission to different parts of the body by various locomotions. in: Komi, P.V. (ed.) 537–543 Internat. Ser. Biomechan., Vol. 1A, Biomechan. V-A, Proc. 5th Internat. Conrg. Biomechan, (1975).

Ferrington, D.G., Nail, B.S., Rowe, M.: Human tactile direction thresholds: Modification by inputs from specific tactile receptor classes. J. Physiol., 272, 415–433, (1977).

Gagnon, R., Hunse, C., Carmichael, L., Fellows, F., Patrick, J.: Effects of vibratory acoustic stimulation on human fetal breathing and gross fetal body movements near term. Am. J. Obstet. Gynecol., 155, 1227–1230, (1986a).

Gagnon, R., Patrick, J., Foreman, J., West, R.: Stimulation of human fetuses with sound and vibration. Am. J. Obstet. Gynecol, 155, 849–851, (1986b).

Gescheider, G.A., Barton, W.G., Bruce, M.R., Goldberg, J.A., Greenspan, M.J.: Effects of simultaneous auditory stimulation on the detection of tactile stimuli. J. Exper. Psychol., 81, 120–125, (1969).

Gescheider, G.A., Niblette, R.K.: Cross-modality masking for touch and hearing. J. Exper. Psychol., 74, 313–320, (1967).

Goodwin, G.M., McCloskey, D.I., Matthews, P.B.C.: The contribution of muscle afferents to kinesthesia shown by vibration induced illusions of movement and by the effects of paralyzing joint afferents. Brain, 95, 705–708, (1972).

Gratsianskaya, L.N., Eroshenko, E.A., Libertorvch, A.P.: Influence of high-frequency vibration on the genital region of females. Gigiena Tnuda: Profess. Zabol., 19, 7–10, (1974).

Green, B.G.: The effect of skin temperature on vibrotactile sensitivity. Percept. Psychophys., 21, 243–248, (1977).
Green, B.G.: The effect of cooling on the vibrotactile sensitivity of the tongue. Percept. Psychophys., 42, 423–430, (1987).
Green, B.G., Lederman, S.J., Stevens, J.C.: the effect of skin temperature on the perception of roughness. Sens. Processes, 3, 327–333, (1979).
Griffin, M.J.: Subjective equivalence of sinusoidal and raudom whole-body vibration. J. Acoust. Soc. Am, 60, 1140–1145, (1976).
Griffin, M.J.: Handbook of Human Vibration, London: Academic Press, (1990).
Griffin, M.J., Lewis, C.H.: A review of effects of vibration on visual acuity and continuous manual control. Part I: visual acuity. J. Sound Vib., 56, 383–413, (1978).
Griffin, M.J., McLeod, R.W., Moseley, M.J., Lewis, C.H.: Whole-body vibrations and aircrew performance. Tech-Rpt. No. 132. Instit. Sound Vib. Res., U. Southampton, (1986).
Guignard, J.C.: Vibration in Cralley and Cnalley (eds.) 653–724, Chapt. 15, Patty's Industrial Hygiene and Toxicology, Vol. 3b: Biological Responses, 2nd Ed. New York: Wiley, 1985.
Guignard, J.C., McCauley, M.E.: Motion sickness incidence induced by complex periodic waveforms. Aviat. Space Environ. Med., 53, 554–563, (1982).
Gurram, R., Rakheja, R., Brammer, A.J.: Driving-point mechanical impedance of the human hand-arm system: synthesis and model development. J. Sound Vib., 180:437–458, (1995).
Hagbarth, K.E., Eklund, G.: The effects of muscle vibration in spasticity, rigidity, and cerebellar disorders. J. Neurol. Neurosurg. Psychiat., 31, 207–213, (1968).
Heide, R., Seidel, H.: Consequence of long-term occupational exposure to whole-body vibration (abridged literature survey). Zeitsch. Hygience Grenz., 24, 153–159, (1978).
Jazor, K., Baicev, H., Panica, D.: State of visual analyzer by vibration diseases. Folia Medica, 17, 25–27, (1975).
Johnson, W.H., Taylor, N.B.G.: Some experiments on the relative effectiveness of various types of acceleration on motion sickness. Aerospace Med., 32, 205–208, (1961).
Johnston, M.E.: A preliminary investigation of the effect of 2–20 Hz whole body vibration on spoken numbers. Tech. Memo. FS152, Royal Aircraft Estab., Farnaborough, (1979).
Keith, S.E., Brammer, A.J.: Rock-drill handle vibration: Assessing the influence of operating parameters and high frequencies. J. Sound Vib., 174:475–491, (1994).
Krestler, O.C.: Musculoskeletal disorders in orthopedic and geriatric patients. Gen. Pract., 26, 6, (1963).
Kublanora, P.S.: The effect of vibration on the state of hearing and the vestibular apparatus. Vestnik Otolaryngol. (Moscow), 22, 15–21, (1960).

Lawther, A., Griffin, M.J.: Prediction of the incidence of motion sickness from the magnitude, frequency and duration of vertical oscillation. J. Acoust. Soc. Am., 82, 957–966, (1987).

Lerashov, M.M., Nalimova, T.A., Ermakova, L.A.: Latent vestibular disfunction caused by vibration. Proc. Sec. Internat. CISM-IFT. MM Sympos., Moscow, 2–12 Apr., 177–180, (1985).

Lindblom, U., Verrillo, R.T.: Sensory functions in chronic neuralgia. J. Neurol. Neurosurg. Psychiat., 42, 422–435, (1979).

Loewentstein, W.R.: Excitation and inactivation in a receptor membrane. Ann. N.Y. Acad. Sci., 94, 510–534, (1961).

Loveless, N.E., Brehner, J., Hamilton, P.: Bisensory presentation of information. Psychol. Bull., 73, 161–199, (1970).

Martin, B.J., Roll, J.P., Gauthier, G.M.: Spinal reflex alterations as a function of intensity and frequency of vibration applied to the feet of seated subjects. Aviat. Space Environ. Med., 55, 8–12, (1984).

McLeod, R.W., Griffin, M.J.: A review of the effects of whole-body translation vibration on continuous manual control performance. J. Sound Vib., 133, 55–115, (1989).

Messenger, A.J.: Effect of pelvic angle on transmission of vertical vibration to the head of seated subjects. U.K. Inform. Group Mtng. Hum. Resp. Vib., Royal Military Coll. Sci, Shrivenham, 21–22 Sept., (1987).

Messenger, A.J., Griffin, M.J.: Effects of anthropometric and postural variables on the transmission of whole-body vertical vibration from seat to head. Tech. Report No. 172, Instit. Sound Vib. Res., U. Southampton, (1989).

N1OSH: Criteria for a Recommended Standard Occupational Exposure to Hand-Arm Vibration, U.S. Department of Health and Human Services, National Institute for Occupational Safety and Health, Cincinnati, Ohio, (1989).

O'Hanlon, J.F., MaCauley, M.E.: Motion sickness incident as a function of the frequency and acceleration of vertical sinusoidal motion. Aerosp. Med., 45, 366–369, (1974).

Osborn, W.C., Sheldon, R.W., Baker, R.A.: Vigilance performance under conditions of redundant and nonredundant signal presentation. J. Appl. Psychol. 4, 130–134, (1963).

Riopelle, A.J., Hines, M., Lawrence, M.: The effects of intense vibration Report No.,358, AD 203637, U.S. Army Med. Res. Lab., Fort Knox, KT, (1958).

Rood, N.O.: On contraction of the muscle, induced by contact with bodies in vibration. Am. J. Sci. Arts, 24, 449, (1860).

Sandover, J.: Dynamic loading as a possible source of low-back disorders. Spine, 8, 652–658, (1983).

Sato, M.: Response of Pacinian corpuscles to sinusoidal vibrations. J. Physiol., 159, 391–409, (1961).

Seidel, H., Bluethner, R., Hinz, B.: Effects of sinusoidal whole-body vibration on the lumbar spine: the stress-stain relationship. Int. Arch. Occup. Environ. Helath, 57, 207–223, (1986).

Stevens, J.C.: Temperature can sharpen tactile acuity. Percept. Psychophys., 31, 577–580, (1982).
Stevens, J.C.: Temperature and the two-point threshold. Somatosens. Mot. Res., 6, 275–284, (1989).
Stevens, J.C., Hooper, J.E.: How skin and object temperature influence touch sensation. Percept. Psychophys., 32, 282–285, (1982).
Verrillo, R.T., Bolanowski, S.J.: The effects of skin temperature on the psychophysical responses to vibration in glabrous and hairy skin. J. Acoust. Soc. Am., 80, 528–532, (1986).
Verrillo, R.T., Bolanowski, S.J.: Effects of temperature on the subjective magnitude of vibration. Somatosens. Mot. Res., 20, 133–137, (2003).
Verrillo, R.T., Bolanowski, S.J., Baran, F., Smith, P.F.: Effects of underwater environmental conditions on virbrotactile systems. J. Acoust. Soc. Am., 100, 651–658, (1996).
Verrillo, R.T., Bolanowski, S.J., Checkosky, C.M., McGlone, F.P.: Effects of hydration on tactile sensation. Somatosens. Mot. Res., 15, 93–108, (1998).
Verrillo, R.T., Bolanowski, S.J., McGlone, F.P.: Subjective magnitude of tactile roughness. Somatosens. Mot. Res., 16, 352–360, (1999).
Verrillo, R.T, Capraro, A.J.: Effect of simultaneous auditory stimulation on vibrotactile thresholds. Percept. Psychophys., 16, 597–600, (1974).
Verrillo, R.T., Ecker, A.D.: Effects of root or nerve destruction on vibrotactile sensitivity in trigemireal neuralgia. Pain, 3, 239–255, (1977).
Verrillo, R.T., Gescheider, G.a., Calman, B.G., Van Doren, C.L.: Vibrotactile masking: Effects of one-and two-site stimulation. Percept. Psychophys., 33, 379–387, (1983).
Wasserman, D.E., Taylor, W. (eds.): Proceedings of the International Occupational Hand-Arm Vibration Conference. U.S. Department of Health, Education, and Welfare, National Institute for Occupational Safety and Health, Cincinnati, Ohio, 1977.
Weitz, J.: Vibratory sensitivity as a function of skin temperature. J. Exp. Psychol., 28, 21–36, (1941).
Whitman, E.M., Griffin, M.J.: The effects of vibration frequency an direction on the location of areas of discomfort caused by whole-body vibration. App. Ergon., 9, 231–239, (1978).
Young, D.G.: The effects of mechanical cycloid vibration on EMG potential during maximal muscle tension. Report, Niagara Therapy Ltd., (1968).

64

Sensory Substitution

Ronald T. Verrillo

Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA

1 • Introduction

The idea that the cutaneous surface may be employed as a substitute for the eyes and ears is by no means a modern notion. Although the sense of touch has long been considered as a surrogate for both the visual and auditory modalities, the focus of this chapter will be on the efforts to develop a tactile substitute for hearing, especially that of human speech. The visual system is our primary means of processing information about environmental space such as orientation, distance, direction and size. It is much less effective in making temporal discriminations. The auditory system is unparalleled in processing information that involves rapid sequences of temporal events, such as speech and music. The tactile sense is capable of processing both spatial and temporal information although not as effective in either domain as the eye or the ear.

Of course, the alleviation of hearing impairment due to deafness is not the only venue in which tactile aids may be used. The military has a vested interest in tactile communication. Such devices are important where communication via radio signals would endanger the receiver (navy divers on wartime missions) or where information must be provided to supplement overtaxed visual and auditory systems that utilize tactile feedback as a means of informing the operator of the attitude of the vehicle under his or her control.

Although they are important, our concern here will not be on operating or alerting signals or on aspects of localization but rather on the transmission of speech. Human speech is a most complex, sophisticated and subtle form of communication, without parallel in any other species of animal. A severe handicap is imposed upon those whose ability to use it is limited. Speech acquisition, normally acquired in the first several years of life, is predicated upon a feedback system between the sensing ear and the motor apparatus designed for producing speech sounds. Although hand signing alleviates the problem of communication among deaf persons, there still exists a considerable gap in oral communication between the world of the deaf and the hearing. The problem for the prelingually deaf is especially perplexing because they never have had the experience of hearing speech sounds and so are at a disadvantage in interpreting the meaning of tactile signals. The use of the skin as a substitute for the auditory system has a relatively long history (Gault, 1924; Gault and Crane, 1928; Geldard, 1968b). In spite of the efforts of cutaneous scientists and engineers, the advances in the development of transducers specifically designed for the communication of speech sounds by way of the skin has been meager (Sherrick, 1984). Many names have been applied to these devices by manufacturers, but they all fall into the generic rubric of "tactile vocoder".

A substantial but limited effort has been focused on the stimulation of skin by means of electrical impulses delivered through surface electrodes attached to the skin (see Brown and Stevens, 1992; Gilmer, 1961; Hawkes, 1961; Foulke, 1964; Saunders, 1974; Saunders et al., 1980; Sparks 1979; Tillmann and Piroth, 1986). These devices have the dual advantage of more compact size and much lower power requirements than electromechanical units. Disadvantages include the need of a gel to ensure an efficient transfer of current from the electrode to the skin, and there is the possibility (as often happens) of developing a rash under the electrode. Stimulation at a single site or at multiple sites on the skin has been employed by both mechanical and electrical devices.

One of the first decisions that must be made in the design of a tactile vocoder is where on the surface of the body the device will be placed. Some investigators have chosen the fingers because of their high sensitivity and exquisite tactile acuity. Others have designed units to be held in the whole hand. In spite of the advantages, the hands are in almost constant use for other tasks and hand-held devices soon produce fatigue and perspiration unless released periodically to "breathe". The wrist is a preferred site

for some designers because people are accustomed to wearing wristwatches and jewelry on that site. Other sites that have been tested are the forearm, the thigh and the abdomen. Of these sites, the forearm offers several advantages. Although less sensitive than the fingers and palm, it is more sensitive than the thigh or abdomen and it is of a sufficient size to permit the use of multiple stimuli and stimulus localization as a part of the translation code. It can also be covered (hidden) conveniently by clothing if concealment is desired. In the past, placement of tactile aids has been on the fingers (Blamey et al., 1988), the wrist (Dodgson et al., 1983), the forearm (Tillmann and Piroth, 1986), the abdomen (Sparks, 1979; Saunders, 1974; Saunders, 1980) and even the ear canal (Bochenek et al., 1989).

In order to design an effective tactile aid, it is essential to know the capabilities and limitations of the cutaneous tactile systems in respect of four stimulus characteristics: location, intensity, duration and frequency.

1.1 Location

Geldard (1966) developed a system for decoding letters of the alphabet by means of a pattern of excitation of nine vibrators widely distributed over the surface of the body (the Optohapt). Intended only as a laboratory instrument to study the ability to make useful pattern discriminations, subjects achieved a modicum of success.

The Optacon (Bliss, 1969; Linville and Bliss, 1966) is a device that enables the blind to read the printed page without the intervening step of translation into Braille. It utilizes a densely packed array of piezoceramic reeds that directly mimic the configuration of the printed letters through a vibrating pattern imposed on the fingertip. The speed at which text can be read depends on the age at which training is begun. Most studies report that on average the speed is comparable to Braille reading in adult subjects with equal training in both methods.

For the most part, subjects are extremely accurate in pinpointing the locus of a stimulus impressed on some part of the body's surface.

1.2 Intensity

Geldard (1965) tested subjects using the Optohapt system and found that most errors were made on the intensive continuum. Subjects could make only three absolute judgments of intensity between the absolute threshold of detectability and the onset of pain. Therefore, the usefulness of intensity as a cue in any tactile code appears to be extremely limited.

1.3 Duration

Estimates of the duration of a tactile stimulation are poor. Only four or five absolute discriminations can be made between 0.1 and 2.0 s (Geldard, 1965)

1.4 Frequency

At frequencies below 50 Hz, subjects were fairly accurate in making judgments of vibration frequency (Goff, 1967). However, at higher frequencies, estimates of frequency became very poor. In addition, Békésy (1959) showed that perceived pitch on the skin is dependent on the intensity of the vibration, which adds a complicating factor. The ability to distinguish events closely separated in time is related to the discrimination of frequency. The ear is able to distinguish two auditory clicks separated by as little as 1.8 ms, whereas on the fingertip the stimuli must be at least 10 ms apart (Gescheider, 1966; 1967). If the tactile aid is to use brief successive stimuli, it is important that the intensity be well above threshold because the tactile resolution of clicks is very dependent upon the intensity of the stimulus, much more so than stimuli delivered to the auditory system.

2 • Temporal Gap Detection

Temporal acuity in hearing is often assessed by measuring the ability of subjects to detect a silent period or gap in bursts of sinusoids or bursts of band-limited noise (temporal gap detection) (Plomp, 1964; Buunen and van Valkenberg, 1979; Fitzgibbons, 1983, 1984). Gap detection in hearing correlates with some results of speech tests (Tyler et al., 1982; Hochmair-Desoyer et al., 1985). Because the understanding of speech is the primary goal of tactile aid development, a number of investigators have studied gap detection in the touch modality. Gescheider (1966, 1967) determined that gap detection improves as a function of the time interval and the intensity of clicks, thresholds being about 10 ms.

Van Doren et al. (1990) tested subjects from 8 to 75 years of age using bursts of sinusoids and band-limited noise and found that as gap durations increased they became easier to detect and that gaps in sinusoids were more easily detected than were gaps in noise. Gescheider et al. (2003) confirmed this finding and reported that gap detection is more dependent on temporal cues for sinusoids and on amplitude changes in bursts of noise. They found no differences in temporal acuity between the Pacinian and non-Pacinian receptor channels (see Chapt. 62).

Gap-detection thresholds are substantially higher in touch than those of the auditory system, which implies that the temporally modulated information in speech may be lost in tactile communication devices.

3 • User Variables

It is incumbent upon the designer of tactile aids to consider the characteristics of the intended user. Foremost among these are age and sex.

3.1 Age

This variable alone could easily occupy a whole chapter, but here it will be dealt with very briefly. As we grow older, the senses become progressively less able to detect and make discriminations of every stimulus characteristic that has been measured in the laboratory and in field studies. If a tactile aid is to be useful to the widest range of potential users, it is essential that the designer be acquainted with what is known about the effects of aging. This is especially true because a greater number of those who will need assistance grows with each advancing year. The reader is directed to several summary books and chapters for further information about aging effects (Charness, 1985; Verrillo and Verrillo, 1985).

3.2 Sex

There are no significant differences between men and women in their absolute detection thresholds of vibration (Verrillo, 1979) although at suprathreshold levels women judge the same levels of amplitude to be more intense than do men (Verrillo, 1979). Also, the absolute thresholds of women vary at different phases of the menstrual cycle (Gescheider et al., 1984). The size of the tactile display must also be considered because on average, women tend to be smaller than men, the mean area of the skin on the female body being approximately 1.65 m^2 whereas that of males is about 1.85 m^2.

4 • Practical Design Considerations

A number of constraints are imposed upon the designer of an effective, practical tactile device. In order to be effective, a wearable tactile aid must satisfy certain conditions:

1. The transducer, whether mechanical or electrical, must be efficient. Because it is portable and therefore battery powered, the power source must be of an appropriate size and longevity.
2. The transducer must be able to reproduce accurately the necessary waveforms without undue distortion.

3. The transducer must be able to track intensity changes without distortion.
4. It must be economically priced, which means that production strategies have to be considered in the design and development phases.
5. To be worn comfortably, the device must be small, light, without undue pressure on the skin and it must be easily concealed. If it cannot be hidden completely, it must be cosmetically acceptable.

There are a substantial number of devices and materials developed for aiding deaf persons to understand the spoken word. Some have been designed specifically for use in the laboratory and others are available commercially for consumer use. To review these in detail would be beyond the scope of this chapter; however, a number of them are given below with broad details. The reader is referred to several reviews and discussions that may be helpful (Cholewiak and Wollowitz, 1992; Sherrick, 1984, 1991).

5 • Mechanical Stimulators

In this category are the activators that utilize armature coils and magnets to move a piston upon which is attached a contactor that imposes displacements to the skin. These include vibrators (shakers) and solenoids. Our primary concerns are those systems specifically designed for use in tactile hearing-aid devices, easily portable and reasonably priced.

6 • Electrical Stimulators

The use of electrical stimulation in a controlled and systematic manner dates from efforts beginning in the 1930's to teach deaf-mute children to understand speech. High-frequency (>1500 Hz) and low-frequency (<1500 Hz) electrical pulses were delivered to two fingers. His subjects were able to learn to distinguish the different vowels. Interest in electrical devices re-emerged in the 1960s (Gilmer, 1961; Hawkes, 1961; Foulke, 1964). They reported that subjects could discriminate several intensities. Reviews of the literature on electrotactile stimulation are given in Geddes (1972), Pethig (1979) and Donaldson and Donaldson (1986). McAdams (1987) offers a comprehensive, in-depth study of electrotactile stimulators.

There are several drawbacks to stimulating the skin with electrical impulses. Direct electrical stimulation of the skin can excite the end-organ receptors, terminal nerve fibers, neural

pathways and large nerve trunks depending on the size of the electrode and the amount of current applied. These factors will also determine the size and locus of the resulting sensations. If not carefully controlled, it can lead to confusion as to just where the sensation is coming from. In addition, if nerve trunks are stimulated, efferent (motor) fibers may be stimulated as well as sensory fibers which can cause muscle twitch and contractions.

The power requirements of electrotactile devices are low, about 10 mA. A wearable aid can run for many days on a single miniature battery.

7 • Coding

One of the first decisions that the designer of a tactile aid must make is what coding strategy to use. One such strategy utilized features of the speech signal, such as the voice fundamental frequency (Boothroyd and Hnath-Chisholm, 1988). "Pitch" is displayed in single- or multi-channel devices. Such a code provides the user with information not available in other vocoder models. Direct electrical stimulation of the skin can excite the end-organ receptors, terminal nerve fibers, neural pathways and large nerve trunks depending on the size of the electrode and the amount of current applied. These factors will also determine the size and locus of the resulting sensations. If not carefully controlled, it can lead to confusion as to just where the sensation is coming from.

An excellent detailed discussion of a number of signal processing strategies is in Summers (1992b), who considers the information in the speech signal that is most useful and the capacity of the vibrotactile system to decode the signals imposed upon it. Of special interest are representations of voice fundamental frequencies for two sentences spoken by male speakers. Shown in Figures 1 and 2, these illustrate how the voice fundamental frequency can be transformed into patterns that facilitate detection in a meaningful way by tactile receptors.

Another code emphasizes the second formant of speech sounds. Experiments performed using this coding strategy reported substantial benefits in single-item and connected speech after training (Cowan et al. 1988, 1990; Brenner and Weissenberger, 1990; Osberger et al., 1991).

Perhaps the most promising coding strategy is principle component analysis. A small set of independent parameters is obtained by doing correlations of the spectral energies in a waveform, which has the effect of reducing the redundancy of speech sounds. This process yields information about the entire

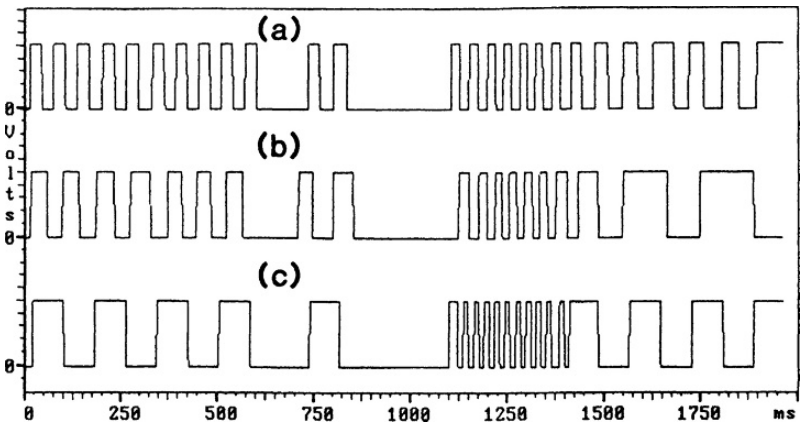


FIGURE 1 Voice fundamental frequencies for the sentence "My horse has three white legs" by a male speaker. (**a**) Square wave at FO; (**b**) after a linear transformation of the frequency range 100–200 Hz into the range 40–220 Hz; (**c**) after an extreme non-linear transformation to a "low–high" (50/300 Hz) representation. In each case, the stress on the word "three" is indicated by a rise in frequency. For greater clarity of reproduction, square-wave frequencies are one-eighth of the true frequencies quoted above. (From Summers, 1992b.)

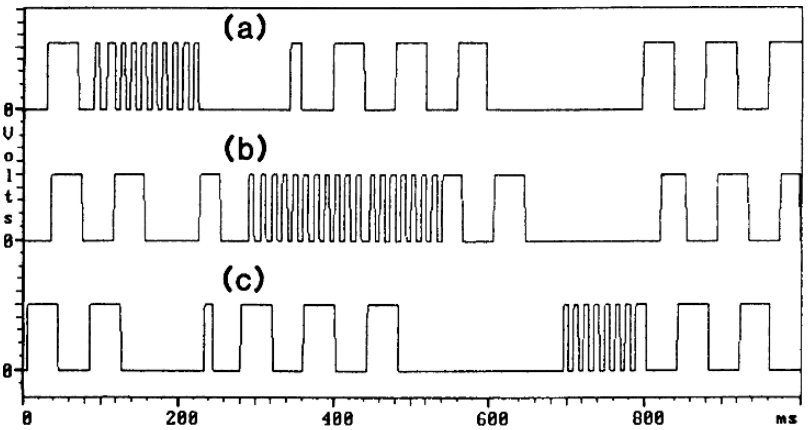


FIGURE 2 Variations of FO for differently stressed versions of the sentence "Look on the chair", presented as square-wave frequency. Only gross FO changes are preserved after reduction to a "low–high" reduction. Tactual perception of FO changes may be facilitated by the transformation. The stress in (**a**), (**b**) and (**c**) falls on the first, second and fourth words; the stressed word is indicated by a rise in frequency. The "low" and "high" outputs are shown as square-wave frequencies of 12.5 and 75 Hz rather than 40/400 Hz, which would be better for tactile presentation. (From Summers, 1992b.)

shape of the waveform rather than just the peaks in its spectrum. Using this coding strategy with a multi-channel tactile aid, Weissenberger et al. (1991) reported impressive performances in the identification of vowels, consonant features as well as improvements in lip reading.

8 • The Systems

It is beyond the scope of this chapter to review the many materials, systems and devices that have been developed and tested for tactile speech appreciation. The driving mechanisms are electromechanical with wire-wound armatures, magnets and moving elements (vibrators, shakers); solenoid configurations with hold and release mechanisms and finally devices that utilize piezoelectric materials in bimorph form that bend in response to the passage of electrical currents or thin-film sheets that bend or expand and contract when current is applied. Listed here are a number of references containing detailed information about the design, construction and performance of many of these systems.

In alphabetical order, these are as follows: Brooks et al., 1986a, b; Cholewiak and Wollowitz, 1992; Cowen et al., 1988; Leysieffer, 1986; Mason and Frost, 1992; Osberger et al., 1991; Sherrick, 1984, 1991; Spens and Plant, 1983; Summers, 1992a; Van Doren, 1987, 1989.

9 • Training

The successful use of any of these devices (as with any unfamiliar instrument) is predicated upon the adequate training of the user. Obviously, the person who became deaf after having heard speech and had spoken previously to becoming deaf is going to be a more rapid learner than the person who became deaf prelingually. Hundreds of studies have been performed to evaluate the adequacy of these devices and all report a certain degree of success or lack thereof. For an excellent discussion of the general problem of tactile communication, see Weisenberger (1992), and for a discussion of the training of users, see Plant (1992). Most studies show that with training, the greatest degree of success in the use of tactile vocoders is achieved when the device is used in conjunction with another sensory input, such as lip reading, a cochlear implant or a conventional hearing aid.

The selection of an aid to be used for training depends upon what features of speech the therapist wishes to concentrate on, the size, location and cosmetic acceptability of the aid and the characteristics of the trainee. Electrical stimulation is often

rejected for children and an aid deemed acceptable for a congenitally deaf child might be unacceptable for training a postlingually deaf adult. One option may not be suitable for all potential users.

There are two opinions concerning the selection of persons to receive tactile aids. A restrictive opinion (official policy in some places) limits the use of tactile aids only to persons who cannot benefit from conventional hearing aids. A more liberal policy places no restrictions and considers tactile aids to be a supplement to, rather than a substitution for conventional devices. In particular, this applies to the training of children learning speech patterns and learning to be alert to acoustic events in the environment. Prior to the decision that a tactile aid would be advisable, it is essential that the potential user be given a thorough audiological workup in addition to having determined threshold and suprathreshold measurements of vibrotactile functions.

It is difficult to determine from the literature the specifics of a single training program in any great detail because the devices that were used differed from report to report and the characteristics of the users (age, sex, age of onset of deafness, audiometric profiles, etc.) differed widely. However, some general principles do emerge, which include the following:

1. Detection of sound.
2. Discrimination of tactile patterns of speech and non-speech sounds.
3. Discrimination of the syllabic structure of words.
4. Detection of high frequencies.
5. Training in the improvement of speech production.

1. It is important that the user be made aware of the presence of a wide variety of sounds in the environment. This is true especially because a sense of relatedness to the world is in large part a function of the ability to hear that world (Ramsdell, 1978). The awareness of background events obtained by hearing is essential to a sense of security and well-being.
2. In order to develop an appreciation for speech sounds, the user must learn to discriminate between speech and non-speech sounds. Within speech the difference between short and long syllables, fast and slow speech and different intensity levels, etc., must be learned.
3. Syllabic structure such as monosyllabic, trochaic, spondaic and trisyllabic differences are important in understanding speech. These are taught slowly, building from single to multiple word combinations.
4. The sound of consonants is determined by high-frequency components of the sound, as in the voicing of s, t, z, etc. The

user must be trained to detect these features and to discriminate between words with and without these components.

5. The last phase of training is practice and improvement of speech production. Although this improvement is a part of every phase of training, there must be a more concentrated effort to monitor speech-sound production in the final period of training.

The exclusive use of tactile aids is not advocated by anyone; they are designed and produced in order to train in the development of speech skills and to be used in conjunction with conventional hearing aids, lip reading and in the aftermath of cochlear implantation. It is doubtful that the ideal tactile aid that is not a supplement to other devices and procedures will ever be realized. But, to abandon the effort would be as senseless as the attempt to close the patent office at the turn of the 20th century.

Several words of caution are in order concerning the investigators and designers of tactile aids. It is too often the case that persons working in this field are engineers who know little or nothing about the anatomy and physiology of cutaneous receptors or the perceptual responses of humans to tactile stimulation. On the other hand, the investigators are well versed in these aspects of vibrotaction but are not knowledgeable about the engineering constraints. Very often, neither of the foregoing groups are expert in the sciences of speech and hearing. The result has been that a great deal of time, effort and funds have been expended that led to dead ends in terms of effective, practical tactile aids to human oral communication. The most practical approach to the problem should ideally be made by a team of scientists comprised of engineers, experts in speech and hearing and scientists who are knowledgeable in the field of human responses to cutaneous stimulation.

References

Békésy, G. von: Similarities between hearing and skin sensations. Psyhcol. Rev., 66, 1–22, (1959).

Blamey, P.J., Cowan, R.S.C., Alcantara, J.I., Clark, G.M.: Phonemic information transmitted by a multichannel electrotactile speech processor. J. Speech Hear. Res, 31, 620–629, (1988).

Bliss, J.C.: A relatively high resolution aid for the blind. IEEE Transactions. Man-Machine Systems, MMS-10, 1–18, (1969).

Bochenek, W., Chrozempa, A., Hazell, J.W.P., Kieiak, J., Kukwa, A.: Noninvasive electrical stimulation of the ear canal as a communication aid in aquired total deafness. Brit. J. Audiol., 23, 285–291, (1989).

Boothroyd, A., Hnath-Chisolm, T.: Spatial, tactile presentation of voice fundamental frequency as a supplement to lipreading: results of extended training with a single subject. J. Rehabil. Res. Dev. 25, 51–56, (1988).

Brenner, C., Weissenberger, J.M.: Evaluation of a wearable multichannel tactile aid. J. Acoust. Soc. Am., 88, s192(A), (1990).
Brooks, P.I., Frost, B.J., Mason, J.I., Gibson, D.M.: Continuing evaluation of the Queen's University tactile vocoder. J. Rehabil. Res. Dev., 23, 119–128, (1986a).
Brooks, P.I., Frost, B.J., Mason, J.L., Gibson, D.M.: Continuing evaluation of the Queen's University Tactile Vocoder. II Identification of open-set sentences and tracking narrative. J. Rehabil. Res. Dev., 23, 129–138, (1986b).
Brown, B.H., Stevens, J.C.: Electrical stimulation of the skin. in: Summers, I.R. (ed.) 376–561. Tactile Aids for the Hearing Impaired, London: Whurr, (1992).
Buunen, T.J.F., vanValkenberg, D.A.: Auditory detection of a single gap in noise. J. Acoust. Soc. Am., 65, 534–537, (1979).
Charness, N. (Ed.) Aging and Human Performance, John Wiley, Chichester, (1985).
Cholewiak, R.W., Wollowitz, M.: The design of vibrotactile transducers. in: Summers, I.R. (ed.) 57–82 Tactile Aids for the Hearing Impaired. Whurr, London, (1992).
Cowan, R.S.C., Alcantara, J.I., Blamey, P.J., Clark, G.M.: Preliminary evaluation of a multichannel electrotactile speech processor. J. Aoust. Soc. Am. 83, 2328–2338, (1988).
Cowan, R.S.C., Blamey, P.J., Galvin, K.L., Sarant, J.Z., Alcantara, J.I., Clark, G.M.: Perception of sentences, words, and speech features by profoundly hearing-impaired children using a multichannel speech processor. J. Acoutst. Soc. Am. 88, 1374–1384, (1990).
Dodgson, G.S., Brown, B.H., Freeston, I.L., Stevens, J.C.: Electrical stimulation at the wrist as an aid for the profoundly deaf. Clin. Phys. Physiol. Meas. 4, 403–416, (1983).
Donaldson, N., Donaldson, P.E.K.: When are actively balanced biphasic ("Lilly") stimulating pulses necessary in a neurological prosthesis? I. Historical background; pt resting potential; Q studies. Med. Biol. Eng. Comput., 24, 41–49, (1986).
Fitzgibbons, P.J.: Temporal gap-detection in noise as a function of frequency, bandwidth, and level. J. Acoust. Soc. Am., 74, 67–72, (1983).
Fitzgibbons, P.J.: Temporal gap resolution in masked normal ears as a function of masker level. J. Acoust. Soc. Am., 76, 67–70, (1984).
Foulke, E.: The locus dimension as a basis for electrocutaneous communication. J. Psychol., 57, 253–257, (1964).
Gault, R.H.: Progress in experiments on tactile inerpretation of oral speech. J. Abnorm. Soc. Psychol., 19, 155–159, (1924).
Gault, R.H., Crane, G.W.: Tactual patterns from certain vowel qualities instrumentally communicated from a speaker to a speaker's fingers. J. Gen. Psychol. 1, 353–359, (1928).
Geddes, L.A.: Electrodes and the Measurement of Bioelectric Events. Wiley, New York, (1972).
Geldard, F.A.: Cutaneous coding of optical signals: The optohapt. Percept. Psychophys. 1, 377–381, (1966).
Geldard, F.A.: The perception of mechanical vibration I–IV. J. Gen. Psychol., 22, 243–308, 1968a.

Geldard, F.A.: Pattern Perception by the skin senses. in: The Skin Senses, Kenshalo, D.R. (ed.) 304–321 Springfield, Ill., Thomas, C.C., (1968b).
Geldard, F.A., Sherrick, C.E.: Multiple cutaneous stimulation: The discrimination of vibratory patterns. Acoust. Soc. Am., 37, 797–801, (1965).
Gescheider, G.A.: Resolving of successive clicks by ears and the skin. J. Exp. Psychol., 71, 378–381, (1966).
Gescheider, G.A.: Auditory and cutaneous resolution of successive brief stimuli. J. Exp. Psychol., 75, 570–572, (1967).
Gescheider, G.A., Bolanowski, S.J., Chatterton, S.K.:Temporal gap detection in tactile channels. Somatosens. Mot. Res., 20, 239–247, (2003).
Gescheider, G.A., Verrillo, R.T., McCann, J.T., Aldrich, E.M.: Effects of the menstrual cycle on vibrotactile sensitivity. Percept Pshchophys., 36, 586–592, (1984).
Gilmer, B.H.: Toward cutaneous electro-pulse communication. J. Psychol., 52, 211–222, (1961).
Goff, G.D.: Differential discrimination of frequency of cutaneous mechanical vibration. J. Exper. Psychol., 74, 294–299, (1967).
Hawkes, G.R.: Cutaneous discrimination of electrical intensity. Am. J. Psychol., 74, 45–53, (1961).
Hochmair-Desoyer, I.J., Hochmair, E.S., Stiglebrunner, H.K.: Psychoacoustic temporal processing and speech understanding in cochlear implant patients. in: Schindler, R.A., Merzenich, M.M. (eds.) 291–304 Cochlear Implants. Raven, New York, (1985).
Leysieffer, H.: A wearable multi-channel auditory prosthesis with vibrotactile skin stimulation. Audiol. Acoust., 25, 230–251, (1986).
Linville, J.G., Bliss, J.C.: A direct translation reading aid for the blind. IEEE Proc., 54, 40–51, (1966).
Mason, J.L., Frost, B.J.: Signal processing strategies for multichannel systems. in: Summers, I.R. (ed.) 128–145 Tactile Aids for the Hearing Impaired. Whurr, London, (1992).
McAdams, E.T.: A study of electrode-tissue impedances encountered in cardiac pacing. PhD Thesis, University of Leeds, (1987).
Osberger, M.J., Robbins, A.M., Todd, S.I., Brown, C.J.: Initial findings with a wearable multichannel vibrotactile aid. Am. J. Otol. 12(Suppl. 1.), 179–182, (1991).
Pethig, R.: Dielectric and Electronic Properties of Biological Materials. Chichester, John Wiley, (1979).
Plant, G.: The selection and training of tactile aid users. in: Summers, I.R. (ed.) 146–166 Tactile Aids for the Hearing Impaired. Whurr, London, (1992).
Plomp, R.: Rate of decay of auditory sensation. J. Acoust. Soc. Am., 65, 534–537, (1964).
Ramsdell, S.A. The Psychology of the hard of hearing and the deafened adult. in: Davis, H., Silverman, S.R. (eds.) 435–446, Hearing and Deafness. Holt, Rinehart and Winston, New York, (1978).
Saunders, F.A.: Electrocutaneous displays. in Geldard, F.A. (ed.) 20–26 Cutaneous Communication Systems and Devices. The Psychonomic Society, Austin, TX, (1974).

Saunders, F., Hill, W., Simpson, C.: Speech perception via the tactile mode: Progress report. IEEE Conference on Acoustics, Speech and Signal Processing (Reprinted in Levitt, H., Pickett, S.M., Houde, R. (eds.) 594–597 Sensory aids for the hearing impaired. IEEE Press, New York, (1980).

Sherrick, C.E.: Basic and applied research on tactile aids for deaf people: Progress and prospects. J. Acoust. Soc. Am., 75, 1325–1342, (1984).

Sherrick, C.E.: Tactual sound-and speech-analyzing aids for deaf persons. in: Christman, C.L., Albert, E.N. (eds.) 189–219 Cochlear Implants: A Model for the Regulation of Emerging Medical Device Technologies. Kluwer Academ. Publ., Norwell, MA, (1991).

Sparks, D.W., Ardell, L.A., Bourgeois, M., Wiedmer, R., Kuhl, P.K.: Investigating the MESA (Multipoint Electrotactile Speech Aid): The transmission of connected discourse. J. Acoust. Soc. Am., 65, 810–815, (1979).

Spens, K.E., Plant, G.: A tactual "Hearing" aid for the deaf. Paper presented at the 10th International Congress of Phonetic Sciences, Utrecht, 1–6 August. And Speech Trans. Lab. Q. Prog. Stat. Rep. 1. Stockholm Royal Institute of Technology, 52–56, (1983).

Summers, I.R.: Tactile Aids for the Hearing Impaired. Whurr, London, (1992a).

Summers, I.R.: Signal processing strategies for single-channel systems. in: Summers, I.R. (ed.) 110–127 Tactical Aids for the Hearing Impaired. Whurr, London, (1992b).

Tillmann, H.G., Piroth, H.G.: System for electrocutaneous stimulation. J. Acoust. Soc. Am. 79 (Suppl. 1.), 73–74, (1986).

Tyler, R.S., Summerfield, A.Q. Wood, E.J., Fernandes, M.: Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. J. Acoust. Soc. Am., 72, 740–752, (1982)

Van Doren, C.L.: A device for measuring spatiotemporal sensitivity. J. Acoust. Soc. Am., 81, 1906–1916, (1987)

Van Doren, C.L.: A model of spatiotemporal tactile sensitivity linking psychophysics to tissue mechanics. J. Acoust. Soc. Am., 85, 2065–2080, (1989).

Van Doren, C.L., Gescheider, G.A., Verrillo, R.T.; Vibrotactile temporal gap detection as a function of age. J. Acoust. Soc. Am., 87, 2201–2206, (1990).

Verrillo, R.T.: Comparison of vibrotactile threshold and suprathreshold responses in men and women. Percept. Psychophys., 26, 20–24, (1979).

Verrillo, R.T., Verrillo, V.: Sensory and perceptual performance. in: Charness, N (ed.)1–46 Aging and Human Performance. John Wiley, Chichester, (1985).

Weissenberger, J.M.: Communication of the acoustic environment via tactile stimuli. in: Summers, I.R. (ed.) 83–109 Tactile Aids for the Hearing Impaired. Whurr, London, (1992).

Weissenberger, J.M., Craig, J.C., Abbot, G.D.: Evolution and evaluation of a principal-components tactile aid for the hearing impaired. J. Acoust. Soc. Am., 90, 1944–1957, (1991).

PART XII
TRANSDUCERS FOR ACOUSTIC MEASUREMENT

Hideo Suzuki

Chiba Institute of Technology, Narashino, Japan

65 Condenser Microphones

Erling Frederiksen 1247

Introduction • General Principle • Transduction • Preamplifiers • Sensitivity • Frequency Response • Microphones and Sound Fields • Critical Design Parameters and Materials • Acoustic Impedance of Diaphragm • Dynamic Range • Microphones and International Microphone Standards • Microphones of Other Designs and Materials

66 Dynamic Pressure Microphones

Erhard Werner 1267

Introduction • Driving the Membrane • Generating the Electric Output • Advantages and Limitations

67 Dynamic Pressure Difference Microphones

Erhard Werner 1275

Introduction • The Frequency Response of a Dynamic Pressure Difference Microphone • The Directional Response of a Pressure Difference Microphone • The Ribbon Microphone • Cardioids and Other Directivity Patterns • Interference Microphones • Reflectors • Fixed and Variable Arrays

68 The Microflown Particle Velocity Sensor

Finn Jacobsen and Hans-Elias de Bree 1283

Introduction • The Microflown Transducer • Limitations • Calibration • Applications

69 Microphone Calibration

Erling Frederiksen 1293

Introduction • Calibrations Performed at Primary and Secondary Levels • Principal Types of Sound Field and Microphone Responses • Laboratory Standard Microphones Suited for Primary Calibration • Principle of Microphone Reciprocity Calibration • Primary Pressure Response Calibration by Reciprocity Technique • Primary Free-Field Response Calibration by Reciprocity Technique • Sound Field Corrections of Laboratory Standard and Measurement Microphones • Determination of Diffuse-Field Frequency Response • Working Standard Microphones • Sensitivity and Frequency Response Calibration by Comparison • Frequency Response Calibration by Electrostatic Actuator • Free-Field Response Calibration by Free-Field Comparison • Traceability and Accreditation • Dynamic Linearity • Field Calibration with Sound Field Calibrators

70 Piezoceramic Accelerometer

James Mathews 1313

Introduction • Technology Fundamentals • Construction • Selecting and Specifying • Mounting • Latest and Future Developments • List of Applicable Standards

71 Laser Doppler Velocimeter

Nobutoshi Yoshida 1329

Introduction • Operation Principle • Basic Construction • Applications

72 Capacitive Sensors

Masahiko Sakai 1339

Principle of Operation • Some Useful Advice for Capacitive Sensors • Applications

65

Condenser Microphones

Erling Frederiksen

Brüel & Kjær, Sound and Vibration Measurement A/S, Skodsborgvej 307, 2850 Nærum, Denmark

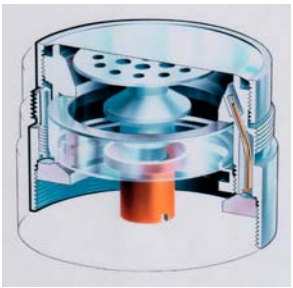
1 • Introduction

This chapter describes operation principle, design and performance of the most commonly applied types of measurement microphone, which are all pressure-sensing condenser microphones. These microphones that are designed for different sound fields and applications cover together a range that with respect to frequency and dynamics widely exceed the audible range of the human ear.

2 • General Principle

Most modern measurement microphones are pressure-sensing condenser microphones. Their principle of operation, which was invented by E. C. Wente in 1917 [1], is based on variation of the distance, and thereby the capacitance, between a moveable and a stationary plate, which are electrically conducting and form a capacitor. A stretched diaphragm that is typically 2–8 μm thick is mounted at one end of the, generally cylindrical, microphone body. This diaphragm constitutes the plate that is displaced by the sound pressure, while the stationary, parallel plate, called the back-plate, is placed behind it, inside a closed cavity. The distance between the plates, at rest, is typically 15–25 μm; see Figure 1a and b. The capacitance is from about 3 to 70 pF, depending on the diameter of the microphone, which, for the common models, is in the range between 3 and 24 mm.

(a)



(b)

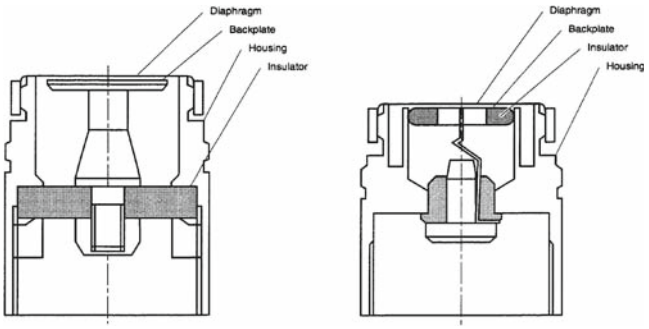


FIGURE 1 (a) Condenser microphone of traditional design. Its back-plate is mounted on a quartz insulator about 20 μm apart from the screwed-on diaphragm. (b) Modern condenser microphones that have their diaphragms laser-welded to rings, which are press-fitted onto the microphone housings. Insulator with back-plate (a metal film on its top) is also press-mounted into the right lower-cost microphone.

3 • Transduction

The transduction from sound pressure to voltage of a condenser microphone is, generally, based on the principle of a condenser that carries a constant electrical charge. With constant charge, any displacement of the microphone diaphragm will produce a proportional output voltage. This is shown by the equations below. Equation (1) is the fundamental equation for a condenser with constant charge while Equation (2) is deduced from this.

$$U(t)\cdot C(t) = U_0 \cdot C_0 \quad \text{or} \quad (U_0 + u(t))\cdot \frac{\varepsilon \cdot S}{d_0 + \Delta d(t)} = U_0 \cdot \frac{\varepsilon \cdot S}{d_0} \qquad (1)$$

$$u(t) = \frac{\Delta d(t)}{d_0}\cdot U_0 \qquad (2)$$

$C(t)$	Instantaneous capacitance of microphone
C_0	Capacitance of microphone with diaphragm at rest position ($\Delta d = 0$)
d_0	Rest distance between diaphragm and stationary plate
$\Delta d(t)$	Displacement of diaphragm from rest position
S	Area of stationary plate facing the diaphragm
$U(t)$	Instantaneous voltage between diaphragm and stationary plate

U_0 Rest voltage between diaphragm and stationary plate ($\Delta d = 0$)
$u(t)$ Output voltage of microphone ($U(t) - U_0$)
ε Dielectric constant of air (8.85×10^{-12} F/m).

The proportionality of the transduction exists for displacements of the diaphragm in both directions from its rest position. In the direction to the back-plate it is valid to a point, just in front of this, where there is no contact between the two electrodes. Displacement and voltage are proportional, even if the corresponding changes of the capacitance are, in fact, highly nonlinear. This relation between the parameters is often not recognised. See the description of distortion later under Dynamic Range and Reference [2] for more details.

The magnitude of the diaphragm displacement and the microphone sensitivity are functions of the diaphragm compliance that is inversely proportional to the tension in the diaphragm foil. The sensitivity range of typical models of condenser microphone goes from about 1 mV/Pa for the smaller models up to 100 mV/Pa for the larger ones.

A condenser microphone may either obtain its charge from an external polarisation voltage source, via a high resistance, or from a built-in electret, which is a permanently charged polymer or other type of high-quality electrical insulator; see Figures 2 and 3.

Electret microphones may either have their electret on the back-plate or have their polymer diaphragm charged to make up an electret. The first solution makes it possible to use metal diaphragms and obtain a temperature range and a stability that is similar to that of externally charged measurement microphones [3]. Very cheap, but less stable, microphones can be made with the latter design that is widely applied for microphones that are mass-produced for consumer products.

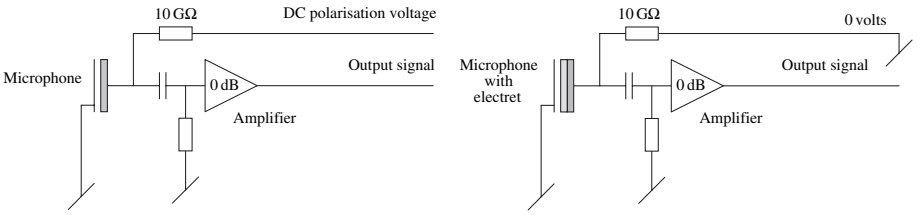


FIGURE 2 Principle of charging the air-gap capacitance by using an external voltage source (*left*) and by using an electret (*right*).

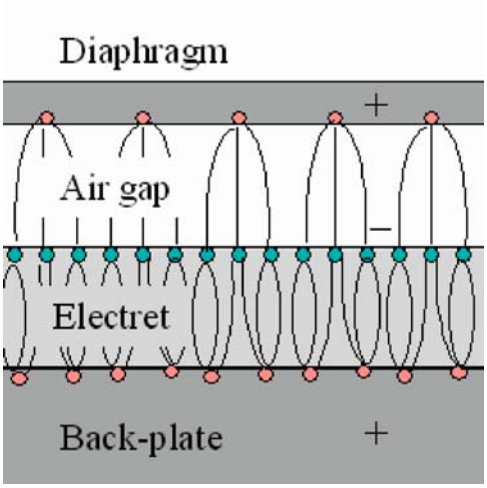


FIGURE 3 Charges of condenser microphone with back-plate-mounted electret. The electric field of the air gap is generated by the "frozen" charge (–) on the back-plate and that fraction of the image charge (+) that is located on the diaphragm. The polarisation is generally opposite to that of externally polarised microphones.

4 • Preamplifiers

The low capacitance of a condenser microphone makes this a source that can only withstand minor loading. Therefore, the microphone must be directly connected to a preamplifier that has a high input impedance and also a low output impedance that can drive the capacitive loading of the needed cable (typically 100 pF/m) – see Figure 4. Generally, microphones work over very large dynamic ranges that should also be covered by their

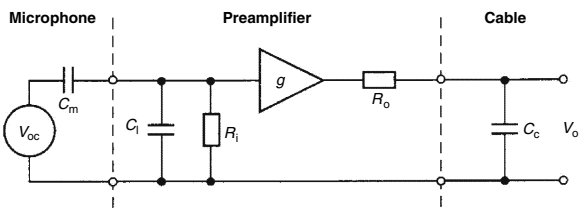


FIGURE 4 Circuit diagram with the elements that determine the electrical response of microphone, preamplifier and cable. Microphone capacitance (C_m), input capacitance (C_i) and resistance (R_i), voltage gain (g), output resistance (R_o) and cable loading (C_c).

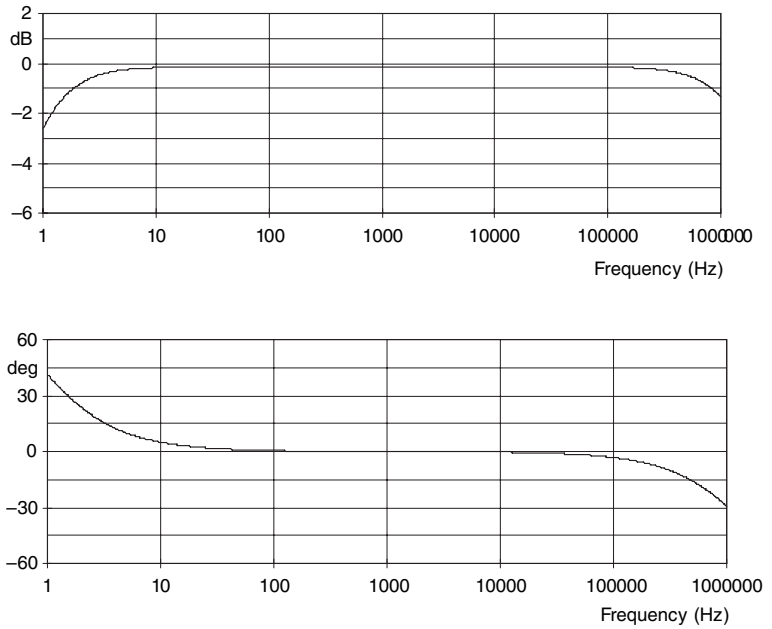


FIGURE 5 Magnitude and phase frequency responses of microphone–preamplifier circuit with the following typical parameters: $C_m = 18\,\text{pF}$, $C_i = 0.2\,\text{pF}$, $R_i = 10\,\text{G}\Omega$, $g = 0\,\text{dB}$ and $R_o = 30\,\Omega$, $C_c = 3\,\text{nF}$ (≈30-m cable).

preamplifiers. They are, in general, designed with unity voltage gain. Some types can cover a range from about 1 μV to 50 V, which corresponds to a difference of 154 dB between the lowest and the highest level. Modern preamplifiers use field-effect transistors (FETs) to obtain low inherent noise and high input impedance. A typical preamplifier circuit and its magnitude and phase frequency responses are shown in Figures 4 and 5, respectively. The low and the high cut-off frequency (−3 dB points) depend on the microphone and cable capacitance, respectively [4].

Instead of ordinary preamplifier, different types of high-frequency circuits may be used for converting the varying microphone capacitance to electrical signals. The microphone capacitance may make a part of a high-frequency circuit that can generate frequency or phase-modulated signals that are functions of the changes of the capacitance and thereby of the sound pressure. Such circuits are rarely used for general purposes, but they may be used down to the mHz range, which is not covered by ordinary preamplifiers.

5 • Sensitivity

The sensitivity of a microphone element consisting of a circular stretched diaphragm with a circular concentric back-plate is given by the equation below (2). The equation is valid for a parabolic displacement mode, which, in practice, means from low frequencies up to a frequency that makes 50–70 % of the diaphragm resonance frequency.

$$M_e = U_0 \cdot \frac{r_{diap}^2 \cdot \left(2 - \frac{r_{bp}^2}{r_{diap}^2}\right)}{8 \cdot \tau \cdot d_0}, \tag{3}$$

where

M_e	Sensitivity of microphone element (V/Pa)
U_0	Polarisation voltage (V)
r_{diap}	Diaphragm radius (m)
r_{bp}	Back-plate radius (m)
τ	Diaphragm tension (N/m)
d_0	Rest distance between diaphragm and back-plate (m)

Note: The diaphragm tension (τ) is the tension across an imaginary cut through the foil.

This equation indicates that a certain microphone sensitivity can be obtained in many ways, but any solution that is selected implies various consequences for frequency response, capacitance, inherent noise, reliability and stability of the microphone.

6 • Frequency Response

The main elements that determine the frequency response of a condenser microphone are the tension in and the mass of its diaphragm foil and also the mass and damping made up by the air in the narrow gap behind the diaphragm. These elements are parts of the frequently applied microphone model that is shown in Figure 6.

Another element is the internal cavity, which acts as a stiffness in addition to that of the diaphragm. The displacement of diaphragm creates a pressure in the cavity that reduces the pressure across the diaphragm itself and thus also the sensitivity of the microphone. The cavity should not be too small, as this will increase its stiffness, but, on the other hand, it should not be too large either, as this may lead to disturbing cavity resonances

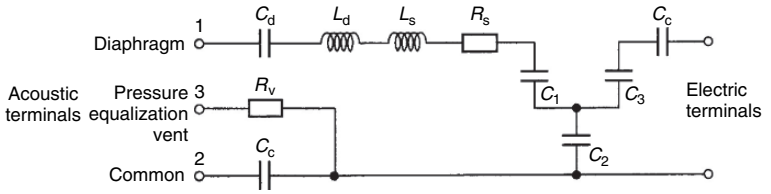


FIGURE 6 Impedance-analogy model of charged condenser microphone. The elements are diaphragm compliance (1/stiffness) and mass (C_d, L_d), air-gap mass and resistance (L_s, R_s), coupling elements (C_1, C_2, C_3), cavity compliance and vent resistance (C_c, R_v) and the electrical capacitance (C_e).

and a less ideal response at higher frequencies. The fraction of air stiffness is typically between 3 % and 18 % for small and large measurement microphones, respectively. The cavity must have a well-controlled venting that can partly ensure equalisation of static pressure changes and partly prevent equalisation of the low-frequency sound pressure that must be measured. Typically, the lower limiting frequency (−3 dB point) is between 1 and 3 Hz, but special microphone models are available for measurements at very low frequencies.

The electro-mechanical coupling of a charge-operated microphone can be represented by a T-network of three equally large capacitors, where the one that is common for the acoustical and the electrical branch has opposite sign of the other ones – see Figure 6 and Table 1. The last of the major model elements is the electrical capacitance across the air gap (valid with the diaphragm blocked at its rest position).

An example of a condenser microphone model is given in Table 1. The values are representative for several types of commonly used half-inch microphones that have a sensitivity of

TABLE 1 Model of half-inch microphone (sensitivity: 50 mV/Pa)

C_d	Diaphragm compliance	0.314×10^{-12} m^5/N
L_d	Diaphragm mass	600 kg/m^4
L_s	Air-gap mass	206 kg/m^4
R_s	Air-gap resistance (Q = 1)	56.2×10^6 Ns/m^5
C_c	Compliance of internal air-filled cavity	2.83×10^{-12} m^5/N
R_v	Resistance of pressure equalisation vent	36×10^9 Ns/m^5
C_1	Coupling compliance	5.66×10^{-12} m^5/N
C_2	Coupling compliance/capacitance	-5.66×10^{-12} m^5/N
C_3	Coupling capacitance	5.66×10^{-12} m^5/N
C_e	Electrical capacitance (blocked diaphragm)	17×10^{-12} m^5/N

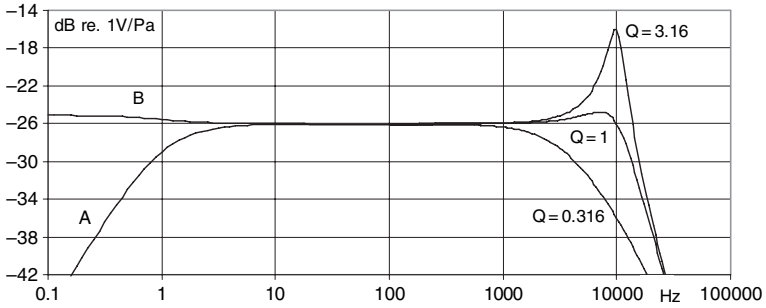


FIGURE 7 Magnitude frequency responses calculated for the model of half-inch microphone specified in Table 1. Curves A and B are valid with and without sound exposure of the static pressure equalisation vent. The influence of diaphragm damping (Q-value) is seen at high frequencies.

50 mV/Pa (corresponding to −26 dB re. 1V/Pa). Figure 7 shows the influence of the diaphragm damping (Q-factor) on the frequency response that, with critical damping, may become essentially flat up to the diaphragm resonance frequency. The response can be modified by changing the air gap and by applying back-plates of different size and pattern of holes – see Figure 1. This makes it possible to design dedicated types of microphone that have flat frequency responses, when applied in specific types of sound field. At low frequencies, the response depends on whether the outer part of the equalisation vent is exposed to the measured sound or not. The response rolls off, if the vent is exposed and equalises the cavity pressure with the ambient sound pressure (curve A), whereas it stays flat or increases just slightly if it is not exposed (curve B).

The obtainable sensitivity and frequency response are mutually dependent. If the flat response, shown in Figure 7, should be extended from say 10 to 20 kHz, the diaphragm resonance should be increased correspondingly. This might be achieved either by increasing the tension in the foil by a factor of 4 or by reducing its diameter by a factor of 2. In both cases, the sensitivity will be reduced by 12 dB – see Equation (7) and Figure 8. Other consequences would be lower diaphragm compliance and increased inherent noise.

Typically, LS1 and WS1 (Ø 23.77 mm) microphones work to 10–20 kHz, LS2 and WS2 to 20–40 kHz, WS3 (Ø 6.35 mm) to about 70 kHz and eighth-inch (Ø 3.175 mm) microphones to nearly 150 kHz.

Extended descriptions of condenser microphones are given in References [5,6].

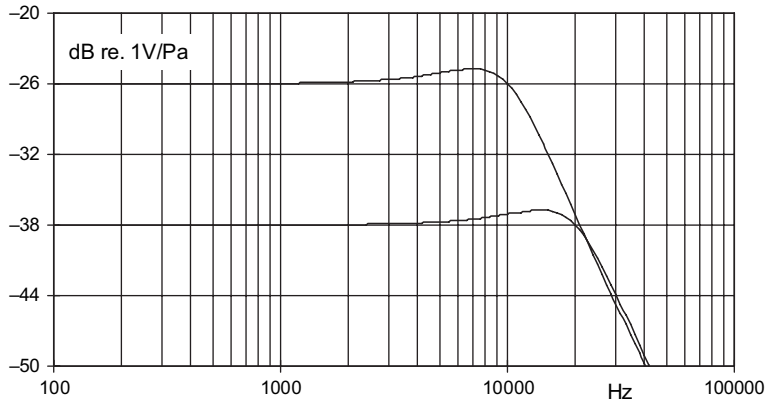


FIGURE 8 The change from the upper to the lower sensitivity and frequency response can be obtained by increasing the diaphragm tension by a factor of 4 or by reducing its diameter by a factor of 2.

7 • Microphones and Sound Fields

The type of frequency response discussed above is the pressure response or the response of the microphone diaphragm system itself. But, as is the case for other transducers and probes, the microphone influences the parameter it is to measure, i.e. the sound pressure. This influence, which depends on the type of sound field, must, therefore, be taken into account. There are three principal types of field – the pressure-field, the free-field and the diffuse-field [7] – and three corresponding types of microphone sensitivity and frequency response [8]. The terms "random incidence sound field" and "random incidence frequency response" are often used synonymously with diffuse sound field and diffuse-field frequency response, respectively. In practice, this is fully acceptable, even if there is a minor theoretical difference between the definitions of the fields [9].

The pressure response refers to a uniformly distributed pressure over the microphone diaphragm.

The free-field response refers to the pressure that would exist in the measurement point of a free sound field, in the absence of the microphone. The free-field response is a function of angle of sound incidence on the microphone body. Generally, the responses of free-field microphones are optimised to be flat for sound waves that impinge perpendicular to the front of their diaphragms. This incidence is usually named "zero degrees".

The diffuse-field response refers to the pressure that would exist in the measurement point, of a diffuse sound field, in the absence of the microphone.

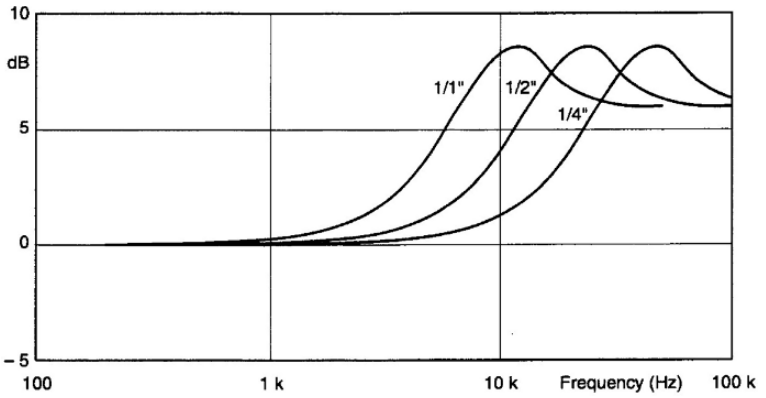


FIGURE 9 Typical changes of pressure at diaphragms of 1/1″, 1/2″ and 1/4″ microphones valid for free-field conditions and for 0° sound incidence.

The three responses are essentially equal at frequencies where the microphone diameter (d) is small compared with the wavelength (λ) of the sound, say for $d/\lambda < 0.07$. The differences at higher frequencies are due to changes of the sound pressure, at the microphone diaphragm, caused by the presence of the microphone body, which diffracts and reflects sound waves. Fortunately, due to equal size and shape, these pressure changes are essentially the same for all microphones of a certain model. Therefore, they do not need to be determined for every single unit but just for each model of microphone – see Figure 9.

Microphones are most often used with grids to protect their diaphragms. Such grids act as resonators and may, depending on angle of sound incidence, contribute to the pressure increase at higher frequencies – see Figure 10.

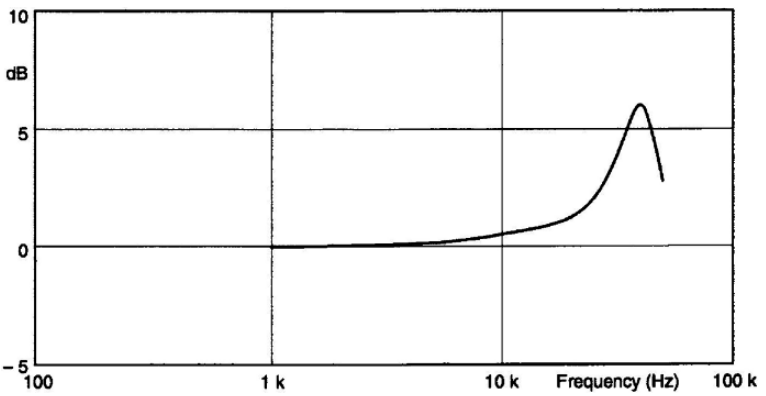


FIGURE 10 Influence of typical protection grid of 1/2″ microphone valid for free-field conditions and for 0° sound incidence.

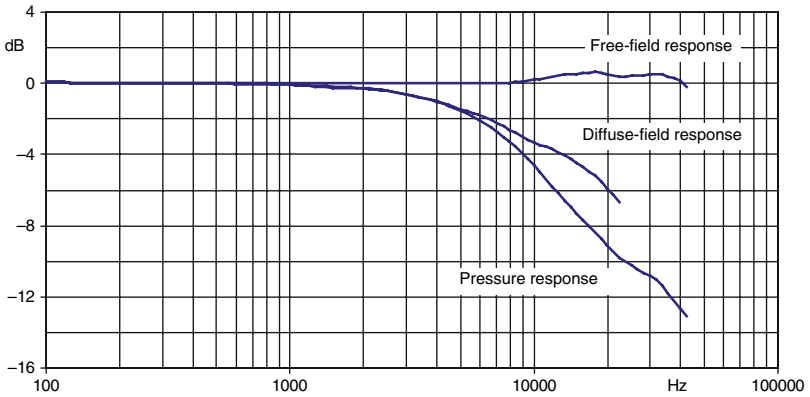


FIGURE 11 Free-field (0° incidence) and pressure responses of a microphone (Ø 12.7 mm and 12.5 mV/Pa) that is optimised for applications under free-field conditions.

The combined effect of reflection and grid lead to a gradually increasing change of pressure level that at the highest frequencies may become as high as 12–14 dB. This change is the reason for the difference between the pressure-field and the free-field response of a microphone. If the microphone has a flat pressure response, the free-field response will thus increase drastically at high frequencies. However, flat free-field frequency responses can be obtained by using back-plate designs, which creates a very heavy damping of the diaphragm resonance. The effect of this is that the pressure response rolls off with increasing frequency and thereby compensates for the influence of the pressure increase at the diaphragm – Figure 11.

When a microphone is placed in a diffuse sound field, the pressure will change, as it does in a free-field, but the change will be far less and will typically not exceed 3 dB – see Figure 11. Microphones, with optimised diffuse-field responses [4], are available, but often diffuse-field measurements are performed with pressure-field microphones, especially, when only low and medium–high frequencies are of interest.

8 • Critical Design Parameters and Materials

High stability and reliability are major requirements of measurement microphones. The sensitivity and frequency response should ideally not vary with ambient conditions and permanent changes should neither occur accidentally nor over time. The design parameters that are most critical in these respects are polarisation voltage, tension in diaphragm foil and

distance between diaphragm and back-plate; – see Equation (2). Any relative changes in these parameters cause equally large relative changes of the sensitivity.

As the sensitivity is proportional to the microphone charge, this must be supplied from a stable electret [3] or a stable voltage source. Therefore, in the latter case, the microphone insulation resistance must be at least ten thousand times higher than the resistance of the source, in order to avoid voltage attenuation and a disturbing inherent noise caused by a leakage current. The back-plate should thus be mounted on a high-quality insulator (quarts, sapphire or else). However, a leakage current might also flow through the air in the microphone air gap, where the field-strength is typically as high as 100,000 V/cm. Current through the air can be prevented by making highly polished, flat and clean surfaces on the two electrodes.

Optimisation of the sensitivity and the frequency response of a microphone leads to a thin diaphragm that is highly stressed (high force/length) and to a very high loading of the material itself, typically $5 \times 10^8 \mathrm{N/m^2}$. This loading may cause a sagging in the material and a tension in the foil that tends to decay over time. The decreasing tension and the associated increasing sensitivity may be prevented by proper selection of material and by artificial aging of the microphone. Long-term changes may be as low as 1 dB/1000 years. However, random short-term changes, caused by mutual displacement between the parts of a microphone, may appear and be up to few tenths of a dB depending on the environment of use and on the level of care in the microphone handling.

Thermal expansion may influence the diaphragm tension and thus also the temperature coefficient of the microphone. Therefore, the thermal expansion coefficients of the diaphragm and of the housing, to which it is attached, should match each other very closely.

The last of the critical parameters is the small distance (typically 20 μm) between diaphragm and back-plate. Proper design and selection of materials may ensure this to stay within a fraction of a micrometer over time and over the temperature range of operation.

9 • Acoustic Impedance of Diaphragm

The acoustic impedance of the diaphragm is important in some connections. This is especially the case when microphones are used for measuring the sound pressure inside small cavities, as may be the case during microphone calibration and during

testing of hearing aids. In such cases, the microphone often influences the sound pressure by its diaphragm impedance, which should, therefore, be taken into account.

A microphone model, like the one described under Frequency Response, may be used for calculation of acoustic diaphragm impedance. The equation is valid for non-loaded electrical microphone terminals, but the loading effect of common types of preamplifier is very small and can generally be neglected – see Equation (3),

$$Z_d = \frac{1}{j\omega \frac{C_d \cdot C_c}{C_d + C_c}} + j\omega(L_d + L_s) + R_s \quad [\mathrm{Ns/m^5}], \tag{4}$$

where ω is the angular frequency, while all other parameters are defined in connection with Equation (2).

In practice, the term "equivalent (air) volume" is often used instead of acoustic compliance or impedance, especially at low and medium frequencies, as the value of this parameter is more directly comparable with the volume of the cavity, where the microphone is inserted. Equation 4 defines the equivalent volume and Table 2 gives typical values,

$$V_e = \frac{\gamma \cdot P_s}{j\omega \cdot Z_d} \quad [\mathrm{m^3}], \tag{5}$$

where

V_e Equivalent air volume
γ Ratio of specific heats of air (1.402)
P_s Ambient static pressure.

When "equivalent volume" is stated, the value is most often meant as being valid at the nominal static pressure (101.325 kPa). The equivalent volume of a diaphragm is, in principle, a function of the static pressure.

TABLE 2 Typical values of equivalent air volume of microphone diaphragms

Microphone diameter	LS1/WS1 23.77 mm	LS2/WS2 12.7 mm		WS3 6.35 mm		3.18 mm
Sensitivity	50 mV/Pa	50 mV/Pa	12 mV/Pa	4 mV/Pa	1.6 mV/Pa	1 mV/Pa
Equivalent Volume	140 mm^3	40 mm^3	10 mm^3	0.6 mm^3	0.25 mm^3	<0.1 mm^3

10 • Dynamic Range

The lower limit of the dynamic range of a measurement system is generally determined by the inherent noise of the microphone and the associated preamplifier.

The microphone itself produces a noise that is due to thermal or Brownian movements of its diaphragm. The shape of the noise spectrum can be explained by the model described under Frequency Response – see Figure 6. The major noise source of the model is the diaphragm damping resistance (R_s), which produces Nyquist (or Johnson) noise. Its contribution to the noise at the microphone output can be understood by moving its series noise source to the microphone input terminals. At the output, this noise source that has a flat spectrum will produce a noise with a spectrum shaped like the microphone frequency response.

The power of the broadband noise of a microphone, which is concentrated within its normal frequency operation range, is, in a simple way, related to the compliance (or equivalent volume) of its diaphragm system – see Reference [10] and Equation 4.

$$\bar{p}_n^2 = \frac{k \cdot T}{C_m} = \frac{k \cdot T \cdot \gamma \cdot P_s}{V_e}, \tag{6}$$

where

$\bar{p}_n$	Time average of broad-band noise pressure on diaphragm (RMS)
k	Boltzmann constant (1.38×10^{-23} Nm/K)
T	Absolute temperature (K)
C_m	Equivalent volume of microphone diaphragm system (m^3).

Calculation of the broadband noise of the microphone model above (50 mV/Pa and 12.7 mm diameter) gives a noise pressure of 120 μPa or 15.6 dB SPL, while a 24 mm microphone with an equivalent volume of 140 mm^3 generates a level of 10.2 dB.

Note: It is interesting to notice that the Brownian noise related to the equivalent volume (1.3 cm^3) of the human eardrum is 21 μPa and thus very close to the threshold of hearing (20 μPa).

The linear (20 Hz–200 kHz) and A-weighted electrical noise of a modern FET preamplifier, with source capacitance like that of a typical 12.7 mm microphone (18 pF), are of the order 8 and 2 μV, respectively. The equivalent inherent noise levels depend on the sensitivity of the applied microphone. For a sensitivity of 50 mV/Pa, the levels are 18 and 6 dB, respectively. In general,

the levels of preamplifier noise, for 1-inch and 0.5-inch microphones, are lower than those of the microphone itself, while the preamplifier noise is dominating for smaller microphones with low capacitance and sensitivity.

The upper limit of the dynamic range is either set by the second harmonic distortion of the microphone, which increases proportionally with the sound pressure, or by signal clipping of the preamplifier. Capacitance in parallel with the active capacitance between diaphragm and back-plate is the major reason for microphone distortion within the frequency range where the diaphragm displacement is controlled by its stiffness. The parallel capacitance leads to a second harmonic component, whose level increases proportionally with the sound level. Even if this capacitance is kept low, distortion is created due to the essentially parabolic mode of diaphragm displacement. As the displacement

TABLE 3 Specifications of typical measurement microphones

Sound field/ application	Diameter (mm)	Sensitivity (mV/Pa)	Maximum frequency (kHz)	Level of inherent noise (dB re. 20 μPa)	Level of 3 % distortion (dB re. 20 μPa)	Capacitance (pF)
Free-field/general	23.77	50	20	11	147	66
	12.7	50	20	15	148	18
	12.7	12.5	40	21	162	18
	6.35	4	100	34	164	6.4
Pressure- and diffuse-field/general	23.77	50	10	10	149	55
	12.7	50	10	15	148	18
	12.7	12.5	20	21	163	18
	6.35	1.6	70	38	172	6.4
Pressure-field/confined space, high spatial resolution	3.175	1	140	45	168	3.5
Free-field/low Level	23.77	1000**	14**	0*	102*	40
Pressure-field/high Level	6.35	0.1	25	55	187	3.3
Pressure-field/surface pressure	20*	11.2	20	30***	140	–

* axial length: 2.5 mm.
** with dedicated preamplifier.
*** includes noise of built-in CCLD preamplifier.

at the edge of the back-plate is relatively smaller than at the centre, the less active edge capacitance constitutes a loading for that of the centre and thereby causes distortion – see reference [2]. In general, microphone manufacturers specify the level of 3% distortion, but typically microphones can be used up to 10 dB higher levels, where the distortion is approximately 10%, unless the associated preamplifier causes clipping of the microphone output signal. See typical distortion levels in Table 3.

11 • Microphones and International Microphone Standards

Microphones are used over wide ranges of level and frequency (see Figure 12), under different sound field conditions and under ambient conditions that may vary from controlled laboratory

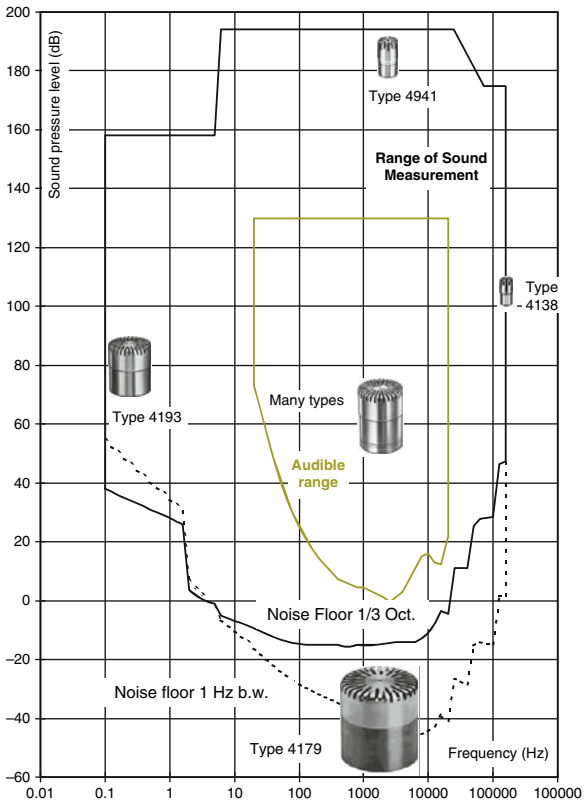


FIGURE 12 Range of sound level and frequency covered by condenser measurement microphones.

conditions to extreme wind, temperature and pressure conditions that, for instance, occur in the upper atmosphere during flight testing. Therefore, there is a need for many different models of microphone with properties that match the large variety of applications.

Two categories of measurement microphone are internationally standardised. They are called laboratory standard microphones [8] and working standard microphones [11] and are designated the symbols LS and WS, respectively. An added number indicates the diameter of the type. This may be "1" for one inch (23.77 mm), "2" for half-inch or "3" for quarter-inch diameter. Laboratory standards (LS1 and LS2) are generally used for primary calibration and as national reference standards, while working standards (WS1, WS2 and WS3) are used as secondary standards and for many different types of measurement in the field. WS2 microphones are used for the most common measurements within the audible range. Larger microphones are generally used for measurement at lower levels and lower frequencies, while smaller microphones are used at higher levels and higher frequencies. Properties of general and specific types of microphone are listed in Table 3. The specific types mentioned are designed for measurements of high spatial resolution, within confined spaces, at very low levels [12], at very high levels and at outer surfaces of a flying aircraft.

12 • Microphones of Other Designs and Materials

Most measurement microphones are traditional condenser microphones, basically consisting of turned metal parts and high-quality insulators. Some of them are equipped with diaphragms of sophisticated alloys that are mounted by laser welding and are stable and resistant to heat and to corrosion. They are designed for both general and specific purposes, but they may, for some applications, be too large and too expensive.

Therefore, large microphone arrays, which are used for sound mapping and spatial transformation of sound fields, may be equipped with microphones that are mass-produced for consumer products. These microphones are generally small electret microphones with polymer diaphragms and built-in preamplifiers. As not all units of such types may meet the requirements of the application, with respect to stability and uniformity of properties, the applied units are artificially stabilised and selected from large batches. Arrays with such microphone may contain hundreds of microphones.

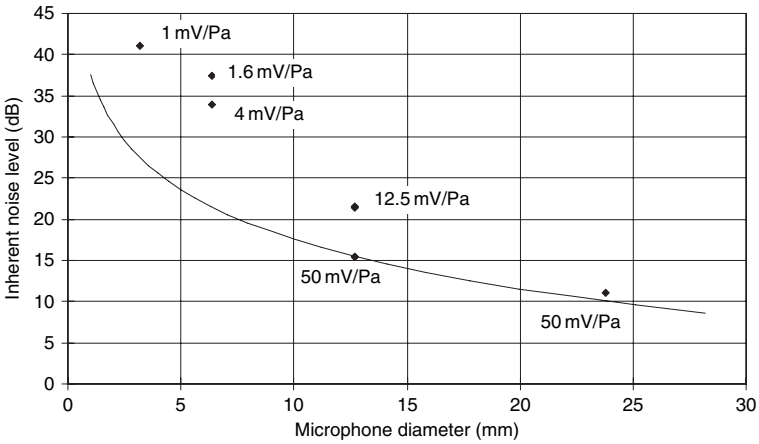


FIGURE 13 Achievable inherent noise level as a function microphone diameter. A frequency range up to 20 kHz and a fraction of air stiffness of 12.5 % are considered. The points and the sensitivities indicate noise levels of existing types of microphone.

The inherent noise of a microphone is related to its equivalent diaphragm volume and, therefore, also to its mechanical dimensions – see under Dynamic Range and Equation (5). The result of a microphone noise analysis, based on this fact, is shown in Figure 13. The graph shows the achievable inherent noise level as a function of the diameter of the microphone, considering that the microphone should operate up to 20 kHz and that the fraction of air stiffness should be 12.5 %, which is typical for measurement microphones. The graph is taken from Ref. [13] that gives more details about the calculation conditions.

The figure shows that there is room for improvement of smaller microphones (for $d < 12$ mm) with respect to inherent noise and sensitivity. Improved microphones should be made with a much lower diaphragm tension. In practice, however, a low tension is not easy to control. The related small lengthening of the diaphragm material makes its tension highly sensitive to thermally and mechanically caused dimension changes, and thus, the diaphragm unstable and not suited for a measurement microphone. This problem might be solved by making new types of microphone of only one, highly uniform material, such as monocrystalline silicon that is used for micro-mechanical devices [14]. Such microphones are now being produced, but they are yet not known to be applied for measurement purposes.

More details about microphone designs, accessories and applications may be found in References [4,5,15].

References

1. E. C. Wente, "A Condenser Transmitter as a Uniformly Sensitive Instrument for the Absolute Measurement of Sound Intensity," Phys. Rev. 10 (1), 39–63 (1917)
2. E. Frederiksen, "Reduction of Non-linear Distortion in Condenser Microphones by Using Negative Load Capacitance", in Brüel & Kjaer Technical Review No. 1 (1996)
3. E. Frederiksen, N. Eirby and H. Mathiasen, "Prepolarized Condenser Microphones for Measurement Purposes," in Brüel & Kjaer Technical Review No. 4 (1979)
4. Brüel & Kjaer Microphone Handbook (BA5105), Theory (Vol. 1) and Data (Vol. 2) (1996)
5. M. S. Hawley, F. F. Romanow and J. E. Warren, "The Western Electric 640AA Capacitance Microphone: Its History and Theory of Operation," in G. S. K. Wong and T. F. W. Embleton (Ed.), AIP Handbook of Condenser Microphones, American Institute of Physics, New York (1994), Chapter 2
6. A. J. Zuckerwar, "Principles of Operation of Condenser Microphones" in G. S. K. Wong and T. F. W. Embleton (Ed.), AIP Handbook of Condenser Microphones, American Institute of Physics, New York (1994), Chapter 2
7. IEC (International Electrotechnical Commission) 50(801), Vocabulary: Acoustics and electroacoustics, parameters 23–28 and 23–31, 3 Rue de Varembé, CH-1211 Geneva, Switzerland
8. IEC (International Electrotechnical Commission) 61094-1, Measurement microphones, Part 1: Specifications for laboratory standard microphones, pp. 8–11, 3 Rue de Varembé, CH-1211 Geneva, Switzerland
9. IEC (International Electrotechnical Commission) 61183, Electroacoustics – Random-incidence and diffuse-field calibration of sound level meters, p. 9, 3 Rue de Varembé, CH-1211 Geneva, Switzerland
10. V. Tarnow, "Thermal Noise in Microphones and Preamplifiers," in Brüel & Kjaer Technical Review No. 3 (1972)
11. IEC (International Electrotechnical Commission) 61094-4, Measurement microphones, Part 4: Specifications for working standard microphones, 3 Rue de Varembé, CH-1211 Geneva, Switzerland
12. E. Frederiksen, "Microphone System for Extremely Low Sound Levels," in Brüel & Kjaer Technical Review No. 3 (1984)
13. E. Frederiksen, "Measurement Microphones and Determination of their Properties, "Simposio Brasileiro de Metrologia em Acustica e Vibracoes – SIBRAMA, Rio de Janeiro (2002)
14. P. R. Scheeper, AGH van der. Donk,. W. Olthuis, P. Bergveld. "A Review of Silicon Microphones." Sensors and Actuators A., 44, 1–11 (1994)
15. Michael Gayford, "Microphone Engineering Handbook" Focal Press, Oxford (1994)

66

Dynamic Pressure Microphones

Erhard Werner

Tannenweg 16, D 29693 Hademstorf, Germany

1 • Introduction

In 1876, Alexander Graham Bell described his first telephone with a microphone using magnetic induction to convert the voice input into an electric output signal. The basic principle led to a variety of designs optimized for different needs, from hearing impaired users to singers or broadcast announcers. From the various sound pressure versions, only the moving coil design is still in mass production for speech and music application. It is correctly classified as electrodynamic,[1] but commonly called the dynamic microphone.

Only four parts are necessary to build a simple version of this principle (Figure 1): The membrane, the coil, a magnet and an iron cup that defines the outer border of the gap for the moving coil and also makes it possible to fix the rim of the membrane. This simple design will however have a restricted performance. Low mechanical damping results in a narrow frequency response around the resonance, somewhere within the frequency range of interest. Extreme damping of the resonance, on the contrary, will flatten and widen the response range but reduce the output

[1]Reference [1], cl.5.1

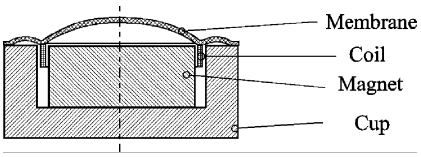


FIGURE 1 Basic components of an electrodynamic microphone.

level. Low acoustical input may then suffer from noise contributions of microphone and amplifier. With additional elements such as resonance chambers and acoustic resistances, amplitude response and sensitivity can be optimized, but with disadvantages for the phase performance. Despite these facts, dynamic microphones are used in great numbers for speech and music as an alternative to the "top" class of condensor studio microphones.

2 • Driving the Membrane

Either the pressure p or velocity v of an acoustic field can be of interest as input for a microphone. Membrane-based transducers always need a driving force to move the membrane. Reference to sound pressure is therefore preferable even when the relation between pressure and velocity is known from the wave propagation conditions. Except for special cases, the microphone response is always given for plane waves under free field conditions in calm air.

Pressure and velocity of a freely propagating plane wave are in phase and have constant amplitude relation. Microphones will cause effects such as reflection, refraction, and diffraction to an extent depending on their size, shape and material properties. The pressure amplitude P of a plane wave traveling in the x-direction at frequency f and speed c, which results in an acoustic field with pressure

$$p(x,t) = P\cos(2\pi f\,(t - x/c)), \tag{1}$$

must be replaced by the local amplitudes P_{m} at the surface of the microphone membrane (Figure 2).[2]

The ratio of P_m to P is dependent on frequency . Algebraic solutions are possible for very simple models only: however, some estimations are available (e.g., the total reflection at an extended rigid plane surface will cause 6 dB pressure increase) and today's computing power allows numerical results based on

[2]Reference [2], p. 11

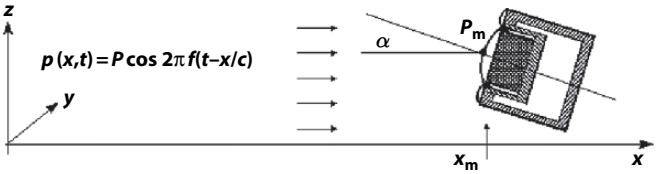


FIGURE 2 Dynamic pressure transducer (schematic cross section) and plane wave.

known boundary conditions for nearly any shape and material properties.

With the microphone axis inclined at an angle α to the direction of wave propagation, the phase of the pressure will vary across the membrane. Diffraction due to shape and extension of mechanical parts in front of the membrane usually keeps the "phase loss"[3] sufficiently below total extinction at critical membrane diameters. For accurate measurements, the remaining difference between on-axis and off-axis responses is accounted for in microphones that are specifically designed for operation at 0° incidence (facing the source) or 90° incidence.[4]

For sound pressure measurement, the backside of the membrane has to be acoustically separated from the source. Only a very small hole in the otherwise closed cup is needed for barometric equalization. The enclosed air volume is substantial part of the acousto-mechanical design.

The total force f_m on the membrane at a reference position x_m results from the integration of the differential forces $p_\mathrm{m}(x,y,z)$ $\mathbf{d}\boldsymbol{A}$ over its surface A. The local pressure amplitudes P_m also include phase shifts from off-axis wave incidence and membrane curvature. In a well-designed transducer, the membrane is only free to move in the axis direction and so the driving force is assumed to be in the normal direction n.

$$f_{\mathrm{m,n}}(x_\mathrm{m},t) = \int\int p_\mathrm{m}(x,y,z)\mathrm{d}A_\mathrm{n} = A_n P_\mathrm{me}(f,\alpha)\cos(2\pi f\ t - \varphi_\mathrm{me}) \quad (2)$$

Membranes are usually circular, with radius r, and so the effective surface area A_n is $r^2\pi$. The differences of the effective pressure amplitude P_me and the overall phase shift φ_me from the free-field values P and $2\pi f x_\mathrm{m}/c$ are negligible for small ratios of membrane diameter $2r$ to wavelength λ. Even at the upper end of the operating range of a microphone, their effect to the frequency

[3]Reference [2], p. 12
[4]Reference [2], p. 77

response is often far less than that of the mechanical system. The simplest mechanical model for this system contains mass m, damping coefficient b, and spring force coefficient k. It leads to the well-known differential equation (3) for the displacement x of the membrane from its zero position when a force $f(t)$ is applied.

$$\begin{aligned} &m\frac{d^2x}{dt^2}+b\frac{dx}{dt}+k\,x=f(t)\\ &\text{or}\\ &dm\frac{dv}{dt}+bv+k\int v dt=f(t)\\ &\text{with}\\ &v(t,\omega)=\frac{dx}{dt} \end{aligned} \tag{3}$$

For sinusoidal excitation $f(t)$, the solution of (3) can be obtained by using either trigonometric functions or complex variables in exponential presentation. This is a convenient way to describe the mechanical performance by electric analogies and to combine the design with accompanying electrical circuitry. Because there are different possible analogies, the corresponding characteristics and conversion multipliers should be carefully inspected.

3 • Generating the Electric Output

The solution of equation (3) for the velocity of the membrane and the connected coil is depending on frequency in both amplitude and phase φ (4),

$$\begin{aligned} &v(t,\omega)=\frac{\omega F/m}{\sqrt{(\omega^2-\omega_0^2)^2+\gamma^2\omega^2}}\cos(\omega t-\varphi(\omega))\\ &\text{with}\\ &\varphi(t,\omega)=\arctan((\omega_0^2-\omega)/(\omega\gamma)), \end{aligned} \tag{4}$$

where $F=A_{\mathrm{n}}P_{\mathrm{me}}$, $\omega_0^2=k/m$ and $\gamma=b/m$. Maximum velocity is always reached at resonance ω_0. Assuming constant magnetic induction B across the air gap the moving coil e.m.f. $e(t)$ between the ends of its total length l is computed by a multiplication of its velocity values (4) with B and l (5).

$$e(t)=Blv(t)=Bl\frac{F}{m}\frac{\omega}{\sqrt{(\omega^2-\omega_0^2)^2+\gamma^2\omega^2}}\cos(\omega t-\varphi(\omega)). \tag{5}$$

From this, the amplitude response E/P_{me} can be separated and written as

$$G=\frac{E}{P_{\mathrm{m}}}=Bl\frac{A_{\mathrm{n}}}{m}\frac{\omega}{\sqrt{(\omega^2-\omega_0^2)^2+\gamma^2\omega^2}} \tag{6}$$

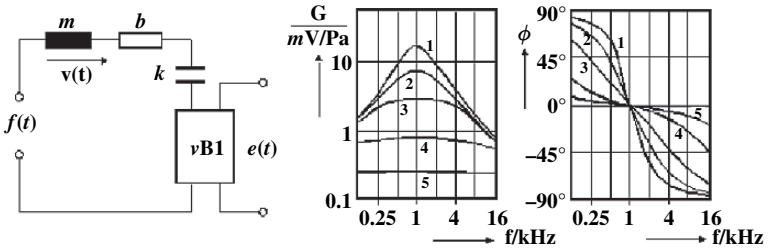


FIGURE 3 Equivalent electrical circuit, amplitude response, and phase response.

The phase response has been given in (4). Figure 3 shows one of the various electrical analogies with both the amplitude and the phase responses for a numerical example. With given values for magnetic induction, membrane area, and length of the coil wire, the maximum amplitude response at resonance frequency ω_0 is depends on the reciprocal of the damping coefficient b (7):

$$G_{\max} = G(\omega_0) = Bl\, A_n/(m\gamma) = Bl\, A_n/b \tag{7}$$

The 3-dB bandwidth results from $G^2_{3dB}/G^2_{\max} = 0.5$ in (6) and can be influenced by changing any of the three characteristics in (3). Possible changes however are limited for acoustical and design reasons. For audio application, it is preferred to match the maximum response with human hearing performance. This determines the ratio of mass and stiffness. Geometrical limitations and material characteristics then lead to a remaining range to vary the damping. With higher damping, the bandwidth can be extended to suitable values albeit with loss of sensitivity (Figure 3).

To overcome this limitation, additional coupled resonance chambers can be introduced to widen the output range without losing sensitivity. Despite the substantial changes of the phase response that result, such microphones are used to advantage for speech and music.

4 • Advantages and Limitations

Dynamic microphones need no power supply and can be constructed for many years of reliable operation with highly constant characteristics under many different environmental conditions. As with other microphone principles used outdoors, the influence of moisture and weather exposure can be reduced by proper choice and surface treatment of materials such as

metals and damping tissue. The absence of electronic circuitry provides high immunity against high-frequency electromagnetic fields if state-of-the-art design is obeyed and suitable cabling is used. However, external magnetic fields (mainly from a.c. mains power) may induce interfering voltages in the coil. An effective remedy is the use of fixed coils near the moving coil, connected in antiseries with it, to extinguish the unwanted interference. Although dynamic microphones are mainly designed for speech and music application, they can also be used for scientific and industrial purposes, such as for monitoring remote locations, if their limitations are taken into account.

High-quality dynamic microphones not only have long individual life, but famous models designed and built 50 and more years ago are still in production. All the main design aspects for the acoustic performance are still used, but remarkable progress has been made with new materials. Rare earth alloys offer far higher magnetic energy than earlier iron or oxides. New tissues from synthetics have been introduced, with less variability in damping properties and less variation with aging. New production technologies have brought advanced design features and reduced tolerances. There is still opportunity for even higher performance and precision. Giving up the zero power solution in dynamic microphones by introduction of active electronic circuits for more flexibility in signal processing would open another level of competition with condensor microphones and their highly advanced capsule and circuit design.

An example of an actual pressure microphone designed for outdoors broadcast application is shown in Figure 4. The transducer is made of more components than the basic setup of



FIGURE 4 Pressure microphone for OB application (courtesy Sennheiser).

Figure 1. These include layers to reduce wind and pop noise and to protect the membrane against dirt and other particles. The housing volume is part of the design for low frequency response.

The on-axis response is increased by about 6 dB starting from 1000 Hz. This leads to improved speech intelligibility. The drop above 12 kHz reduces signals from sources other than speech at all directions. The membrane dimensions and the surrounding mechanical components contributing to highly directive behavior at frequencies above the human speech spectrum, which provides an improved signal to noise ratio.

References

1. IEC 60268-4 (2004–02) Sound system equipment – Part 4: Microphones
2. Gayford, Michael (Ed.), Microphone Engineering Handbook, Butterworth and Heinemann: London 1994.

Further Reading

Abbagnaro, Louis A. (Ed.), *Microphones, AES Anthology*, New York 1979
Bauer, B.B., "A Century of Microphones," Proceedings of the IRE, 1962, pp. 719–729.
Beranek, L.L., *Acoustic Measurements*, Wiley / Chapmann and Hall, London 1949
Griese, H.-J., "Die Entwicklung der dynamischen Mikrofone in den letzten 25 Jahren," Funk-Technik 25 (1970), No. 11, pp. 413–416
Olson, Harry F., *Acoustical Engineering*, Princeton 1957
Olson, Harry F., "Microphones for Recording," Journal of the Audio Engineering Society, 25 (1977), pp. 676–684.
Reichardt W., *Grundlagen der Elektroakustik*, 3. Ed., Leipzig 1960 and following edititions

67

Dynamic Pressure Difference Microphones

Erhard Werner

Tannenweg 16, D-29693 Hademstorf, Germany

1 • Introduction

A microphone with a diaphragm that is exposed to the sound field only on one side responds essentially to the sound pressure. This quantity is a scalar, and thus, pressure microphones are essentially omnidirectional (see Chapter 66). However, directional microphones are useful because they make it possible to focus on a source and suppress background noise from other directions. Most directional microphones respond to the gradient of the sound pressure, to combinations of the sound pressure and its gradient, or to combinations of higher order spatial derivatives of the sound pressure.

2 • The Frequency Response of a Dynamic Pressure Difference Microphone

A microphone with a diaphragm that is exposed to the sound pressure from both sides will respond to the difference between the two pressures. For a plane wave that propagates in the x-axis direction, the sound pressure, the gradient of the pressure, and a pressure difference are respectively

$$p(x) = A\mathrm{e}^{\mathrm{j}(\omega t - kx)}, \tag{1}$$

$$\frac{\partial p}{\partial x} = -\mathrm{j}kA\mathrm{e}^{\mathrm{j}(\omega t - kx)} = -\mathrm{j}kp(x), \tag{2}$$

$$\begin{aligned}\Delta p &= A\left(\mathrm{e}^{\mathrm{j}(\omega t - k(x+\Delta x/2)} - \mathrm{e}^{\mathrm{j}(\omega t - k(x-\Delta x/2)}\right) = -2A\mathrm{j}\sin(k\Delta x/2)\mathrm{e}^{\mathrm{j}(\omega t - kx)} \\ &= -2\mathrm{j}\sin(k\Delta x/2)p(x),\end{aligned} \tag{3}$$

where $k = \omega/c$ is the wavenumber. Obviously,

$$\Delta p \to \frac{\partial p}{\partial x}\Delta x \tag{4}$$

when the distance Δx is much smaller than a wavelength, which is usually the case. Under these conditions, the microphone responds essentially to the gradient of the pressure; hence, such microphones are generally known as "pressure gradient" microphones. Because the difference pressure Δp and thus the force that drives the microphone diaphragm is proportional to the frequency, it follows that the expression for the electromotive force of a dynamic pressure microphone with mechanical impedance Z_m (Eq. (6) in Chapter 66) should be modified accordingly,

$$\begin{aligned}e(\omega) &= \frac{BlF}{Z_\mathrm{m}} = -\frac{\mathrm{j}kBl\Delta xSp}{Z_\mathrm{m}} = \frac{BlS\Delta xp\omega^2/mc}{\omega_0^2 - \omega^2 + 2\mathrm{j}\omega\delta} \\ &= \frac{BlS\Delta xp\omega^2/mc}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\omega\delta)^2}}\mathrm{e}^{\mathrm{j}\varphi},\end{aligned} \tag{5}$$

where S is the area of the diaphragm, Bl is the product of the magnetic induction and the length of the wire of the coil, ω_0 and δ are the natural frequency and damping coefficient of the mechanical system respectively, and

$$\varphi = \arctan\left(\frac{\omega_0^2 - \omega^2}{2\omega\delta}\right) - \frac{\pi}{2} \tag{6}$$

is the phase. The approximation of the pressure gradient by a pressure difference gives rise to an additional factor of $\sin(k\Delta x/2)/(k\Delta x/2)$, which is close to unity in the working frequency range of the microphone but decays rapidly to zero when Δx approaches one wavelength.

Inspection of Eq. (5) leads to the conclusion that the diaphragm should be mass controlled if the microphone is to have a flat frequency response. This is obtained by shifting the resonance frequency to a low frequency.

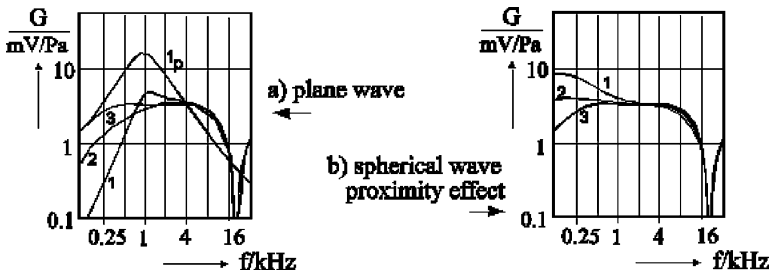


FIGURE 1 On-axis amplitude response of a dynamic pressure difference transducer (**a**) in a plane wave with a resonance at 1000 Hz (*curve 1*), 250 Hz (*curve 2*), and 125 Hz (*curve 3*), and (**b**) at a distance of 5 cm (*curve 1*), 15 cm (*curve 2*), and 2 m (*curve 3*) from a point source. The response of a dynamic pressure transducer (non-difference) corresponding to curve 1 is shown in curve 1_p.

The frequency response curves shown in Figure 1 correspond to the curves for dynamic pressure microphones shown in Figure 3 in Chapter 66, that is, they have been calculated assuming similar values of the various parameters. With the original resonance at 1 kHz and a damping coefficient of $2500\,s^{-1}$, the maximum response is reduced by about 10 dB compared with the pressure case (compare curves 1 and 1_p). With the resonance shifted to 250 Hz, the useful frequency range is extended (curve 2), and the response gets better still with the resonance shifted to 125 Hz and reduced damping (curve 3). Above 16 kHz, there is a steep roll off in sensitivity because Δx equals half a wavelength at 10 kHz.

If the microphone is exposed to a simple spherical wave instead of a plane wave Eqs. (1) and (2) become, respectively,

$$p(r) = \frac{A e^{\mathrm{j}(\omega t - kr)}}{r}, \tag{7}$$

$$\frac{\partial p}{\partial r} = -\frac{\mathrm{j}kA e^{\mathrm{j}(\omega t - kr)}}{r}\left(1 + \frac{1}{\mathrm{j}kr}\right) = -\mathrm{j}kp(r)\left(1 + \frac{1}{\mathrm{j}kr}\right). \tag{8}$$

It can be seen that compared with the plane wave case, the pressure gradient is increased by a factor of $\sqrt{1 + 1/(kr)^2}$. The resulting boost at low frequencies is substantial if the microphone is close to the source (the "proximity effect"), as for instance when a singer's mouth is a few centimeters from the microphone. Figure 1(b) shows the response of the dynamic pressure difference microphone whose plane wave response is shown in curve 3 of Figure 1(a) with a distance to the source of 5 cm (curve 1), 15 cm (curve 2), and 200 cm (curve 3). At 2 m distance, the response is practically identical with the response in a plane wave.

3 • The Directional Response of a Pressure Difference Microphone

The response of a dynamic pressure difference microphone to an incident plane wave at an angle θ is found by replacing Δx with $\Delta x \cos\theta$. This gives a simple cosine dependence ("figure-of-eight" directionality) in the frequency range where Δx is much shorter than the wavelength.

4 • The Ribbon Microphone

Ribbon microphones are dynamic pressure difference microphones in which a ribbon acts both as the diaphragm and as the induced conductor in a magnetic field. Both sides of the ribbon are exposed to the sound field, and thus, the microphone responds to the gradient of the sound pressure. This microphone principle was described in 1924 [1] and realized in 1931 [2]. A low resonance frequency is achieved with a ribbon of thin metal foil at a very low tension; see Figure 2. This mechanical design leads to a very short acoustic path difference between the front and the back of the ribbon and consequently to a rather weak driving force; hence, the ribbon must be very light. Care should be taken to ensure that higher order modes in the ribbon do not influence the response within the working frequency range of the microphone. Because the ribbon is far shorter than the wire length of a coil, ribbon microphones usually include a transformer in order to achieve a reasonable output voltage.

The microphone shown in Figure 3 is symmetrical about the vertical axis. It uses a helical yoke *Y* and the ribbon *R* can be seen between the pole pieces *P*. The frequency response is very flat, and the directional response is very close to the perfect figure-of-eight, although a small asymmetry causes a marginal difference above 10 kHz. This is an example of a microphone that has been in production for decades.

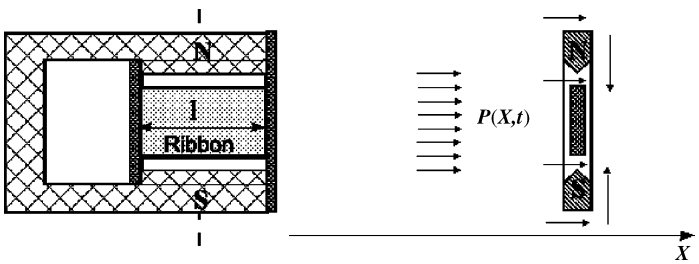


FIGURE 2 Basic design of a ribbon microphone.

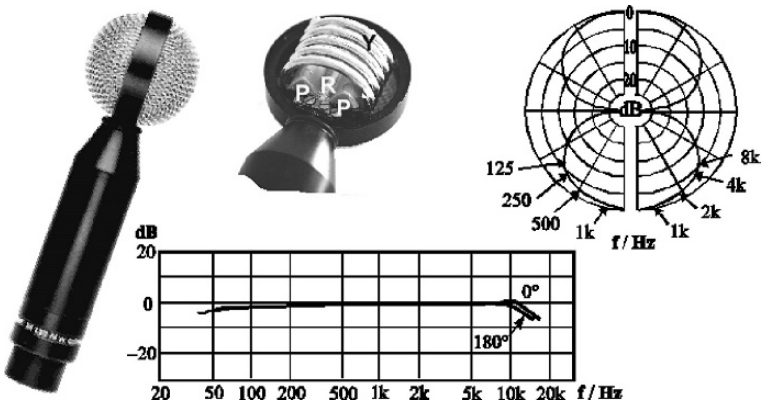


FIGURE 3 Bidirectional ribbon microphone (courtesy Beyerdynamic).

5 • Cardioids and Other Directivity Patterns

The symmetrical figure-of-eight directivity is favorable if two sources face each other and the unwanted background noise is diffuse or arriving from the sides. However, other directivity patters can be useful, for example, for a singer with a monitor loudspeaker in front, requiring a microphone with a maximum attenuation at 180°.

If one side of the diaphragm of a dynamic pressure difference microphone is directly exposed to the sound field whereas the other side is exposed to sound via an extra acoustic path, the resulting phase shift between the pressures on the two sides of the diaphragm will change the directivity. In this way, directivity varying from omnidirectional to figure-of-eight can be obtained. The characteristics shown in Figure 4 are specified in international standards [3]. They can also be achieved by adding the signal from a pressure-sensitive omnidirectional microphone to the signal from a microphone with a figure-of-eight directivity. By adjusting the proportions in the sum, it is possible to vary

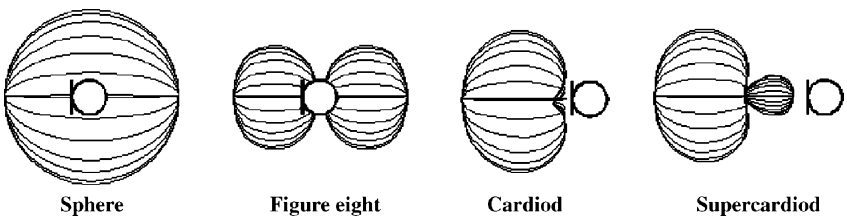


FIGURE 4 Examples of polar responses in wire mesh representation.

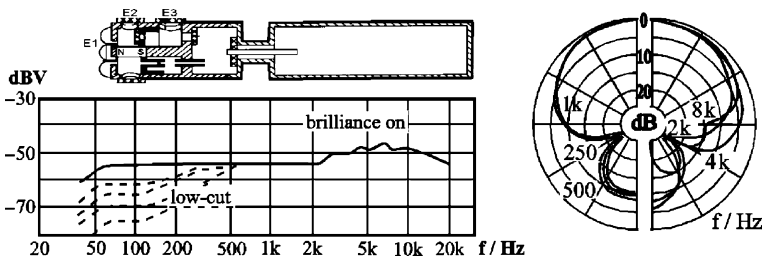


FIGURE 5 Moving coil microphone with supercardioid directivity.

the directional behavior as desired. Microphones with adjustable directivity are mainly top class condenser types.

It is also possible to extend the usable frequency range with frequency depending delay and phase shift. The "variable *d*" principle is one example [4].

Figure 5 shows an example of a successful moving coil microphone from 1971 (which is still in production) [5]. Despite an outer cross section of only 33×36 mm, a supercardioid with high sensitivity (1.8 mV per Pa) has been realized in the frequency range from 30 Hz to 20 kHz. The acoustic design includes several resonator chambers, two back entries, and a membrane resonance frequency of less than 100 Hz.

The interior of the microphone is shock isolated from the housing with two damped metal springs *S* between the back volume *V*, which contains additional passive circuitry for equalization, including a low-cut filter with five different responses and a two-position "brilliance" setting; see Figure 6. The response shown in Figure 5 has been measured with "brilliance" on.

6 • Interference Microphones

An extremely high directivity can be obtained with an interference tube in front of the microphone diaphragm. Early laboratory versions had a bundle of tubes with different lengths in front of the diaphragm, but today only one tube is used. An example of such a microphone is shown in Figure 7. Sound waves emanating from the holes along the tube interfere with each other before exciting the transducer at the end [6]. If the sound incidence is off axis then the interference is partly destructive.

Interference tube microphones are widely used for radio and TV production as well as for cinema sound recording. Because of their very small recording angle, experience is required for a user to obtain optimum results. If the microphone is not aimed properly, the level is reduced and the response spectrum

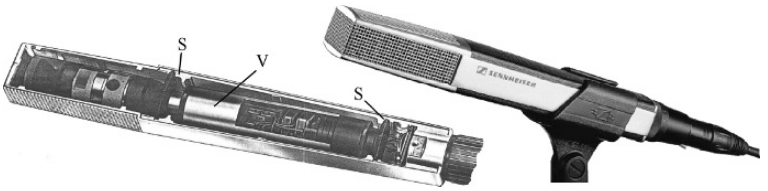


FIGURE 6 Small moving coil supercardioid microphone (courtesy Sennheiser).

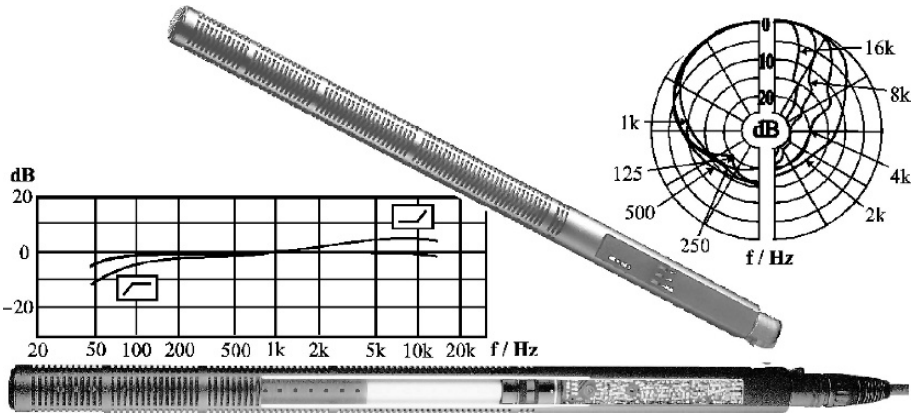


FIGURE 7 Modern interference microphone (courtesy Sennheiser).

is colored. For this reason, optical means are sometimes used for directing the device, for example when the microphone is operated from a crane.

7 • Reflectors

A special solution is to combine a parabolic reflector with a regular pressure microphone placed in its focus. Due to the limited dimensions, the directivity is poor up to medium audio frequencies [7]. Such microphones have been used for instance for recording singing birds.

8 • Fixed and Variable Arrays

High directionality can be obtained with microphone arrays [Ref: Handbook Chapter 9]. The countless possibilities offered by modern signal processing include, for example, signal-controlled source detection [8].

References

1. W. Reichardt: *Grundlagen der Elektroakustik*, 3. Ed., Leipzig, 1960, p.405
2. H.F. Olson: Mass controlled electrodynamic microphones: The ribbon microphone, J. Acoust. Soc. Am. **3**, 56 (1931)
3. IEC 60268-4 (2004–02) Sound system equipment — Part 4: Microphones (IEC, Geneva, 2004)
4. L. Burroughs: *Microphones: Design and Application* (Sagamore Publishing, Plainview, NewYork, 1974)
5. E. Werner: Ein neues dynamisches Studio-Richtmikrofon, Fernseh- und Kinotechnik **4**, 127–129 (1971)
6. W.P. Mason and R.N. Marshall: A tubular directional microphone, J. Acoust. Soc. Am. **10**, 206–215 (1939)
7. H.F. Olson: *Acoustical Engineering* (D. Van Nostrand Company, Princeton, 1957)
8. G.W. Elko and G.V. Frisk: 50 years of progress in microphone arrays for speech processing, J. Acoust. Soc. Am. **116**, 2498–2499 (2004)

68

The Microflown Particle Velocity Sensor

Finn Jacobsen[1] *and Hans-Elias de Bree*[2]

[1]*Technical University of Denmark, Lyngby, Denmark, e-mail: fja@elektro.dtu.dk*
[2]*Microflown Technologies, The Netherlands, USA*

1 • Introduction

The phenomenon of sound involves molecules of a fluid moving back and forth in the direction of propagation (with no net flow), accompanied by changes in the pressure, density, and temperature. The density and pressure variations are directly proportional to the sound pressure [1]. By contrast, the particle velocity is proportional to the *gradient* of the sound pressure, as can be seen from Euler's equation of motion,

$$\nabla p + \rho_0 \frac{\partial \mathbf{u}}{\partial t} = 0, \tag{1}$$

where ∇p is the pressure gradient, ρ_0 is the equilibrium density of the medium, and $\mathbf{u}$ is the particle velocity (which is a vector). In the general case, there is independent information in the sound pressure and in the particle velocity at a given point in a sound field. As an example, it can be mentioned that both quantities are needed in measurement of the sound intensity,

$$\mathbf{I} = \langle p(t)\mathbf{u}(t) \rangle_t \tag{2}$$

where $\langle\ \rangle_t$ indicates averaging with respect to time [2]. The sound intensity is a vector defined as the time average of the net flow of sound energy per unit area.

Condenser microphones for measurement of the sound pressure are more stable and reliable than any other acoustic transducer. By comparison, it has hitherto been difficult to measure the particle velocity in air directly. The only possibility for direct measurement of this quantity has been to use hot-wire anemometers, as e.g. in [3]. Hot-wire anemometers measure the change in resistance of a heated wire due to its cooling by a flow of air, but they cannot distinguish between flows coming in opposite directions, and thus a steady bias mean air velocity is needed for such devices to give a signal that represents the particle velocity [4]. The result of this bias flow is a poor signal-to-noise ratio. Alternatively, one has relied on pressure gradient microphones [see Chapter 67] or finite difference approximations to the pressure gradient, making use of (1). Recording microphones often measure a combination of the pressure and the pressure gradient, but the design criterion of such microphones is to obtain a desired directionality and frequency response rather than a reliable particle velocity signal [see Chapter 67].

In measurement of sound intensity with the established "two-microphone" technique, two closely spaced, phase-matched pressure microphones are used, the pressure is determined from the sum of the two signals, and the particle velocity is determined from the difference signal, which approximates the axial component of the gradient of the pressure [2] [see Chapter 58]. Sound intensity measurement systems of this type do not actually determine the particle velocity itself, though, and therefore the amplitude match of the two microphones is not critical. For other applications, the particle velocity signal itself is of concern. A trustworthy particle velocity signal may also be determined from a finite difference approximation, but then the microphones must be phase *and* amplitude matched very well [5]. Another problem of the finite difference approach is that the signal-to-noise ratio of the resulting particle velocity signal compared with single-microphone measurement of the sound pressure because of the time-integration that follows from (1) is reduced by $10\log(2/(k\Delta r)^2)$ dB [5], where k is the wavenumber and Δr is the microphone separation distance, and thus is rather poor at low frequencies (the finite difference approximation is valid only if $k\Delta r < 1$).

2 • The Microflown Transducer

The Microflown measures the particle velocity directly. This transducer is a micromachined hot wire anemometer, but based on *two* wires, not one as in the classical anemometer; see Figure 1. The two short, thin, closely spaced wires of silicon nitride coated with platinum are heated by a DC current to about 300°C [6–8]. Their resistance depends on the temperature. A particle velocity signal in the perpendicular direction changes the temperature distribution instantaneously, because the upstream wire is cooled more than the downstream wire by the airflow. The resulting resistance difference is measured with a bridge circuit that provides a signal proportional to the particle velocity.

The frequency response of the Microflown relative to a pressure microphone is shown in Figure 2. The frequency response has been measured in a large anechoic room far from a small loudspeaker [9]. At low frequencies, the sensitivity of this particle velocity transducer increases 6 dB per octave. The behavior at very low frequencies is believed to be related to the thermal boundary layer on the wires. Between 100 Hz and 1 kHz the frequency response is relatively flat. Between 1 and 10 kHz, there is a rolloff of 6 dB per octave, caused by diffusion effect related with the distance between the two wires (about 40 μm). Above 10 kHz, the sensitivity decreases an additional 6 dB per octave because of the thermal heat capacity of the wires.

The transducer is combined with a small electret condenser microphone in the ½-in. sound intensity probe shown in Figure 3. The velocity transducer is mounted on a small, solid cylinder, and

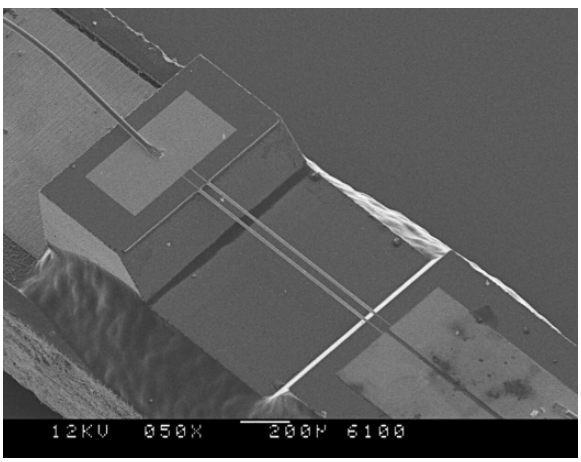


FIGURE 1 The two wires of the Microflown particle velocity transducer.

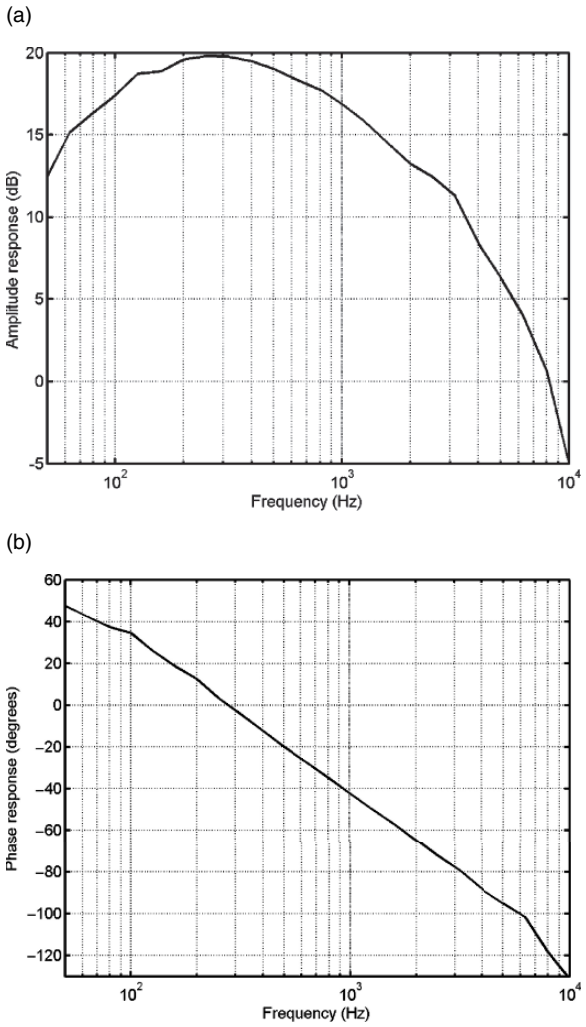


FIGURE 2 Amplitude and phase response relative to a pressure microphone.

the condenser microphone is mounted inside another, hollow cylinder. The geometry of this arrangement increases the sensitivity of the velocity transducer by about 10 dB.

3 • Limitations

The Microflown particle velocity transducer is noisier than standard condenser microphones of comparable sensitivity, but less noisy than a finite difference estimate of the particle velocity in most of the frequency range. Figure 4 shows the equivalent particle

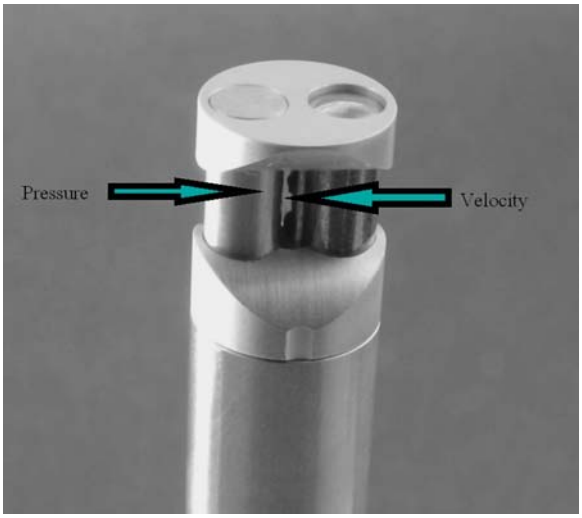


FIGURE 3 A $\frac{1}{2}$-in. $p-u$ sound intensity probe produced by Microflown.

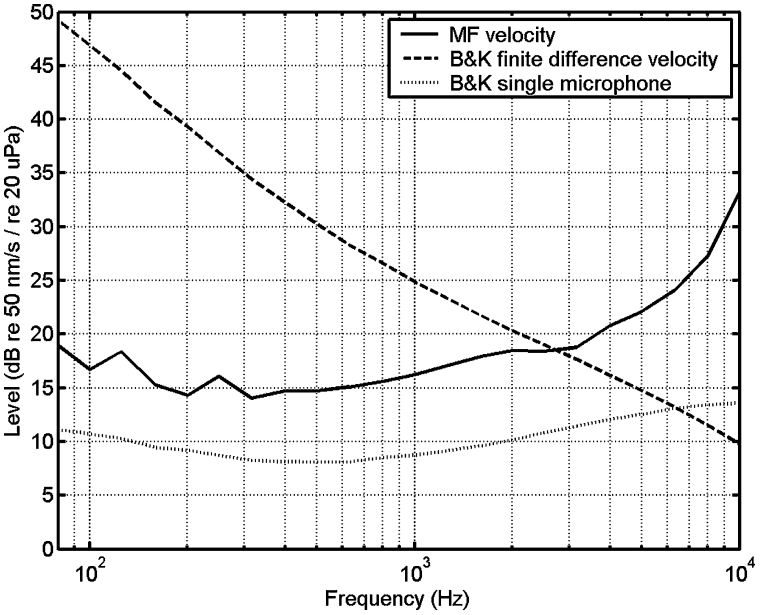


FIGURE 4 Electrical noise of the Microflown in one-third octave bands compared with the noise in the particle velocity measured with a two-microphone Brüel & Kjær sound intensity probe with a 12-mm spacer, and with the noise of a single pressure microphone of type B&K 4181.

velocity level of the noise of the velocity channel of the Microflown. The noise of a particle velocity signal measured with a Brüel & Kjær sound intensity probe of type 3599 with a 12-mm spacer between the two microphones is also shown. Note that the levels are shown relative to a reference velocity of 50 nm/s so as to make them directly comparable with the sound pressure level. The equivalent sound pressure level of the noise of a single $^1/_2$-in. Brüel & Kjær sound intensity microphone of type 4181 is also shown.

The upper limit of the dynamic range of the Microflown is limited by the electronics to about 125 dB re 50 nm/s, but the transducer itself can handle about 10 dB higher levels.

Any particle velocity transducer is more sensitive to unsteady airflow than pressure microphones are because the resulting velocity cannot be distinguished from the velocity associated with the sound waves, and the Microflown is no exception. The remedy is a windscreen of porous foam. However, the Microflown is not suitable for measurements under conditions with mean airflow of more than a few meters per second.

4 • Calibration

Calibration of the transducer is less straightforward than calibration of condenser microphones, and correction of the amplitude and, for some applications, phase response is essential, as demonstrated by Figure 2. One cannot calibrate a particle velocity transducer in a small coupler; the device must be exposed to a sound field with a known relation between the sound pressure and the particle velocity. There are no standardized calibration procedures [2]. However, recent experimental results seem to indicate the possibility of calibrating the particle velocity channel of the Microflown intensity probe in the far field of a loudspeaker in an anechoic room, in the near field of a small source (sound emitted from a small hole in a spherical baffle), and – at low frequencies – in a standing wave tube with a rigid termination [10].

5 • Applications

There is as yet relatively little documented experimental evidence of the practical utility of the Microflown particle velocity transducer, but measurement of the acoustic particle velocity has many applications, although most of them still are at a relatively early stage of development because of the absence of reliable particle velocity transducers until recently. Some applications are mentioned in what follows.

The idea of measuring total acoustic energy density (three components of the particle velocity *and* the pressure) rather than just potential energy density (sound pressure) in a sound field goes back to the 1930s [11]. One advantage is that the total energy density tends to vary less with the position in a reverberant enclosure than the potential energy density [11,12]. Sensors for the total acoustic energy have also been found to be useful in active noise control systems, because minimizing the total energy density can be a more efficient control strategy than minimizing the sound pressure [13]. Because of the directional information, particle velocity measurements have found application in room acoustics [14], as already suggested in [11]. Yet another application is in analyzing complicated sources of noise, where reciprocity experiments for characterizing machinery noise sources are very useful [15,16]. Other applications include measurement of the impedance and absorption coefficient of materials in situ, and measurement of the specific acoustic impedance in tubes [3,17].

The Microflown sound intensity probe seems to have potential in measurement of sound power [9]. Figure 5 shows the results of sound power measurements with the Microflown intensity probe

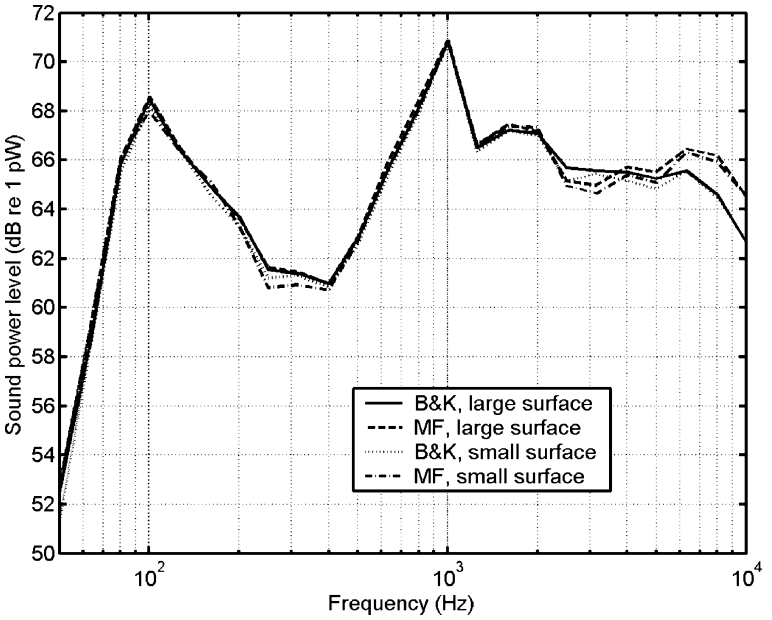


FIGURE 5 Sound power of a source in one-third octave bands measured using the Microflown intensity probe and a two-microphone intensity probe produced by Brüel & Kjær.

shown in Figure 3 in comparison with similar results made with a Brüel & Kjær two-microphone intensity probe of type 3599. The Microflown intensity probe is very small — much smaller than a standard two-microphone intensity probe. Thus, it is possible to measure very close to a vibrating surface with this device. This is advantageous in several applications of sound intensity, for example, in measurement of the radiation factors of sources.

Advanced source identification methods such as near-field acoustic holography and inverse boundary element methods involve solving "ill-posed" inverse problems, and recent results show that such methods tend to become more robust if they are based on particle velocity data than on sound pressure data [18–20].

References

1. A.D. Pierce: *Acoustics: An Introduction to Its Physical Principles and Applications*, 2nd ed. (Acoustical Society America, New York, 1989)
2. F.J. Fahy: *Sound Intensity*, 2nd ed. (E & FN Spon, London, 1995)
3. R.L. Pratt, S.J. Elliott, J.M. Bowsher: The measurement of the acoustic impedance of brass instruments, Acustica **38**, 236–246 (1977)
4. H.H. Bruun: Interpretation of hot-wire probe signals in subsonic airflows, J. Phys. E Sci. Inst. **12**, 1116–1128 (1979)
5. F. Jacobsen: A note on finite difference estimation of acoustic particle velocity, J. Sound Vib. **256**, 849–859 (2002)
6. H.-E. de Bree, P. Leussink, T. Korthorst, H. Jansen, T. Lammerink, M. Elwensoek: The Microflown: A novel device measuring acoustical flows, Sensor. Actuator **SNA054/1-3**, 552–557 (1996)
7. H.-E. de Bree: The Microflown: An acoustic particle velocity sensor, Acoust. Aust. **31**, 91–94 (2003)
8. R. Raangs, W.F. Druyvesteyn, H.-E. de Bree: A low-cost intensity probe, J. Audio Eng. Soc. **51**, 344–357 (2003)
9. F. Jacobsen, H.-E. de Bree: A comparison of two different sound intensity measurement principles, J. Acoust. Soc. Am. **118**, 1510–1517 (2005)
10. F. Jacobsen, and V. Jand: A note on the calibration of pressure-velocity sound intensity probes, J. Acoust. Soc. Am. **120**, 830–837 (2006)
11. I. Wolff, F. Massa: Use of pressure-gradient microphones for acoustical measurements, J. Acoust. Soc. Am. **4**, 217–234 (1933)
12. R.K. Cook, P. Schade: New method of measuring the total energy density of sound waves, in *Proceedings of Inter-Noise 74*, Washington, DC, pp. 101–106 (1974)
13. S.D. Sommerfeldt, P.J. Nashif: A comparison of control strategies for minimizing the sound field in enclosures, in *Proceedings of Noise-Con 91*, Tarrytown, New York, pp. 299–305 (1991)

14. M. Barron: *Auditorium Acoustics and Architectural Design* (E & FN Spon, London, 1993)
15. J.W. Verheij: Inverse and reciprocity methods for machinery noise source characterization and sound path quantification, Int. J. Acoust. Vib. **2**, 103–112 (1997)
16. F. de Beer, J.W. Verheij: High power sound source with large coverage angle for use in reciprocity experiments, in *Proceedings of Seventh International Congresson Sound and Vibration*, Garmisch-Parten-Kirchen, Germany, pp. 3235–3242 (2000)
17. Y. Liu, F. Jacobsen: Measurement of absorption with a $p-u$ sound intensity probe in an impedance tube, J. Acoust. Soc. Am. **118**, 2117–2120 (2005)
18. R. Visser: A boundary element approach to acoustic radiation and source identification (PhD thesis, University of Twente, Enschede, The Netherlands, 2004)
19. F. Jacobsen, Y. Liu: Near field acoustic holography with particle velocity transducers, J. Acoust. Soc. Am. **118**(5), 3139–3144 (2005)
20. F. Jacobsen and V. Jand: Statistically optimized near field acoustic holography using an array of pressure-velocity probes, J. Acoust. Soc. Am. **121**, 1550–1558 (2007)

69

Microphone Calibration

Erling Frederiksen

Brüel & Kjær, Sound and Vibration Measurement A/S, Skodsborgvej 307, 2850 Nærum, Denmark

1 • Introduction

Sound measurements in air are performed over a very wide range of frequencies that goes from infrasound to ultrasound or, say, from a fraction of a hertz to about 200 kHz. Sound is also measured over a wide dynamic range that begins below the human threshold of hearing, below 20 μPa or 0 dB, and goes to more than 20 kPa or 180 dB. Because of these wide ranges and the different types of sound field that occur, there is a need for many different models of microphone and several different calibration methods. This chapter describes the methods that are most commonly applied and generally standardized by the International Electrotechnical Commission (IEC). The less common methods are mentioned in brief.

2 • Calibrations Performed at Primary and Secondary Levels

Most commonly, a microphone is calibrated by comparing its response with that of a previously calibrated reference microphone. The two microphones are exposed to the same sound pressure, either by simultaneous exposure or by exposing one by one (substitution). A comparison calibration is called a secondary calibration. However, the reference microphone, or the one that is at the very top of the calibration hierarchy, cannot be compared

with a higher-ranking microphone. It must be calibrated with reference to non-acoustic standards, i.e., by a primary calibration method that refers to electrical, mechanical and physical parameters [1]. Primary methods are generally more complex and time-consuming than secondary methods. Primary standard microphones must be calibrated with the highest possible precision, as the uncertainty associated with their calibration influences all calibrations made at lower levels in the measurement chain.

3 • Principal Types of Sound Field and Microphone Responses

There are three principal types of sound field and three corresponding types of microphone frequency response. The three types of sound field are pressure field, free field and diffuse field [2]. The corresponding sensitivities [3] or frequency responses of a given microphone are essentially equal from say 20 to 500–2,000 Hz (depending on its dimensions), but the three responses can be very different at higher frequencies; see Figure 1. Therefore, for any particular application and type of sound field, it is important to select a proper type of microphone and a proper method of calibration. The standard IEC50(801) defines the types of sound field, while a series of standards IEC61094 [4] describes the various relevant calibration methods.

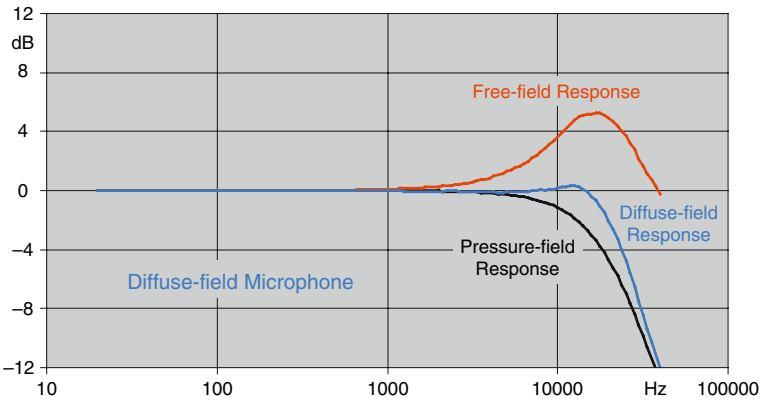


FIGURE 1 Pressure-field, free-field and diffuse-field frequency responses of a half-inch measurement microphone that is optimized for use in diffuse sound fields.

4 • Laboratory Standard Microphones Suited for Primary Calibration

Most measurement microphones are cylindrically shaped condenser microphones with good acoustic properties, such as high stability of their sensitivity and an influence of ambient conditions [5] that is relatively small and predictable. The simple cylindrical shape makes it possible not only to measure but also to calculate several acoustic properties, like diaphragm impedance, sound diffraction/reflection and other parameters that are important in connection with calibration. Another property of the condenser microphone, which is of particular importance for calibration, is that it is a reciprocal transducer [6]. This means that it can be operated both as a receiver and as a transmitter of sound and that it can be calibrated by reciprocity calibration technique; see Figure 2. Today this is the technique that is most commonly applied for primary pressure-field and free-field calibration. The standard IEC61094-1 [3] specifies microphones that are suited for primary calibration and for use as (national) reference standards. They are designated LS1 and LS2 and their diameters are 23.77 and 12.7 mm, respectively.

5 • Principle of Microphone Reciprocity Calibration

Calibration of microphones by reciprocity technique was invented in the 1940s [7,8]. Since then the method has become refined and standardized and is now the dominating primary calibration technique for determining both pressure-field and free-field responses. The methods that are quite complex

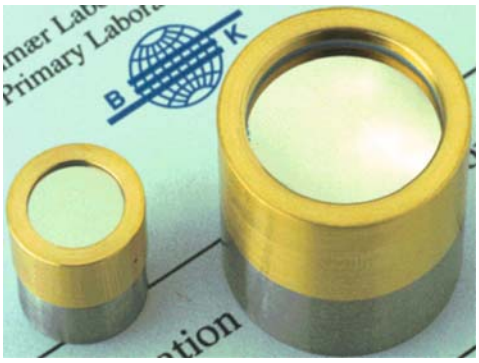


FIGURE 2 Laboratory standard microphones B&K Types 4160 and 4180. These microphones meet the requirements of IEC 61094-1, LS1p and LS2ap, respectively. Such microphones are suited for reciprocity calibration and applied as national standards.

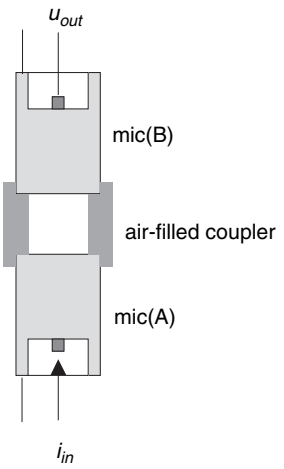


FIGURE 3 Principle of pressure-field coupling applied for reciprocity calibration of microphone pressure sensitivity.

and time consuming to work with are mainly applied by national metrology institutes (NMIs) and by leading microphone manufacturers.

Reciprocity calibration is based on the measurement of the transfer function between two coupled microphones that are operated as a source and as a sensor. The microphones are coupled together in an acoustically well-defined way, while the over-all transfer function, the ratio between sensor output voltage and source input current, is measured; Figures 3 and 4. From this ratio, called the electrical transfer impedance (Z_e), and from the acoustic coupling represented by the acoustic transfer impedance (Z_a), the product of the microphone sensitivities can be determined:

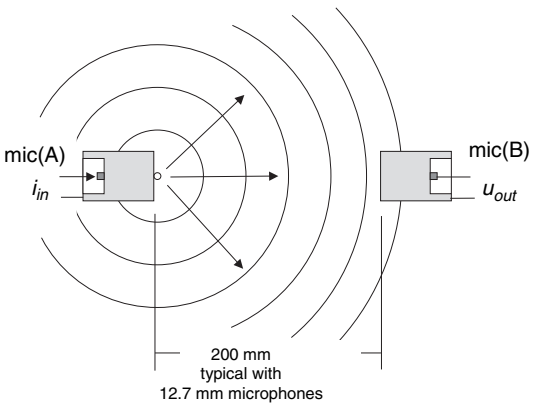


FIGURE 4 Principle of free-field coupling applied for reciprocity calibration of microphone free-field sensitivity.

$$M_1 \cdot M_2 = \left(\frac{Z_e}{Z_a}\right)_A \tag{1}$$

where

M_1 and M_2 are the sensitivities of microphones 1 and 2
Z_e/Z_a is the ratio of electrical and acoustic transfer impedance

By using three microphones (1,2,3) and by making impedance ratio determinations (A, B, C) for the three possible combinations of microphone (1–2, 1–3, 2–3), the sensitivities of all microphones can be calculated by the equations below:

$$M_1 \cdot M_2 = \left(\frac{Z_e}{Z_a}\right)_A \quad M_1 \cdot M_3 = \left(\frac{Z_e}{Z_a}\right)_B \quad M_2 \cdot M_3 = \left(\frac{Z_e}{Z_a}\right)_C \tag{2}$$

Either free-field or pressure sensitivities can be obtained by selection of microphone coupling method. The pressure-field responses are thus obtained by coupling the microphones with the air (or gas) contained in a small closed cavity, while the free-field responses are obtained with an open-air coupling in a room that has no disturbing sound-reflecting surfaces (in an anechoic room).

6 • Primary Pressure Response Calibration by Reciprocity Technique

Several NMIs around the world perform pressure reciprocity calibration [9] of laboratory standard microphones. Typically, the frequency range is from 20 to 10 kHz for "one-inch" (LS1) and 20 to 20 kHz for "half-inch" (LS2) microphones, but some institutes have experience with calibration at both lower and higher frequencies. Standing waves inside the applied acoustic couplers set the upper frequency limit of pressure reciprocity calibration. Wave motion problems may, in practice, be reduced either by filling the coupler cavity with hydrogen, which has a higher speed of sound than air, or by shaping the cavity in a way that leads to a simple wave pattern that facilitates a precise calculation of the acoustic transfer impedance. The latter is valid for a cylindrical cavity with the diameter equal to that of the microphone diaphragms, which close the cavity at the ends. The cavity volume of such couplers is typically 4–8 cm^3 for LS1 and 0.3–0.7 cm^3 for LS2 microphones. These couplers, which are called plane-wave

TABLE 1 Typical uncertainty of pressure reciprocity calibrations performed by an experienced NMI

Uncertainty ($k = 2$)		20 Hz	32 Hz	63 Hz–4 kHz	10 kHz	20 kHz	25 kHz
Pressure-field	LS1	0.06 dB	0.04 dB	0.03 dB	0.08 dB	–	–
	LS2	0.08 dB	0.05 dB	0.04 dB	0.04 dB	0.08 dB	0.12 dB

couplers, can work to high frequencies without hydrogen. The calculation of their transfer impedance is based on the theory of transmission lines. The standard IEC61094-2 [10] describes the method. Table 1 shows the typical calibration uncertainty achieved by an experienced NMI.

The microphone sensitivity is a function of the ambient pressure. This pressure dependence can be measured by placing the microphones and the calibration couplers in a measurement chamber that is pressurized during the reciprocity calibration; see Figure 5.



FIGURE 5 Measurement chamber of B&K Microphone Reciprocity Calibration System Type 9699 mounted with microphones and coupler for pressure reciprocity calibration. The chamber that can be pressurized is intended for measurement of microphone sensitivity as a function of ambient pressure.

7 • Primary Free-Field Response Calibration by Reciprocity Technique

Only a few laboratories in the world perform free-field reciprocity calibration [11]. The theory behind the determination of the acoustic transfer impedance is simpler than that of pressure reciprocity calibration, but the practical measurement of the electrical transfer impedance is technically more difficult due to the very small microphone output voltages that have to be measured. The reason is that condenser microphones are very weak sound sources, when applied in open space; see Figure 6. This often leads to serious measurement problems that are related to ambient or inherent noise and to electrical cross-talk in the measurement system. As this is especially the case at lower frequencies, the lower limit of free-field reciprocity calibration systems is typically between 0.8 and 2 kHz. Free-field calibration is performed up to about 25 kHz and to 50 kHz for one-inch (LS1) and half-inch (LS2) microphones, respectively. The standard IEC 61094-3 [12] describes the method. Table 2 shows typical calibration uncertainty values.

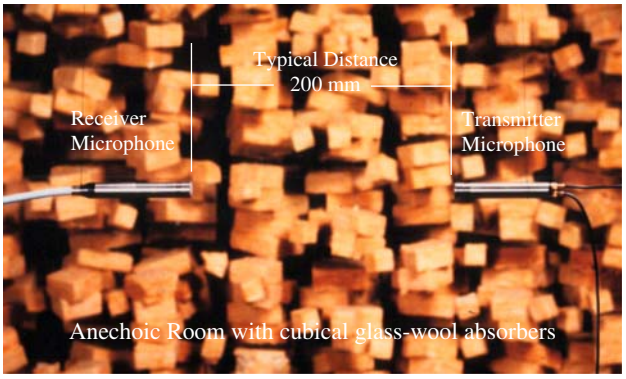


FIGURE 6 Reciprocity calibration setup in an anechoic room for determination of free-microphone field response.

TABLE 2 Typical uncertainty of free-field reciprocity calibrations performed by an experienced NMI

Uncertainty ($k = 2$)		1 kHz	2 kHz	4 kHz	8 kHz	10 kHz	20 kHz	40 kHz
Free field	LS1	0.10 dB	0.08 dB	0.07 dB	0.07 dB	0.08 dB	0.15 dB	–
	LS2	0.12 dB	0.10 dB	0.08 dB	0.07 dB	0.07 dB	0.09 dB	0.15 dB

8 • Sound Field Corrections of Laboratory Standard and Measurement Microphones

The ratio (or the difference in decibels) between the free-field and the pressure-field responses of a microphone is predominantly determined by the outer shape and by the dimensions of its body, which diffracts and reflects the impinging sound wave in a free sound field. As all microphones of a specific model have the same shape and the same dimensions, there is essentially the same ratio between their free-field and pressure-field sensitivities. In the case of models, where this ratio has been determined and published, the free-field response of an individual microphone can be determined by adding a correction to its pressure-field response, which is generally more easily and more accurately measured. Such corrections, which are functions of frequency and angle of sound incidence on the microphone body, are called free-field corrections. They are available for many models of measurement and laboratory standard microphone; see the two examples shown in Figure 7. IEC has issued a technical report IEC61094-7 [13] with free-field corrections for laboratory standard microphones.

9 • Determination of Diffuse-Field Frequency Response

Diffuse-field reciprocity calibration [14] is, in principle, possible, but the method is not sufficiently refined for practical application. Therefore, diffuse-field responses of microphones are

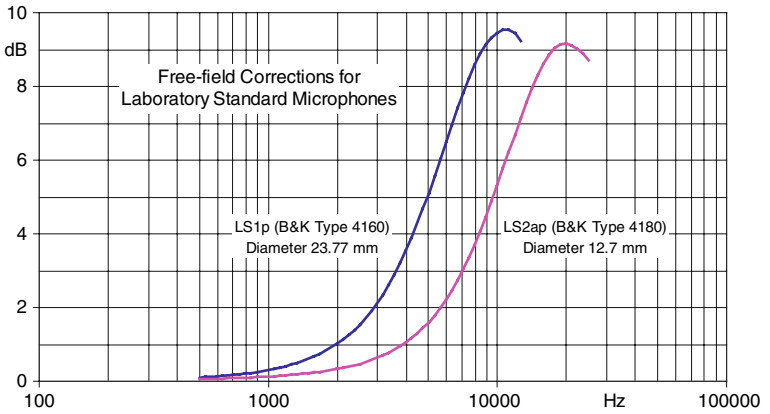


FIGURE 7 Free-field corrections valid for perpendicular sound incidence (0°) on the diaphragms of B&K Type 4160 (LS1) and Type 4180 (LS2). The corrections are to be added to the pressure response to obtain the free-field response.

most often calculated from a series of free-field responses that are measured at selected angles of sound incidence on the microphone body. The angles are typically spaced by 10° and cover the range from 0° to 180°. The calculation of the diffuse-field response, which usually considers the microphone to be rotationally symmetrical, is a power-based summation of the responses that are measured for the selected angles. Different weighting factors are applied for the measured responses, as they represent ranges of solid angles that are not equally large. Thus more weight has to be put on the response of 90° incidence than on that of any other angle.

This method of calculating the diffuse-field response from a series of free-field responses or, alternatively, the diffuse-field correction from a series of free-field corrections is described in the international standard IEC61183 [15]. The method is commonly applied for determining diffuse-field responses of both microphones and sound level meters. A measured series of free-field corrections and the calculated diffuse-field correction that are valid for a half-inch microphone are shown in Figure 8.

The diffuse-field response of a larger microphone might also be measured by comparing its response with a much smaller pressure-field-calibrated reference microphone. If the reference microphone is sufficiently small, the difference between its pressure-field and diffuse-field response may be so small that it, in some cases, may be neglected. The measurement may be made with a practically obtainable, yet not fully perfect, diffuse sound field. During the comparison measurements the microphones should be interchanged between fixed positions in the

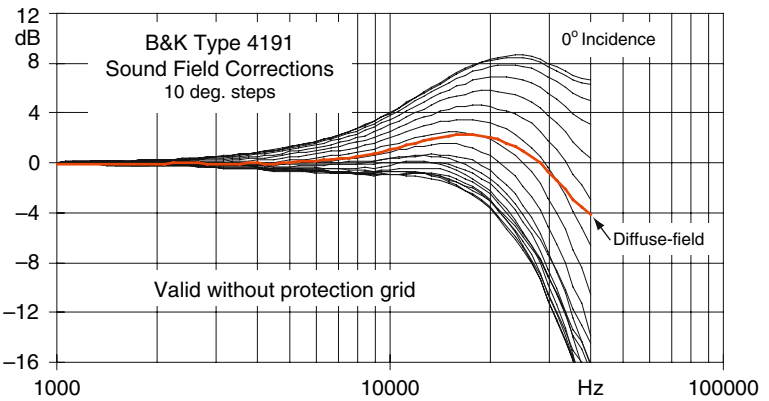


FIGURE 8 Diffuse-field correction for typical half-inch microphone model calculated from free-field corrections measured at angles stepped by 10°.

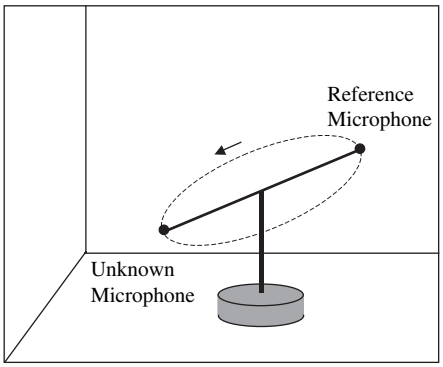


FIGURE 9 Principle of using rotating boom in a reverberant room for diffuse-field comparison calibration of microphones.

sound field or, better, they should be mounted on moving arms that slowly pass them through the same circular orbit within the sound field of a reverberant room [16]; see Figure 9.

10 • Working Standard Microphones

In addition to the laboratory standard microphones mentioned above, another group of microphones has been standardized by IEC. They are named working standard (WS) microphones and are specified in the IEC61094-4 [17]. They are used as transfer standards for the calibration of all kinds of acoustic measurement instruments. Three types/sizes of microphone are specified, WS1, WS2 and WS3 with diameters of 23.77, 12.7 and 6.35 mm, respectively; see Figure 10. The laboratory standard microphones are operated with fully open or non-protected diaphragms, while the working standards, depending on the application, may be used either with or without their diaphragm protection grids. The



FIGURE 10 Typical models of WS microphone, whose frequency responses are optimized for use in pressure, free or diffuse sound fields.

grids make them more resistant to the frequent handling and use in an acoustic calibration service centre. The sensitivities of WS microphones are most often determined by comparison with laboratory standard microphones that are calibrated by the reciprocity technique.

11 • Sensitivity and Frequency Response Calibration by Comparison

The sensitivity and frequency response calibration of a measurement microphone may be performed either by making comparison measurements with a reference microphone at a longer series of frequencies or by using a combined technique that implies comparison at one frequency only and a separate measurement of the frequency response. This response can be measured by excitation with an electrostatic actuator (see below). In the latter case, the sensitivity is generally measured at a frequency that is so low that the pressure-field, free-field and diffuse-field sensitivities of the microphone are essentially equal. The frequency of 250 Hz is most commonly used, as it is also the operation frequency of several models of acoustic calibrator (pistonphones).

Pressure sensitivity comparison calibration of microphones is described in the international standard IEC61094-5 [18]. The standard proposes different methods for mounting of the two microphones to be compared. In all cases, the microphones are lined up, on-axis and opposite each other with their diaphragms placed "face to face" at a very short distance (typically 2–3 mm). One can use either couplers with built-in sound sources or microphone holding jigs with separate, nearby mounted sound sources. It is advantageous with these mountings that they can be applied over wide frequency ranges and that the influence of the microphones on the sound pressure is eliminated. The close mounting of the diaphragms ensures that the two microphones "hear" the same pressure. The expanded uncertainty ($k = 2$) [19] of the method may, below 1,000 Hz, be very low say, <0.03 dB. Comparison calibration of equally large microphones can, with good couplers, be made within 0.3 dB ($k = 2$) up to about 16 kHz. Free-field and diffuse-field frequency responses may be obtained by adding corrections to the measured response. It should be noticed that the sound field corrections that should be applied depend on the type of jig, the coupler or the mounting method and also that they differ from corresponding free-field and diffuse-field corrections valid for other

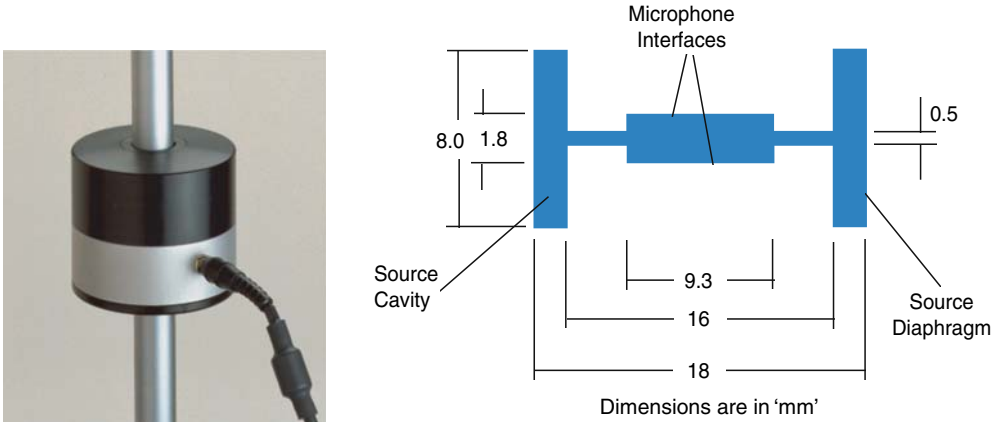


FIGURE 11 Photo of comparison calibration coupler (B&K, WA0817) for half-inch and smaller microphones beside a sketch of its cross-sectional view.

measurement methods. A typical coupler and cross-sectional view are shown in Figure 11.

12 • Frequency Response Calibration by Electrostatic Actuator

Ideally, the frequency response of a microphone should be measured with a sound pressure that is either known as a function of frequency or known to be constant over the frequency range of interest. These requirements are very difficult to meet for an acoustic source, but an electrostatic actuator can produce a force, or say an electrostatic pressure, directly on a microphone diaphragm that is independent of the frequency. The actuator can simulate a sound pressure on the microphone diaphragm, if this is electrically conducting. An actuator is most often a flat and very stiff metallic plate with holes or slits that make it acoustically transparent; see Figure 12. During the operation, the actuator is placed close to and in parallel with the microphone diaphragm. Typically, 800 V DC and 30 V AC are applied between the actuator and the diaphragm; see Figure 13. These voltages will, with the most commonly used models of actuators, generate an equivalent sound pressure of about 1 Pa (94 dB). The equivalent sound pressure may be calculated by the equation below:

$$p(t) = -\frac{\varepsilon_a}{2d^2}\left(U_0 + u\sin(\omega t)\right)^2 \tag{3}$$



FIGURE 12 Electrostatic actuator for frequency response calibration of one-inch microphones. Pressure-field, free-field and diffuse-field responses may be obtained by adding model-dependent corrections to the measured response.

where

$p(t)$	equivalent instantaneous pressure (Pa)
ε	dielectric constant of the surrounding air (8.85×10^{-12} F/m)
a	ratio between the areas of the perforated actuator and the diaphragm
d	distance between actuator and microphone diaphragm (m)
U_0	DC voltage applied between actuator and diaphragm (V)
u	peak value of sinusoidal AC voltage applied between actuator and diaphragm (V)
ω	angular frequency (rad/s)
t	time

In principle, the actuator method is a primary method, but the method is not used for absolute calibration of sensitivity, as the uncertainty of some of the pressure determining parameters (a, d) is, generally, too large.

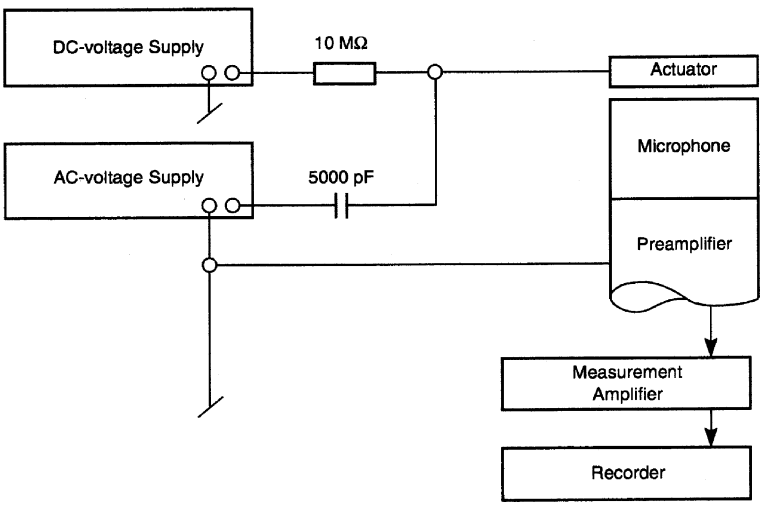


FIGURE 13 Driver circuit for electrostatic actuators.

The frequency-independent and uniform pressure that the actuator produces on the diaphragm leads to a response that is similar to the pressure response of the microphone. However, there are minor differences, whose magnitude depends on the type of microphone. They are due to the sound pressure that, during the electrostatic excitation, is produced by the movement of the microphone diaphragm itself. The deviation between pressure and actuator responses is, therefore, smaller for microphone models with low sensitivity and stiff diaphragms than for models with higher sensitivity and more flexible diaphragms. The deviation of the actuator response from the pressure response is for LS1 and WS1 microphones (50 mV/Pa, Ø23.77 mm): 0 dB at low frequencies, typically 0.1 dB at 1 kHz and less than 0.2 dB below 6 kHz, while it is about 1.5 dB at 10 kHz.

The actuator method is widely used by calibration service centres for determination of free-field, diffuse-field and pressure-field responses. These responses are obtained by adding corrections to measured actuator responses. The corrections, which depend on the type of microphone, are available from leading microphone manufacturers.

The actuator method is generally used between 20 Hz and 100 kHz, but it can be used at both higher and lower frequencies, though with low-frequency calibration (typically $f < 20$ Hz), attention should be paid to the fact that the actuator excites the microphone diaphragm only. This is in accordance with the definition of the microphone pressure response. At low frequencies, this is essentially flat and can, in principle, be measured to DC. When used in free and diffuse sound fields, the microphone is fully surrounded by the sound field. This leads to a roll-off of the microphone response at low frequencies (say below 5–10 Hz), because the pressure in the cavity behind the diaphragm is partly equalized with the surrounding pressure via the static pressure equalization vent of the microphone. As this phenomenon cannot be measured with an electrostatic actuator, the method cannot be applied for calibration of free-field and diffuse-field responses at frequencies lower than say 10 times the lower limiting frequency (−3 dB point) that is specified for the microphone.

The actuator method is relatively cheap to establish, as no specific acoustic laboratory facilities are required. The method is described in detail in the standard IEC61094-6 [20] and in the AIP Microphone Handbook [21].

13 • Free-Field Response Calibration by Free-Field Comparison

Free-field response determination by free-field comparison is a method that is commonly used by metrology institutes, test laboratories and manufacturers for testing of microphones and sound level meters. The comparison is, generally, made by substitution, using so large a distance between the sound source and the microphone position that the wave front is essentially flat where it passes the test object. The distance is typically 0.5 m, when calibrating a microphone, and larger, up to 2.0 m, in case of a bigger object, like a microphone mounted on a sound level meter. The method requires the same positioning of the unknown and the reference microphone as well as a source that stays stable between the measurements. An advantage of this method is that the shape and size of the calibrated microphone is not critical. However, lower uncertainty is, generally, obtained, when the reference and test microphones have the same dimensions. This minimizes errors caused by deviations between the generated sound wave and an ideal plane wave.

The method requires either an anechoic room or some free space in an ordinary, quiet room. In the latter case, a time-selective measurement technique must be used to separate the response of the microphone to the direct sound wave from that of the unwanted, later-arriving, reflected waves.

The use and accuracy of the free-field comparison methods will benefit from two new IEC standards. One of them states internationally agreed free-field corrections [13] for the laboratory standard microphones that are the commonly used references standards. The other one, presently an IEC draft, will describe the calibration method itself, including requirements to instruments, measurement site and positioning of the source and microphones.

14 • Traceability and Accreditation

International trade and communication make it increasingly important to exchange measurement results and have them accepted across country borders. Therefore, more and more calibrations are now performed with reference standards that are traceable back to NMIs. These institutes maintain national standards and ensure their international agreement by performing mutual comparison calibrations. Such calibrations are organized either by Bureau International des Poids et Mesures (BIPM) or by a regional organization. Charts, like the

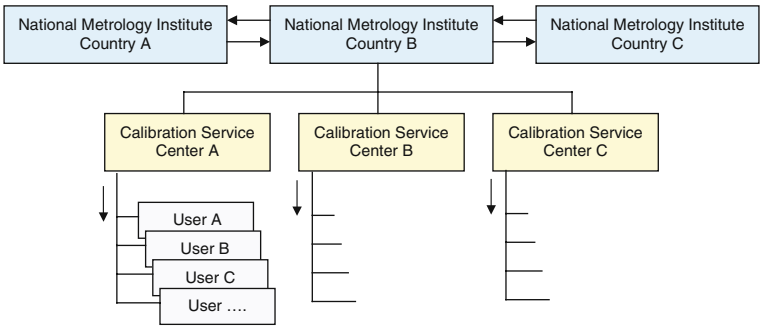


FIGURE 14 Typical calibration hierarchy with NMIs, service centres and microphone users. In some countries there is an extra level, a reference laboratory, between the NMI and the service centres.

one shown in Figure 14, are used to illustrate traceability of calibrations within their hierarchy.

In connection with legal metrology, traceability may not be enough. All laboratories that are links in a chain of the hierarchy must hold an accreditation that is obtained by proving to an official accreditation body that the laboratory follows the international standard ISO17025 [22] for good calibration laboratory practice. The accreditation requirements include a well-established quality system, a detailed analysis of possible calibration errors and calculations of the uncertainties that are associated with the applied methods.

IEC, ISO, BIPM, Organization Internationale de Métrologie Légale (OIML) and other international organizations have issued a common, very comprehensive document [19] about uncertainty calculation of measurements, while the European Cooperation for Accreditation (EA) has issued a shorter description that follows the same principles [23], but is more directly applicable.

15 • Dynamic Linearity

The dynamic range of condenser microphones is generally large compared with the commonly occurring range of measured sound levels. However, sometimes measurements must be made near to the lower limit – the level of the inherent noise of the microphone – or near to the upper limit, which is usually determined by microphone distortion or by signal clipping by its preamplifier.

Inherent noise of a microphone can be measured by placing the microphone in a sound-isolating chamber. This is typically a sealed metal cylinder with walls of 10–20 mm thickness and a

volume of 1,000–2,000 cm^3. Such a chamber can be used for noise measurements down to less than 0 dB (threshold of hearing), if it is acoustically sealed and isolated from vibration.

High-level calibration devices have been designed with horn-drivers, shakers, units for sudden DC pressure release, piston-phones and resonating tubes, which usually work at relatively low frequencies. Linearity of condenser microphones is measured with such systems by comparing their output with that of a ceramic or quartz high-level reference microphone. An example of a resonating tube system is shown in Figure 15. This system can generate a quiet pure tone at 500 Hz and measure level-linearity and distortion from 94 to 174 dB SPL (0.1 bar) with uncertainties less than 0.025 dB and 0.5%, respectively [24].



FIGURE 15 Sound chamber of Brüel & Kjær system Type 9719 for measurement of dynamic level linearity and distortion that can generate a pure tone at 500 Hz (<0.5 % distortion) over the entire range from 94 to 174 dB.

16 • Field Calibration with Sound Field Calibrators

On-site calibration of measurement systems is common and highly recommendable. A system can be either adjusted or calibrated by setting its parameters according to the sensitivity stated for the microphone or according to the sound pressure level stated for a connected sound level calibrator. In some cases either one of these methods may be sufficient, but a more reliable calibration and a better check of the functioning of the system is obtained by applying both methods. After setting of the microphone sensitivity and checking with the calibrator, there should be only minor disagreement, a few tenths of a decibel, between the methods. In general, the uncertainties of the two methods are about 0.15 – 0.30 dB. In case that the uncertainty of the calibrator method is estimated to be the lowest one, the system should finally be adjusted to show the level of the calibrator.

Specifications of sound calibrators are given in national and international standards [25]. Most types of sound calibrator are intended for calibration at one frequency only. Typically they generate a level between 94 and 124 dB at a frequency of either 250 or 1,000 Hz. The latter frequency is specified for calibration of sound level meters according to the IEC standard 61672-1 [26]. Generally, older types of calibrator contained just an electrical generator and a sound source, while newer types have a built-in microphone that measures and controls the generated pressure by a feedback loop. This makes the newer calibrators more stable and less dependent on ambient conditions and on the loading made up by the inserted microphone.

Pistonphones are calibrators that generate the sound by pistons that are driven by an electric motor and are moving into and out of a coupler cavity. Generation of a well-defined sound pressure implies a relatively high level and a relatively low frequency – most commonly 124 dB and 250 Hz. Previously pistonphones were widely used as field calibrators. Now they are mainly applied as reference standards by metrology institutes.

Multi-frequency calibrators are typically applied for frequency response verification of sound level meters and other measurement systems; an example is the B&K Calibrator Type 4226; see Figure 16. This calibrator applies a monitoring microphone and a feedback loop for controlling the sound pressure. The calibrator, which produces a pressure field, can verify free-field and diffuse-field responses of commonly used models of microphone by applying pre-installed sound level corrections. The calibrator works from 31.5 Hz to 16 kHz.

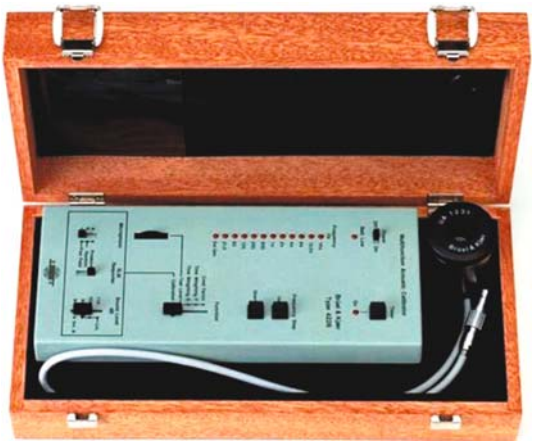


FIGURE 16
Multi-frequency calibrator (with feedback microphone) that is applied for laboratory and field verification of sound level meters and other measurement systems.

References

1. International Vocabulary of Basic and General Terms in Metrology, BIPM, IEC and others, primary and secondary standard, p. 46
2. IEC 50(801), Vocabulary: Acoustics and electroacoustics, [23-28, 23-31]
3. IEC 61094-1, Measurement microphones, Part 1: Specifications for laboratory standard microphones, pp. 8–11
4. IEC 61094-1/2/3/4/5/6, Measurement microphones, Series of microphone and calibration standards
5. K. Rasmussen, "The Influence of Environmental Conditions on the Pressure Sensitivity of Measurement Microphones", Brüel & Kjaer Technical Review, No. 1–2001
6. IEC 50(801), Vocabulary: Acoustics and electroacoustics, 25-08 Reciprocal Transducer
7. R. K. Cook, "Absolute pressure calibration of microphones", Journal of Research of the National Bureau of Standards, 25, 489–505, 1940 and Journal of the Acoustical Society of America, 12, 415–420, 1941
8. L. L. Beranek, Acoustic Measurements, John Wiley & Sons, Inc., New York, pp. 113–148
9. G. S. K. Wong, "Primary Pressure Calibration by Reciprocity", in G. S. K. Wong and T. F. W. Embleton (Ed.), AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1994, Chapter 4
10. IEC 61094-2, Measurement microphones, Part 2: Primary method for pressure calibration of laboratory standard microphones by the reciprocity technique
11. V. Nedzelnitsky, "Primary Method for Calibrating Free-field Response", in G. S. K. Wong and T. F. W. Embleton (Ed.), AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1994, Chapter 4

12. IEC 61094-3, Measurement microphones – Part 3: Primary method for free-field calibration of laboratory standard microphones by the reciprocity technique
13. IEC 61094-7, Measurement microphones – Part 7: Values for the difference between free-field and pressure sensitivity levels of laboratory standard microphones
14. H. G. Diestel, "Reciprocity Calibration of Microphones in a Diffuse Sound Field", Journal of the Acoustical Society of America, 33 (4), 1961
15. IEC 61183, Electroacoustics – Random-incidence and diffuse-field calibration of sound level meters
16. K. Brinkmann and H. Goydke, "Random-incidence and Diffuse-field Calibration" in G. S. K. Wong and T. F. W. Embleton (Ed.), AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1994, pp. 120–135
17. IEC 61094-4, Measurement microphones, Part 4: Specifications for working standard microphones
18. IEC 61094-5, Measurement microphones – Part 5: Methods for pressure calibration of working standard microphones by comparison
19. GUM, Guide to the Expression of Uncertainty in Measurement, 1993 (1995), International Organization for Standardization, Geneva, Switzerland
20. IEC 61094-6, Measurement microphones – Part 6: Electrostatic actuators for determination of frequency response
21. E. Frederiksen, "Electrostatic Actuator" in G. S. K. Wong and T. F. W. Embleton (Ed.), AIP Handbook of Condenser Microphones, American Institute of Physics, New York, 1994, p. 231
22. ISO/IEC 17025, General requirements for the competence of testing and calibration laboratories
23. EA 4/02, Expressions of the Uncertainty of Measurements in Calibration, European Cooperation for Accreditation (previously EAL-R2, 1999)
24. E. Frederiksen, "Verification of High-Pressure Linearity and Distortion of Measurement Microphones", Proceedings of the International Conference on Acoustics, Kyoto, 2004
25. IEC 60942, Electroacoustics – Sound Calibrators
26. IEC 61672-1, Electroacoustics – Sound Level Meters – Part 1: Specifications

70

Piezoceramic Accelerometer

James Mathews

PCB Piezoelectronics Inc., San Clemente, CA, USA

1 • Introduction

Piezoceramics belongs to a sub-group of piezoelectricity. The post-WWII era research and development, in metal oxide-based piezoceramics, ushered in dramatic opportunities for transducer applications. Natural crystal materials such as quartz and tourmaline were still being used, but the new ferroelectrics (man-made crystalline materials) offered substantially higher sensitivities and temperature ranges for new products such as miniature microphones and updated sonar designs. This period experienced major research efforts, by private companies as well as various governments' laboratories, emphasizing processes for poling piezoceramic/ferroelectric materials, i.e. subjecting the piezoceramic material to a high dc voltage to achieve specific properties. The key properties targeted were coupling coefficients, stiffness and dielectric constant, among others. A sample group of ferroelectric piezoceramics and natural crystal types is shown in Table 1.

2 • Technology Fundamentals

On the basis of the preceding review of piezoelectricity, the piezoceramic accelerometer may now be detailed further. Two of the most widely used sensing materials used in the

TABLE 1 Ferroelectric piezoceramics and natural crystal types and their piezoelectric constants

Material	Dielectric constant	Piezoelectric constant(pC/N)	Curie temperature(°C)
Ferroelectric			
Lead zirconate titanate/PZT	1600	300	370
Bizmuth titanate	160	20	680
Natural			
Quartz	4	2	570
Tourmaline	7	2	NA
Langatate/LGT			
Compression	76	6	NA
Shear	19	12	NA

Dielectric constant – measures the amount of charge that an electroded piece of material can store relative to the charge that would be stored by identical electrodes separated by air or vacuum at the same voltage.
Piezoelectric constant – relation of charge collected on the electrodes to the applied mechanical stress
Curie temperature (T_C) – the absolute maximum use temperature, in degree Celsius, beyond which piezoceramic materials transform to a high symmetry state or become non-piezoelectric
Note: The sensitivity in the direction perpendicular to the main direction is approximately 5% of that of the main direction.

manufacture of accelerometers are quartz and piezoceramics. Since quartz is a naturally grown material, although it may also be laboratory grown, it does not fall into the category of piezoceramics. However, a basic comparison of both quartz and piezoceramic/ferroelectric sensing materials will clarify why piezoceramics are the material of choice in the manufacture of transducers such as accelerometers.

Quartz has traditionally been used in applications where thermal transients were known to be encountered. Since quartz is not pyroelectric, that is having a property of signal output due to a sudden temperature change, it is considered more stable in thermally active environments. Quartz is excellent for force and pressure transducer applications.

Additionally, accelerometer designs using quartz as the sensing element provide a desirable feature of long-term stability and repeatability. This is because quartz has a low capacitance value, where the voltage sensitivity is relatively high compared to that

for most ceramic materials, making it ideal for use in voltage-amplified systems.

$$V = Q/C$$

where

V = voltage (V)
Q = charge (pC)
C = capacitance (pF)

However, quartz has a low charge output sensitivity, relative to the physical size of the crystal, which limits its use in several other accelerometer applications. Piezoceramic based accelerometer designs exhibit high-charge sensitivity characteristics. These accelerometers offer a charge sensitivity > 150 times that of quartz. Additionally, these type of accelerometer designs are currently capable of an operating temperature range of > 1,000°F (537°C).

The basic construction of a piezoelectric accelerometer is depicted in Figure 1:

Piezoceramic accelerometers, utilizing ferroelectric materials, consist of a seismic mass that stresses the polycrystalline material when subjected to motion.

The accelerometer's charge sensitivity is given as:

$$S = d_{33}ma$$

where

S = sensitivity in pC/g
d_{33} = piezoelectric charge constant (pC/N)
m = mass
a = acceleration normalized to gravity of 9.8 m/s^2

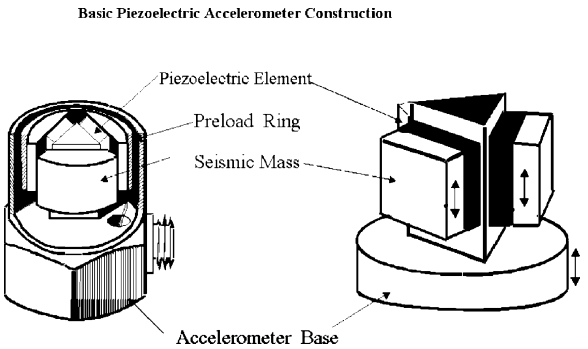


FIGURE 1 Basic construction of a piezoelectric accelerometer.

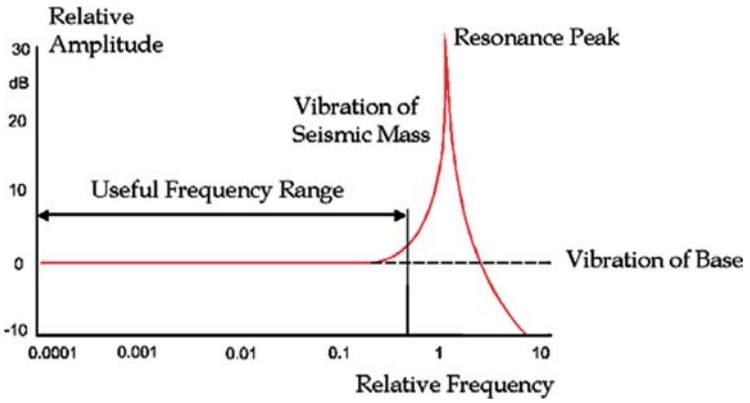


FIGURE 2 Typical piezoelectric accelerometer frequency response.

The active elements of the accelerometer are the piezoelectric elements. The elements acts as springs, which have a stiffness k, and connect the base of the accelerometer to the seismic masses. When an input is present at the base of the accelerometer, a force (F) is created on piezoelectric material that is proportional to the applied acceleration (a) and size of the seismic mass (m). (The sensor is governed by Newton's second law of motion $F = ma$.) The force experienced by the piezoelectric crystal is proportional to the seismic mass times the input acceleration. The more the mass or the acceleration, the higher the applied force and the more the electrical output from the crystal. The frequency response of the sensor is determined by its resonant frequency, which can generally be modeled as a simple single degree of freedom system. Using this system, the resonant frequency (ω) of the sensor can be estimated by:

$$\omega = \sqrt{k/m}$$

The typical frequency response of piezoelectric accelerometer is depicted in Figure 2.

3 • Construction

Piezoceramic accelerometers are ac-coupled devices that provide an output proportional to the acceleration caused by vibration and/or shock. A variety of mechanical configurations that follow the transduction principles of a piezoceramic accelerometer are available. These configurations are defined by the nature in which the inertial force of an accelerated mass acts upon the piezoelectric material. There are three primary configurations in use

today: shear, flexural beam and compression. The shear and flexural modes are the most common, while the compression mode is used less frequently, but is included herein as an alternative configuration.

3.1 Shear Mode

Shear mode accelerometer designs bond, or "sandwich," the sensing material between a center post and seismic mass, as shown in Figure 3. A compression ring or stud applies a preload force required to create a rigid linear structure. Under acceleration, the mass causes a shear stress to be applied to the sensing material. This stress results in a proportional electrical output by the piezoelectric material. The output is then collected by the electrodes and transmitted by lightweight lead wires either to the built-in signal conditioning circuitry of Integral Electronic Piezo-Electric (IEPE) sensors or directly to the electrical connector for a charge mode type. By isolating the piezoceramic from the base and housing, shear accelerometers excel in rejecting thermal transient and base bending effects. Also, the shear geometry lends itself to small size, which promotes a high-frequency response while minimizing mass loading effects on the test structure. With this combination of ideal characteristics, shear mode accelerometers offer optimum performance. There are three main types of shear designs: annular shear, radial shear and flat plate shear. The three types are defined by what face of the ceramic the electrodes

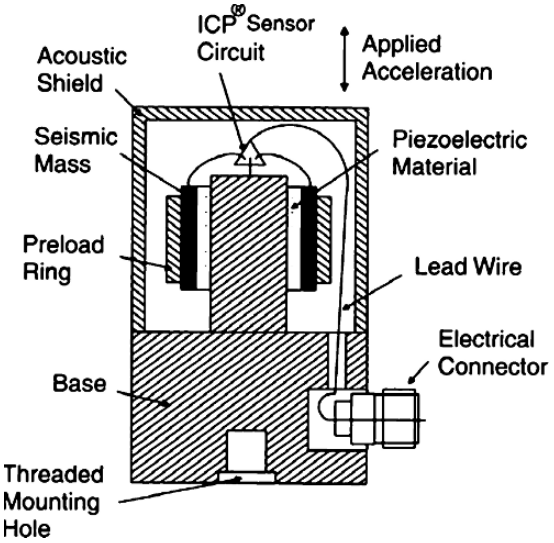


FIGURE 3 Shear mode accelerometer.

TABLE 2 Three types of shear mode accelerometers

	Annular shear	Radial shear	Flat plate shear
Electrodes	Outer diameter of ring	Top and bottom of ring	Outer face of plate
Direction of poling	Same direction as applied acceleration	Perpendicular to direction of applied acceleration	Same direction as applied acceleration

are applied, and the specific direction of poling, whether parallel or perpendicular to the direction of acceleration. See Table 2.

3.2 Flexural Mode

Flexural mode designs utilize beam-shaped sensing materials, which are supported to create strain on the material when accelerated, as shown in Figure 4. The ferroelectric material may be bonded to a carrier beam that increases the amount of strain when accelerated. The flexural mode enables low-profile, lightweight designs to be manufactured at an economical price. Insensitivity to transverse motion is an inherent feature of this design. Generally, flexural beam designs are well suited for low-frequency, low-gravitational (g) acceleration applications like those that may be encountered during structural testing.

3.3 Compression Mode

Compression mode accelerometers are simple structures which that provide high rigidity (see Figure 5). They represent the traditional or historical accelerometer design. Upright compression designs sandwich the piezoceramic between a seismic mass and rigid mounting base. The electrodes are applied to the top and bottom surface of the ring-shaped ceramic. The direction of polarization is parallel to the direction of acceleration. A pre-load stud or screw secures the sensing element(s) to the mounting

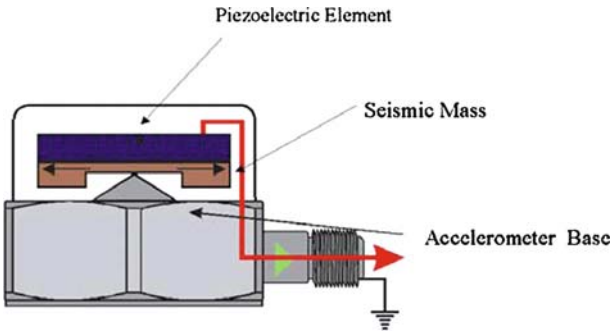


FIGURE 4 Flexural model accelerometer.

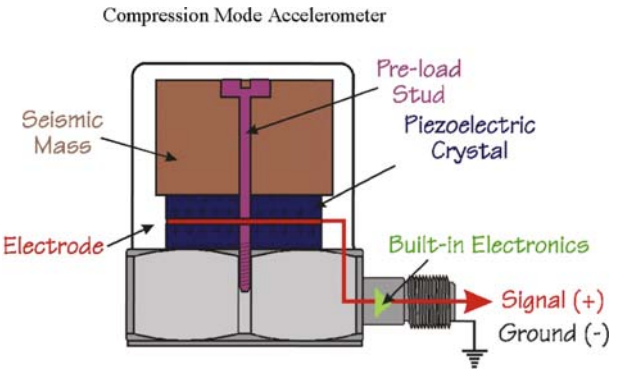


FIGURE 5 Compression mode accelerometer.

base. When the sensor is accelerated, the seismic mass increases or decreases the amount of force acting on the piezoceramic, and a proportional electrical output results. The larger the seismic mass, the greater the stress and, hence, the greater the output.

This design is generally very rugged and can withstand high-*g* shock levels. However, due to the intimate contact of the sensing material with the external mounting base, upright compression designs tend to be more sensitive to base bending (strain). Additionally, expansion and contraction of the internal parts act along the sensitive axis, making the accelerometer more susceptible to thermal transient effects. These effects can contribute to erroneous output signals when used on thin-sheet metal structures or at low frequencies in thermally unstable environments, such as outdoors or near fans and blowers.

4 • Selecting and Specifying

Piezoceramic accelerometers are designed and manufactured in either charge (SEPE, Diff. PE) or low impedance (IEPE) designs. The single-ended piezoelectric (SEPE, see Figure 6) and differential piezoelectric (Diff.PE) designs are high-impedance charge output accelerometers that provide low-level signals measured in pC/g units. Diff.PE designs are typically used in gas turbine test/monitoring environments, as well as in other applications, where common mode noise rejection, for good signal transmission, is paramount. Both SEPE and Diff.PE accelerometers require the use of low-noise-treated cabling. When the shield of the cable separates from the Teflon dielectric, during normal vibratory events, a charge builds up and bleeds into the signal

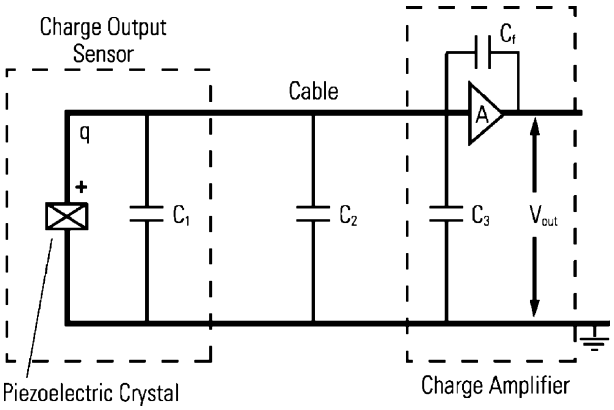


FIGURE 6 Charge amplified system schematic.

conductor, contaminating the normal charge acceleration signal. This is called the triboelectric effect. Low-noise-treated cables have a carbon graphite treatment over the Teflon dielectric, which is a conductive medium between the cable shield and the Teflon insulation. This allows any charge build-up, as described above, to bleed into the shield, or low side of the coaxial cable, thus minimizing any noise or non-acceleration-generated charge inputs to the center conductor.

As shown in Figure 6, the other end of the low-noise cable is connected to the charge amplifier, which is designed so that the high-impedance charge variation of the sensor is converted to a low-impedance voltage suitable for transmission over longer cables. Charge mode accelerometers are typically used when high-temperature survivability is required.

If the measurement signal must be transmitted over long distances, its use as in-line charge converter, placed near the accelerometer, is recommended (see Figure 7). This minimizes the chance of noise. In-line charge converters can be operated from the same constant-current excitation power source as IEPE accelerometers for a reduced system cost. In either case, the use of a special low-noise cable is required between the accelerometer

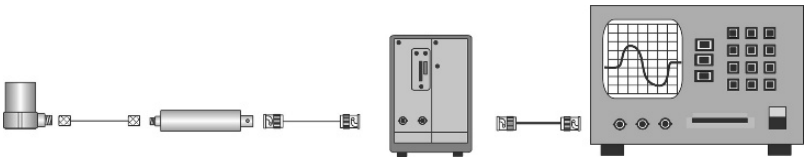


FIGURE 7 Typical in-line charge converter system.

and the charge converter to minimize vibration-induced triboelectric noise.

The low-impedance (IEPE) designs are voltage output accelerometers that provide signals measured in mV/g units. This type of accelerometer may incorporate the same seismic elements as standard charge type accelerometers, but have an added feature of on-board electronic circuitry (see Figure 8). This on-board, hybrid circuitry converts the high-impedance charge output signal to a low-impedance voltage. The same style coaxial output connectors are used for both types of accelerometers. By having the accelerometer provide a positive dc bias level above ground potential, which the accelerometer signal rides on, the system maintains the single-ended design. The dc bias level is then blocked at the input circuitry of the follow-on signal conditioner or remote charge converter, via a blocking capacitor. The IEPE accelerometer designs do not require special low-noise-treated cabling. Both SEPE accelerometer deigns, charge and IEPE, are available in uniaxial and triaxial versions. Manufacturers may provide custom designs based on specific customer requirements. Piezoceramic accelerometers are available in a multitude of sizes, shapes, weights and price range. Accelerometer models are manufactured with various case materials such as stainless steel, aluminum, titanium and inconel. Designs are based on dynamic and environmental factors as well as cost to manufacture and the intended end use. Specific application requirements will ultimately determine which piezoceramic accelerometer design is most appropriate.

In Table 3, some of the typical characteristics of different sensor types are shown. The measurement frequency range is

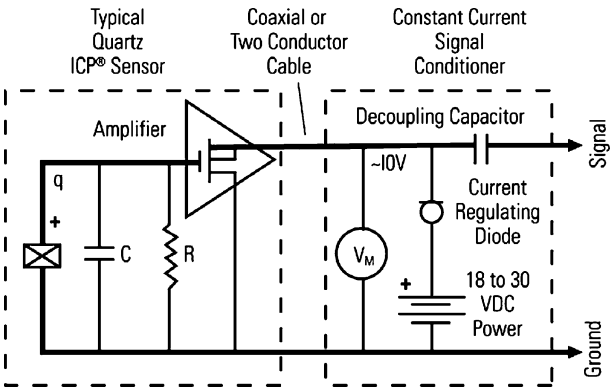


FIGURE 8 Typical ICP sensor system.

TABLE 3 Some of the *typical* characteristics for different sensor types

Accelerometer type	Frequency range	Sensitivity	Measurement range	Dynamic range	Size/weight
IEPE piezoelectric accelerometer	0.5–50,000 Hz	0.05 mV/g–10 V/g	0.000001–100,000 g's	~ 120 dB	0.2 gram to 200 + grams
Charge piezoelectric accelerometer	0.5–50,000 Hz	0.01–100 pC/g	0.00001–100,000 g's	~ 100 dB	0.14 grams to 200 + grams

from 0.5 to 50 kHz. In order to measure high-frequency signals, the size and weight must be small.

The advantages, limitations and typical applications of IEPE piezoelectric and charge piezoelectric accelerometers are listed below:

IEPE piezoelectric accelerometer

Advantages:

- Wide dynamic and frequency ranges. Durable (high shock protection).
- Powered by low-cost constant current source. Fixed output.
- Less susceptible to EMI and RF inference. Can be made very small.
- Less operator attention, training and installation expertise required.
- Long cable driving with out noise increase.
- Operates into many data acquisition devices with built-in constant current input.
- Operates across slip rings. Lower system cost per channel.

Limitations:

- Limited temperature range. Maximum temperature of 175°C (350°F)
- Low-frequency response is fixed within the sensor.
- Built-in amplifier is exposed to same test environment as the element of the sensor.

Typical applications:

- Modal analysis, NVH, engine NVH, flight testing, body-in-white testing.
- Cryogenic. Drop testing. Ground vibration testing. HALT/HASS.
- Seismic testing. Squeak and rattle. Helmet and sport equipment testing.
- Vibration isolation and control.

Charge piezoelectric accelerometer

Advantages:

High operating temperatures to 700°C. Wide dynamic range. Wide frequency range. High shock protection (durable). Flexible output

Simpler design with fewer parts. Charge converter electronics usually at ambient conditions away from test environment.

Limitations:

More care/attention is required to install and maintain.

High-impedance circuitry must be kept clean and dry.

Capacitive loading from long cable run results in noise floor increase.

Powered by charge amp, which can be complicated and expensive.

Need to use special low-noise cable.

Typical applications:

Steam pipes. Turbo machinery. Steam turbine. Exhaust. Brake.

5 • Mounting

One of the most important considerations in dealing with accelerometer mounting is the effect the mounting technique has on the accuracy of the usable frequency response.

5.1 Stud Mounting

For permanent installations, where a very secure attachment of the accelerometer to the test structure is preferred, stud mounting is recommended (Figure 9). First, grind or machine on the test object a smooth, flat area at least the size of the sensor base, according to the manufacturer's specifications. Then, prepare a tapped hole in accordance with the supplied installation drawing, ensuring that the hole is perpendicular to the mounting surface.

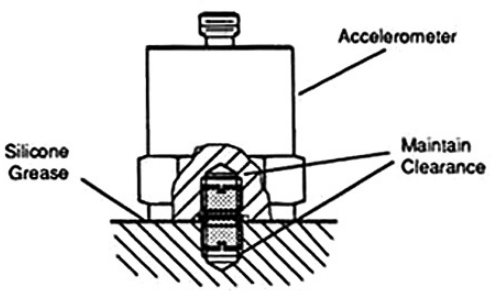


FIGURE 9 Typical stud mounting.

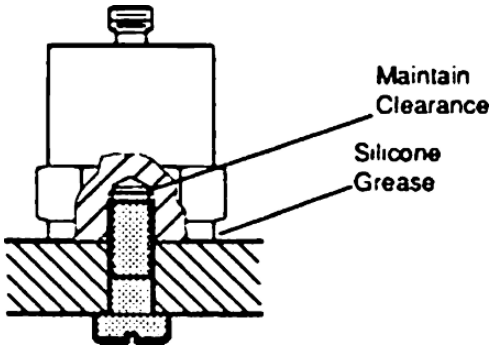


FIGURE 10 Typical screw mounting.

Install accelerometers with the mounting stud and make certain that the stud does not bottom in either the mounting surface or the accelerometer base.

5.2 Screw Mounting

When installing accelerometers onto thin-walled structures, a cap screw passing through a hole of sufficient diameter is an acceptable means for securing the accelerometer to the structure (see Figure 10).

5.3 Adhesive Mounting

Occasionally, mounting by stud or screw is impractical. For such cases, adhesive mounting offers an alternative mounting method (see Figure 11). Petro wax offers a very convenient, easily removable approach for room temperature use.

5.4 Magnetic Mounting

Magnetic mounting bases offer a very convenient, temporary attachment to magnetic surfaces (see Figure 12). Magnets offering high pull strengths provide the best high-frequency response.

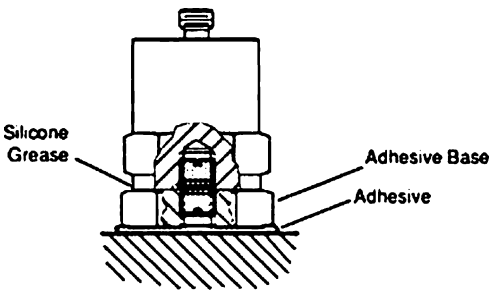


FIGURE 11 Typical adhesive mounting.

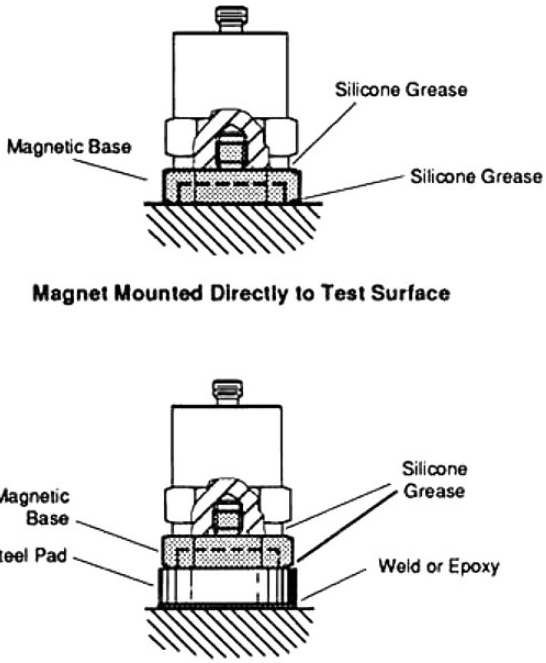


FIGURE 12 Magnet mounting to steel.

6 • Latest and Future Developments

Current developments for piezoceramic accelerometers include the IEEE 1451.4 Smart Transducer Interface, which includes the Transducer Electronic Data Sheet or TEDS capability, and elevated temperature IEPE accelerometer designs including high-temperature ($> +175°C$) hybrid circuitry. Future developments include widespread digital output accelerometer capabilities; network sensing accelerometer modules with on-board data acquisition capability and wireless accelerometers. Applications requirements, for extreme temperature acceleration measurements at very high frequencies, are fielded by transducer manufacturers daily. This in turn fuels new research on piezoceramics by universities , government laboratories and the private sector, pushing the boundaries on the current thinking of how acceleration measurements are going to be made, and how they will be integrated into the overall dynamic measurement matrix of the future.

7 • List of Applicable Standards

ISO 9001 (2000) – Quality system for assurance in design, development and production

ISO 5347 (1997) – Methods for the calibration of vibration and shock pick-ups

ISA-RP37.2 (R1995) – Guide for specifications and tests for piezoelectric acceleration transducers for aerospace testing

ANSI S2.17-1980 (R1997) – Standard concerned with vibration measurement quantities, equipment, and procedures involved in operating machinery

IEST RP-DTE011.1 – Recommended practice (proposed) for shock and vibration transducer selection (Institute of Environmental Sciences and Testing)

Major Manufacturers

Piezoceramic Accelerometers

PCB Piezotronics
3425 Walden Ave
Depew NY 14043
Tel: 716-684-0001
Fax 716-684-0987
Toll Free 1 800 828-8840
Sales@pcb.com
webmaster@pcb.com
www.pcb.com

Brüel & Kjær S & V Denmark (HQ)
Skodsborgvej 307
2850
Naerum
Denmark
Tel: 45 45 80 05 00
Fax: 45 45801405
info@bksv.com
www.bksv.com

Endevco Corporation
30700 Rancho Viejo Road
San Juan Capistrano,
CA 92675
Toll Free: 1-800-982-6732
Tel: 1-949-493-8181
Fax: 1-949-661-7231
applications@endevco.com
http://www.endevco.com

Wilcoxon Research, Incorporated
21 Firstfield Road,
Gaithersburg, MD 20878
(1-800-945-2696)
Tel: 301-330-8811
Fax: 301-330-8873
sensors@wilcoxon.com
www.wilcoxon.com

Kistler Instrument Corporation
75 John Glenn Drive
Amherst, NY
14228-2171/USA
(1-888-547-8537),
Fax: 1-716-691-5226
sales.us@kistler.com

Dytran Instruments Incorporated
21592 Marilla Street,
Chatsworth,
CA 91311 USA
Tel: (818) 700-7818 •
Fax: (818) 700-7880 •
info@dytran.com
www.dytran.com

Columbia Research Labs
1925 Macdade Blvd
Woodlyn PA 19094
Tel: 1-800-813-8471
Fax: 1 610 872 3882
www.columbiaresearchlab.com

Vibro-Meter SARte de Moncor 4
CH-1701 Fribourg
Switzerland
Tel: +41 26 407 11 11
info@vibro-meter.com
www.vibro-meter.com

Vibra-Metrics
195 T Clarksville Rd
Princeton Junction NJ 08550
Tel: 609 716–4000
Fax: 609 716 – 0706

References and Resources

Textbooks

C. M. Harris (ed), Shock and Vibration Handbook, 4th ed., McGraw-Hill, New York, NY 10020, 1996

K.G. McConnell, Vibration Testing Theory and Practice, John Wiley & Sons Inc, New York, NY 10158, 1995

Primary Literature Citations

Institute of Environmental Sciences and Technology, RP-DTE011.1, Shock and Vibration

Aszkler, Craig ,Vibration Division Manager, PCB Piezotronics, Inc., Sensor Technology Handbook, Ch. 05

T. L. Jordan, Nasa Langley research Center, Hampton VA, Ouanaies, Z., ICASE, Hampton, VA

NASA/CR-2001-211225, Piezoelectric Ceramics Characterization, September 2001. http://www.piezo.com/history.html

71

Laser Doppler Velocimeter

Nobutoshi Yoshida

Ono Sokki Co. Ltd., 1-16-1 Hakusan, Midori-ku, 226-8507 Yokohama, Japan, e-mail: yoshi@onosokki.co.jp

Laser Doppler Velocimeter, a vibrometer featuring non-contact, high resolution, and wide dynamic range.

1 • Introduction

Generally, in order to treat vibrations quantitatively, we use the following three parameters: displacement, velocity and acceleration. Each parameter can be transformed into another by differential or integral. If velocity is constant over frequency, displacement and acceleration versus frequency are as described in Figure 1. This graph shows displacement is most sensitive at a low frequency, but that sensitivity of acceleration exceeds that of the others at a high frequency. This means that for sensitive measurement, displacement should be used at low frequencies, and acceleration should be used at high frequencies. For applications such as plant diagnosis, displacement and velocity are used at frequencies under several hundred hertz and acceleration is used at higher frequencies. Moreover, it is necessary to consider the mass of the object, the size of the vibration, the range of the frequency, and the measurement environment, and so on to select an appropriate sensor to detect the vibration in high accuracy. There are two types of sensors, the contact and the non-contact types. The piezoelectric type acceleration pickup

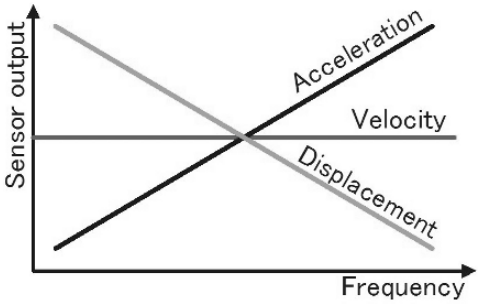


FIGURE 1 Velocity, displacement and acceleration versus frequency.

has been widely used as a contact type vibration sensor in the past [1]. A piezoelectric acceleration pickup uses a piezoelectric device as a spring for a seismic system and a mechano-electric transducer [2]. This sensor outputs an electric signal proportional to vibration acceleration. The system can be achieved at low cost, and measurement is comparatively easy. However, it is necessary to take account of the mass effect and the contact resonance for the contact type. Mass effect refers to the sensor mass ability to affect an object's natural frequency. Because natural frequency depends on mass, an object to which the sensor is attached will have a lower natural frequency than that of the object itself. If a sensor mass is too large, the sensor will decrease the object's natural frequency, thus resulting in measurement error. If the object's mass, sensor mass, and the system's natural frequency be M, m and fe, respectively, and the natural frequency will decrease by Δfe according to the equation shown on Figure 2. It is recommended that the sensor mass be under one-fiftieth of the object's mass. When m is one-fiftieth of M , the frequency deviation $\Delta fe/fe$ amounts to 0.01. It is necessary to note that M is not the mass of the whole object but the mass of the part to which the sensor will be attached. The part may be unexpectedly light. The non-contact sensor, which does not influence the vibration body, includes the capacitive sensor, the eddy current sensor, the optical sensor, and no on. An optical

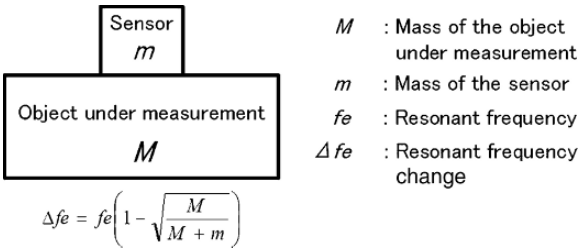


FIGURE 2 Mass effect.

sensor includes the laser displacement meter with triangulation method and the Laser Doppler method. This chapter describes the Laser Doppler method by the velocity measurement whose dynamic range is larger than the displacement measurement by the triangulation method [3]. With the decrease in size and weight of electronic components and mechanical components in recent years, an increasing number of Laser Doppler Vibrometers (LDV) applying the Doppler shift of laser beam are being used as a detector to analyze the action, vibration and structure of the object under measurement. In particular, the LDV has become essential as a detection and analysis tool for micro electrical mechanical systems (MEMS). This chapter explains a micro-vibration measurement system applying the LDV. First of all, the features of LDV are described as follows.

1.1 Features of LDV

The LDV features the following:

(1)Non-contact detection

In many cases, an acceleration detector with a piezo-electric element is used to detect the behavior and vibration of the object under measurement. Usually mainstream vibrometers are used as the acceleration detector, because of ease of use, relatively low price and ample product lineup. However, it is necessary to bring the vibrometer into contact with the object under measurement due to its operating principle. Since even a compact type acceleration detector has a weight of about 0.5 g, the weight of the detector itself can significantly affect the measurement if the object under measurement is small and light. In the case of microstructures such as MEMS, attachment is not physically possible. On the other hand, the LDV makes it possible to detect vibration of the object under measurement by using only irradiating laser light.

(2) High spatial resolution

Spatial resolution refers to how minutely the behavior in a unit area can be measured in the form of coordinates, i.e., how microscopic target object can be measured. Since the LDV uses laser light, light condensing through lenses is possible using optical devices, allowing high spatial resolution to be attained in a relatively easy way. Non-contact detectors include displacement detectors that use capacitance and eddy current for operating principle. Neither type of detector can reduce the detection area to the area of the detector's end face or less. The LDV has a minimum measurement distance of 20 μm or less and, when a specialized object lens is attached, a minimum beam spot diameter of about $\phi 3\,\mu m$.

(3)Wide dynamic range

The dynamic range of the LDV is overwhelmingly wider than that of the capacitance and eddy current sensors of non-contact detectors and laser displacement meters with triangular surveying. Although the LDV is based on the velocity output and therefore simple comparison is difficult, it is provided with a displacement resolution of 0.01 nm when displacement conversion is made in high-frequency regions. With the use of a fringe count displacement meter unit that directly performs displacement demodulation, a resolution of 5 nm can be attained.

2 • Operation Principle

If sound wave, electric wave or light wave with a certain fixed frequency component is applied to an object moving at a certain speed, the frequency changes in proportion to the velocity component of the moving object. This is referred to as the Doppler effect. The relationship between the frequency of the irradiated wave and the frequency of the reflected wave is as follows:

(1) When the object is approaching, the frequency of the irradiated laser light is less than the frequency of the reflected light. When the object is moving away, the frequency of the irradiated light is greater than the frequency of the reflected light
(2) The difference between the frequency of the irradiated light and the frequency of the reflected light is proportional to the moving velocity of the object and increases with increasing speed.

The LDV applies this principle. If laser light is irradiated onto the moving target object, the frequency of the reflected light from the target object reveals the Doppler shift from the frequency of the irradiated light because of the Doppler effect. The frequency shift Δf is given by:

$$\Delta f = f' - f = \frac{2V}{\lambda} \cos\theta. \tag{1}$$

where f and f' are the frequencies of the irradiated and reflected light, respectively, V is the speed of the target object, λ is the wavelength of the irradiated light, and θ is the angle between the application direction of the irradiated light and the moving direction of the target object (see Figure 3).

Since the wavelength of the laser light, λ, used with the LDV is extremely stable, the Doppler frequency, Δf, is proportional to the moving velocity of the target object, V. With the LDV,

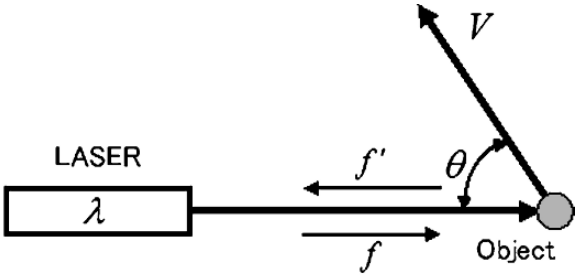


FIGURE 3 Measurement principle.

since the angle between the laser irradiation direction and the moving direction of the target object, θ, is usually set to 0° (only the component of the reflected light in parallel with the incident light is detected: out-of-plane vibration), it is possible to obtain the moving velocity in the irradiation direction of the target object by measuring the Doppler frequency Δf. However, since the frequency of the laser light itself is very high and difficult to measure directly, the Doppler frequency Δf is usually detected through interference between the irradiated light (f) and the reflected light (f').

3 • Basic Construction

The LDV consists of an optical head that irradiates laser light and a converter that processes the Doppler frequency from the reflected light. The signal from the converter becomes a voltage signal proportional to the moving velocity of the object under measurement. When this signal is input to a fast Fourier transform (FFT) analyzer, it can be converted to acceleration and displacement allowing analysis of the behavior of the object under measurement through frequency analysis of the signal. The LDV are classified as follows in terms of the configuration.

3.1 One-Dimensional LDV

The laser beam irradiated from the light source is divided into two by a beam splitter. One beam is the incident beam, which is applied to the object under measurement, and the other is the reference beam, which is returned inside the LDV. The reflected beam returning from the target object causes a Doppler shift corresponding to the vibration velocity of the object under measurement. The reflected beam interferes with the reference beam to which a frequency shift was given in advance by the acoustic optical modulator (AOM), to obtain a beat frequency. This beat signal, with the Doppler-shifted frequency component extracted by the detector circuit, is output as a voltage signal corresponding to the vibration velocity by the FM demodulator

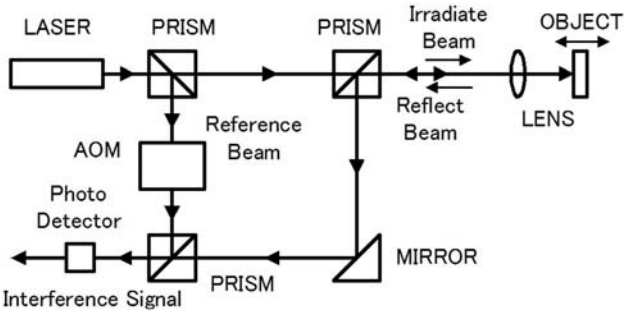


FIGURE 4 Construction block diagram of basic LDV.

circuit. This is the basic configuration of the interference optical system for the LDV. The block diagram of the basic LDV configuration is shown in Figure 4 and the photograph of the corresponding LDV in Figure 5. Example specifications are shown in Table 1. If the luminous energy of the laser light, which is reflected from the object and returns to the light sensor, is below a given level, the signal-to-noise ratio deteriorates [4]. This deterioration sometimes results in a failure to perform measurement. For this reason, the LDV is best used for objects that reflect beams in completely opposite directions. This LDV can also measure objects that scatter beams, if the luminous energy of the reflected laser light, which returns to the object lens in the sensor head, is above the specified level. More specifically, even objects such as black rubber that reflect little luminous energy can be measured using LDV if the sensor head is close to the object. If the luminous energy is likely to be inadequate, reflexive materials such as a reflection sheet or reflective paint must be attached to or coated on the object.



FIGURE 5 Photograph of the one-dimensional LDV system.

TABLE 1 Specification of the one-dimensional LDV

Sensor unit	Wide range type	High-resolution type
Detection demodulation system	Velocity demodulation using optical heterodyne detection	
Light source	He-Ne laser (632.8 nm wavelength)	
Irradiated light output	Within 1 mW (Class2 IEC60825-1:2001)	
Measurement distance	100 mm–5 m	
Laser spot size	20 μm or less (at the shortest measurement distance)	
Converter unit	Wide range Type	High-resolution type
Frequency range	1 Hz–3 MHz	1 Hz–300 kHz
Velocity range	0.01 m/s/V (max. 0.1 m/s) 0.1 m/s/V (max. 1 m/s) 1 m/s/V (max. 10 m/s)	0.002 m/s/V (max. 0.02 m/s) 0.01 m/s/V (max. 0.1 m/s) 0.05 m/s/V (max. 0.5 m/s)
Velocity output	Analogue voltage ± 10 V (DC offset within ± 20 mV)	
Minimum velocity resolution	0.3 μm/s (when the range is 0.01 m/s/V and the LPF 1 MHz or 1 kHz)	0.1 μm/s (when the range is 0.002 m/s/V and 1 kHz filter off)
Low-pass filter	100 kHz `1MHz ` OFF	20 kHz ` 100kHz ` OFF
High-pass filter	10 Hz ` 100Hz ` OFF	

3.2 Three-Dimensional LDV

An object vibrates not necessarily only in one direction but in complicated three-dimensional motion. With the increase in the integration and precision of products in recent years, the necessity of measuring and analyzing such complicated three-dimensional motion is also increasing. The three-dimensional LDV uses three reference light type laser Doppler velocimeters to perform vector operation on the signals obtained from three directions. This makes it possible to simultaneously measure the vibration velocities of the object under measurement in the X, Y, and Z directions and the vibration direction. The photograph of the optical head configuration of the three-dimensional LDV is shown in Figure 6.

One of the three optical heads is installed so that the angle of the incident beam agrees with the angle of the moving direction (Z direction) of the target object. The other two heads are installed at certain angles (Z_X, Z_Y) with respect to the Z axis. The signals obtained from the two heads at certain angles have respective vibration velocities in the Z_X and Z_Y directions and the vibration direction. As a result, the reflected beams from the three optical heads have vibration components in the Z, Z_X, and Z_Y directions, respectively. Through vector operation of the three signals,

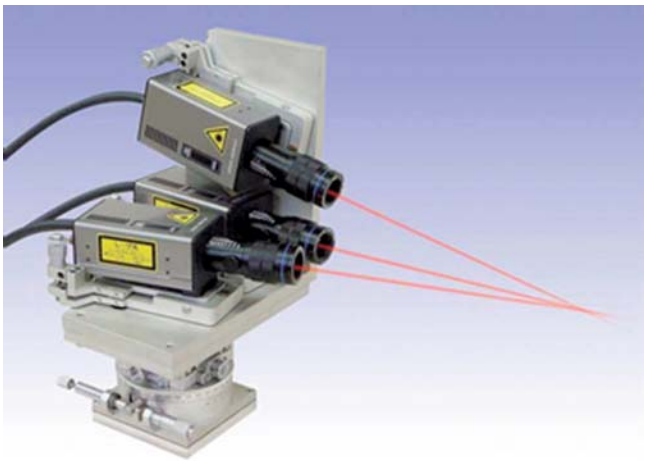


FIGURE 6 Photograph of optical head of three-dimensional LDV.

it becomes possible to simultaneously measure the vibration in each of the X, Y, and Z directions. Vector operation expressions are shown below (see Figure 7):

$$V_{ZX} = V_Z \cos\theta_X + V_X \sin\theta_X \tag{2}$$

$$V_{ZY} = V_Z \cos\theta_Y + V_Y \sin\theta_Y \tag{3}$$

$$V_X = \frac{V_{ZX} - V_Z \cos\theta_X}{\sin\theta_X} \tag{4}$$

$$V_Y = \frac{V_{ZY} - V_Z \cos\theta_Y}{\sin\theta_Y} \tag{5}$$

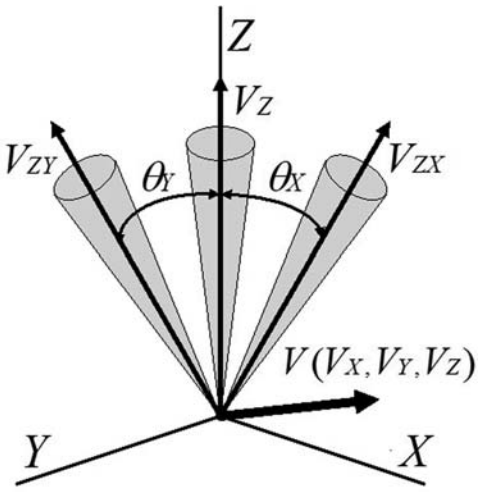


FIGURE 7 Measurement principle of vector operation method.

4 • Applications

4.1 Multipoint Vibration Measurement and Vibration Mode Analysis

Computer control of the LDV measurement system and high-precision stage makes it possible to perform automatically, in a short time, multipoint vibration measurement and vibration mode analysis of a small object under measurement. An example of vibration mode analysis results of a piezo-electric element is shown in Figure 8.

4.2 Micro-object Vibration Measurement System

Figure 9 shows a micro-object vibration measurement system that combines the LDV with a microscope. The microscope optical system is used to irradiate laser light condensed to a maximum diameter of 1.5 μm or less to any portion of the object under measurement and, while monitoring the irradiated portion, to detect the vibration velocity of the object under measurement. Using a coaxial incident-light optical system and a digital still camera or a CCTV camera, this system makes it

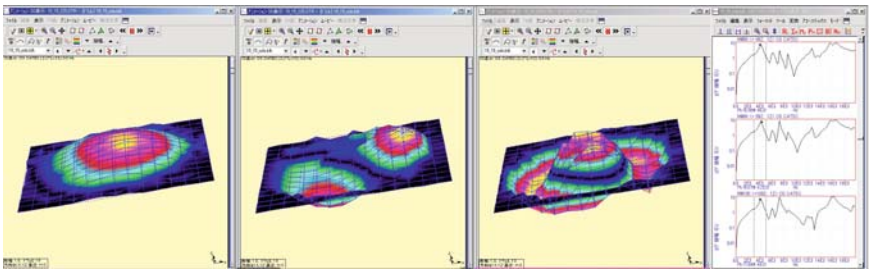


FIGURE 8 Vibration mode analysis result of piezo-electric element.



FIGURE 9 Photograph of micro-object vibration measurement system.

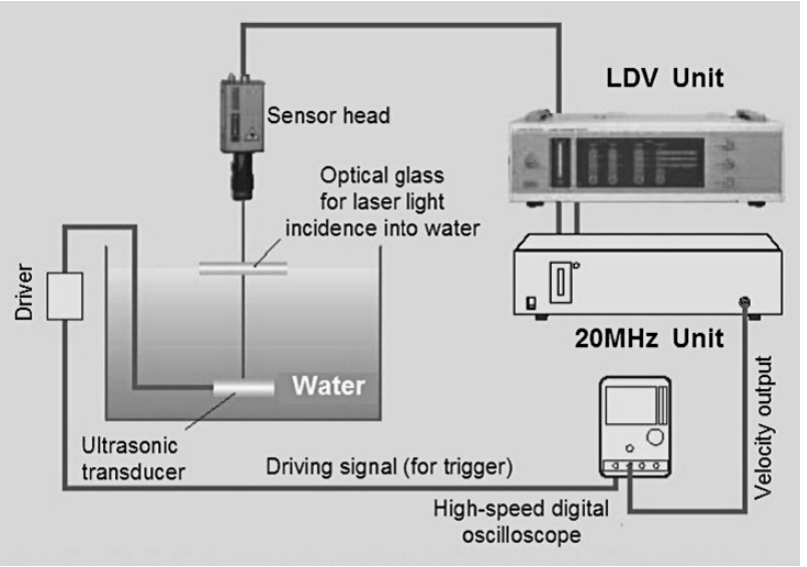


FIGURE 10 Block diagram of submerged measurement system.

possible to observe the detected displacement as a clear image, and record these images as still or moving pictures. It also allows non-contact measurement of vibration and behavior of micro-machines and other small-sized mechanical components as well as IC lead wires. The photograph of the micro-object vibration measurement system is shown in Figure 9.

4.3 Measurement of Ultrasonic Oscillator in Liquid

Taking advantage of the features of non-contact measurement, it is possible to measure the behavior of an ultrasonic transducer when soaked in liquid, which is used both for aerial and for submerged measurement. It can also be applied to vibration analysis of an ultrasonic washing tub or stirrer. The block diagram of the submerged measurement system is shown in Figure 10.

References

1. Kearns, R. W.: Velocities and Displacements Associated with the Transient Responce of Accelerometers. In: *Proceedings. Society for Experimental Stress Analysis* Paper 534 (1995)
2. http://www.onosokki.co.jp/English/hp e/products/keisoku/s v/np series.html
3. http://www.onosokki.co.jp/English/hp e/whats new/SV rpt/SV2/ frames.html
4. Drain, L. E.: The Laser Doppler Technique. (A Wiley-Interscience Publication, 1980) pp. 76–84

72

Capacitive Sensors

Masahiko Sakai

Ono Sokki Co.,Ltd. 1-16-1 Hakusan, Midori-ku, Yokohama, 226-8507 Japan, e-mail: msakai@onosokki.co.jp

Capacitive sensors are used to make non-contact displacement measurements with high precision and high bandwidth as well as inductive sensors and optical sensors [1]. Capacitive sensors are especially effective in measurement and control of spindle runout, vibration, and other demanding dynamic. However, the capacitive sensor has the fault that the sensor output often depends upon environmental conditions. It is necessary to understand the characteristics of capacitive sensors when using it. This chapter describes the principle of operation, gives some useful advice, and presents examples of applications.

1 • Principle of Operation

1.1 Basic Principle

A capacitor is formed when two parallel electrodes are brought near each other as shown in Figure 1. The capacitance $\boldsymbol{C}$ is a function of the electrode area $\boldsymbol{A}$, the dielectric constant $\boldsymbol{\varepsilon}$, and the distance $\boldsymbol{d}$ between the electrodes. The capacitance $\boldsymbol{C}$ can be expressed by the following simple equation:

$$\boldsymbol{C} = \frac{\boldsymbol{\varepsilon A}}{\boldsymbol{d}} \tag{1}$$

In (1), if one electrode is a sensor plate and the other is a target plate, the distance $\boldsymbol{d}$ will be determined by measuring capacitance $\boldsymbol{C}$. A simple capacitive sensor that has only one electrode

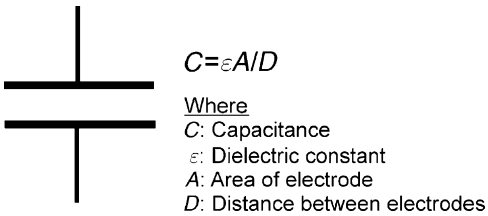


FIGURE 1 Relationships between capacitance $\boldsymbol{C}$ and Distance $\boldsymbol{d}$.

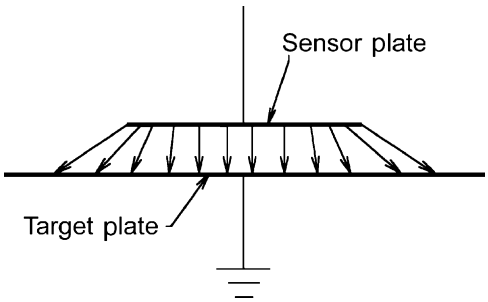


FIGURE 2 A simple capacitive sensor.

called a sensor plate is shown in Figure 2. The facing electrode is a target plate. We can see the electric field lines between the sensor plate and the target plate are curved at the edges of the sensor plate. In such simple case, $\boldsymbol{C}$, $\boldsymbol{\varepsilon}$, $\boldsymbol{A}$, and $\boldsymbol{d}$ will not satisfy (1). Therefore, another electrode called a guard plate is added around the sensor plate as shown in Figure 3. The electric potential of the guard plate shall be driven to the same electric potential of the sensor plate. We can see the electric field lines between the sensor plate and the target plate are straight due to the guard fields. In this case it satisfies (1). One more electrode should be

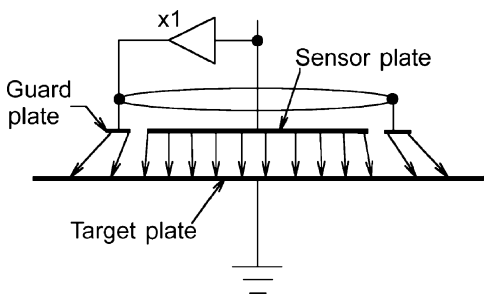


FIGURE 3 Effect of guard plate.

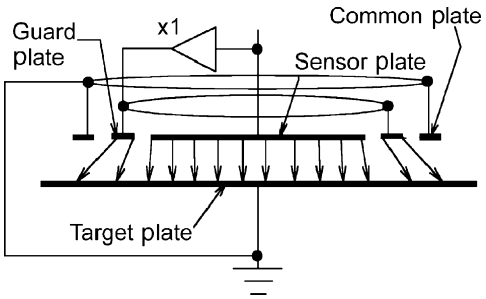


FIGURE 4 A typical capacitive sensor with guard plate and common plate.



FIGURE 5 Those are typical capacitive sensors. Sensor size effect the range. Smaller sensor shall be selected as the range is smaller.

added around the guard plate, and the electric potential of this plate shall be connected to the common level of the capacitive sensor as shown in Figure 4. Most capacitive sensors have the tri-axial structure of electrodes which consist of Sensor Plate, Guard Plate, and Common Plate in Figure 4. Some popular capacitive sensors are shown in Figure 5.

1.2 About the Electric Circuit of the Capacitive Sensor

The block diagram of the circuit is shown in Figure 6. In general, conventional circuits of capacitive sensors will require the use of a linearizer because the distance $\boldsymbol{d}$ will be inversely proportional to the capacitance $\boldsymbol{C}$ as shown in (1). The capacitance $\boldsymbol{C_x}$ is a function of the electrode area $\boldsymbol{A}$, the dielectric constant $\boldsymbol{\varepsilon}$, and the distance $\boldsymbol{d}$ between the sensor plate and the target plate. In Figure 6, *AMP* is an amplifier with unity gain, $\boldsymbol{E_{in}}$ is the input voltage of *AMP*, *OSC* is a voltage oscillator with amplitude $\boldsymbol{E_0}$ and frequency $\omega/2\pi$, $\boldsymbol{C_{ref}}$ is the reference capacitance and $\boldsymbol{i}$ is electric current through $\boldsymbol{C_{ref}}$ and $\boldsymbol{C_x}$. If the electric current into the *AMP* can be ignored, the output voltage $\boldsymbol{V_{out}}$ of the *AMP* can be expressed by the following equation.

$$\boldsymbol{V_{out}} = \boldsymbol{E_0}\sin(\omega t) + \left(\frac{1}{j\omega \boldsymbol{C_{ref}}} + \frac{1}{j\omega \boldsymbol{C_x}}\right)\boldsymbol{i} \tag{2}$$

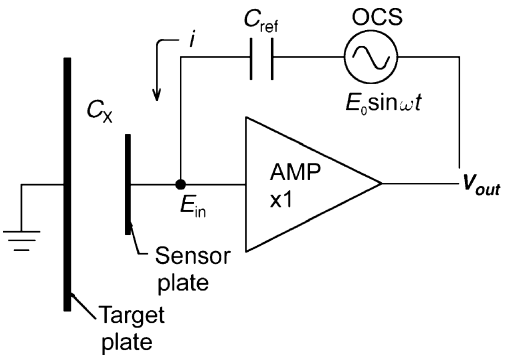


FIGURE 6 Electric circuit of capacitive sensor.

The electric current $\boldsymbol{i}$ through $\boldsymbol{C_x}$ and $\boldsymbol{E_{in}}$ is

$$\boldsymbol{i} = \boldsymbol{E_{in}} j\omega \boldsymbol{C_x} \tag{3}$$

Inserting (3) into (2) gives

$$\boldsymbol{V_{out}} = \boldsymbol{E_0} \sin(\omega t) + \left(\frac{\boldsymbol{C_x}}{\boldsymbol{C_{ref}}} + 1\right)\boldsymbol{E_{in}} \tag{4}$$

In (4) $\boldsymbol{V_{out}}$ is equal to $\boldsymbol{E_{in}}$ because *AMP* is an amplifier with unity gain.

$$\boldsymbol{V_{out}} = -\left(\frac{\boldsymbol{C_{ref}}}{\boldsymbol{C_x}}\right)\boldsymbol{E_0} \sin(\omega t) \tag{5}$$

By (1)

$$\boldsymbol{C_x} = \frac{\boldsymbol{\varepsilon A}}{\boldsymbol{d}} \tag{6}$$

Putting (6) into (5) and $\boldsymbol{K} = -\frac{\boldsymbol{C_{ref}}}{(\varepsilon A)}$ is constant

$$\boldsymbol{V_{out}} = \boldsymbol{K d E_0} \sin(\omega t) \tag{7}$$

We can get proportional output voltage $\mathbf{V_{out}}$ to the distance $\boldsymbol{d}$ between the sensor plate and the target plate by (7), while the output voltage will be inversely proportional to distance $\boldsymbol{d}$ by conventional circuits. As (7) does not depend upon a material characteristic of the target plate, $\mathbf{V_{out}}$ is independent of the target plate material if only it is electric conductor. On the other hand, output voltage of an inductive sensor is a function of target plate conductivity and output voltage of an optical sensor is a function of its reflection ratio or its surface roughness.

2 • Some Useful Advice for Capacitive Sensors

2.1 Where to Connect Sensors' Common

The electric potential of the target plate shall be connected to the sensors' common electric potential because any circuit of capacitive sensor will process voltage signals based on the common level. If not, the output voltage of capacitive sensor will be unstable. This is very important point when using capacitive sensors.

2.2 Environmental Considerations

The environmental consideration for capacitive sensors is that there is a nonconductive material between the sensor and the target. This material is air in most applications. Each material has unique dielectric constant $\boldsymbol{\varepsilon}$. Therefore, the output voltage of any capacitive sensor will change if this material changes from pure air to air filled with vapor. We have to take notice that output voltage $\boldsymbol{V_{out}}$ is a function of the dielectric constant $\boldsymbol{\varepsilon}$ as well as the distance $\boldsymbol{d}$ (see (7)). Also, one should keep the surface of the sensor plate and of the target plate clean. This is also for the same reason. Oil and dust on the sensor and target plates will also make dielectric constant change.

2.3 Selection of Sensor

Typical capacitive sensors are round and vary in size from about 1–40 mm in diameter (see Figure 5). The sensor size shall be chosen by the measurement range. We can get the same output voltage $\boldsymbol{V_{out}}$ at the same $\frac{d}{A}$ in (7). So, advanced capacitive sensors are designed with constant $\frac{d}{A}$. It is that any combinations between various size sensors and the electric circuit of the capacitive sensor will output the same voltage when the distance is equal to the measurement range. We can see some specifications about various size sensors in Table 1.

TABLE 1 Specifications about capacitive sensors

Type	VE-121	VE-222	VE-521	VE-133	VE-231	VE-531
Gap rating	0 ~ 0.1 mm	0 ~ 0.2 mm	0 ~ 0.5 mm	0 ~ 1.0 mm	0 ~ 2.0 mm	0 ~ 5.0 mm
Resolution	0.1 μm	0.1 μm	0.1 μm	0.1 μm	1.0 μm	1.0 μm
Linearity	±0.15%/fs					
Response frequency	DC to 10 kHz					
diameter	Φ3 mm	Φ6 mm	Φ8 mm	Φ10 mm	Φ20 mm	Φ40 mm
Operating temperature range	0~ + 80°C					

3 • Applications

We can see some application examples in Figures 7–10. Figure 7 shows an example of shaft vibration measurement. Shaft vibration and eccentricity can be observed by using an oscilloscope and a couple of sensors. Figure 8 shows an example of surface vibration measurement. Using an FFT analyzer, vibration and displacement analysis can be performed [2]. Non-contact

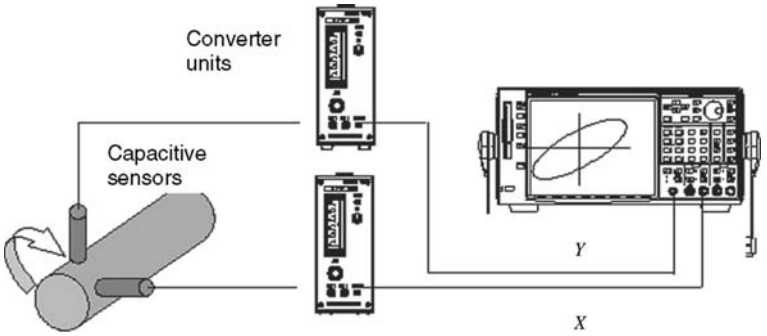


FIGURE 7 Application for shaft vibration and eccentricity.

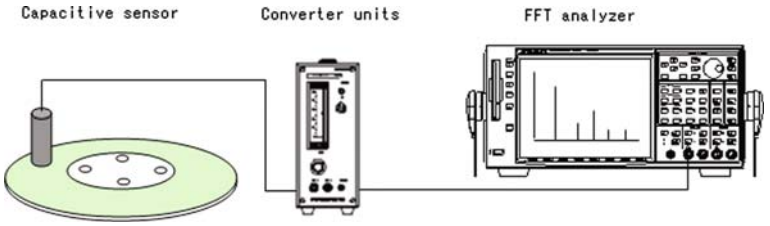


FIGURE 8 Application for spindle and disk runout.

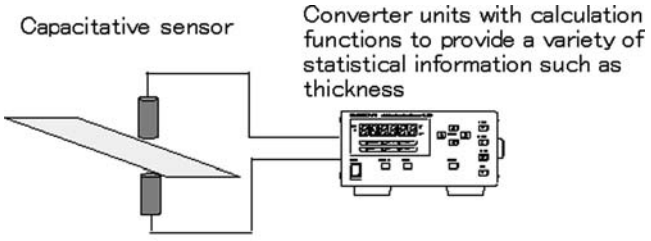


FIGURE 9 Application for measuring thickness.

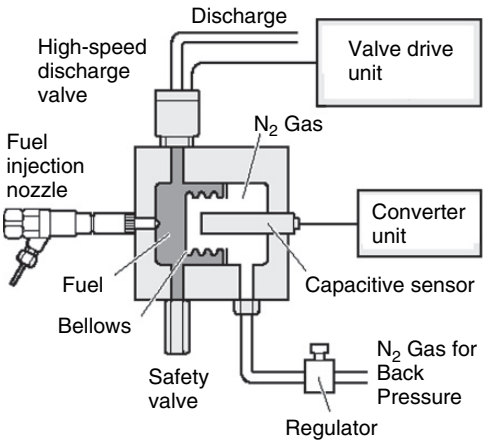


FIGURE 10 Application for measuring fuel amount of diesel engine.

method provides a solution for a dynamic measurement of spindle and disk vibration for hard disk drives, etc. Figure 9 shows an example of non-contact thickness measurement of silicon wafers. As the other silicon applications, deflection, warping, and flatness are measured by capacitive sensors. The measuring point shall be scanned for detecting the variation of displacement in these applications [3]. Figure 10 shows application of fuel amount measurement for a diesel engine. When fuel is injected inside the vessel with the bellows, the bellows gets longer proportionally to the amount of fuel. Thus, it is possible to determine the amount of fuel by detecting the displacement of bellows. The Zeuch method is currently applied for this application [5]. However, the method in which capacitive sensors are available for detecting the displacement of bellows is more accurate than the Zeuch method as a result of high precision and excellent stability of capacitive sensors [4].

References

1. Larry K. Boxer: Capacitive Sensor: Design and Applications, IEEE Press Series on Electronics Technology pp. 67–68
2. http://www.onosokki.net/Pages/ffts.html
3. http://www.adesemiconductor.com/8150.shtml
4. T. Kamimoto: Measurement of Rate of Multiple-Injection in CDI Diesel Engines, SAE Paper, 2000-01-1257.2000
5. W. Zeuch: Neue Verfahren zur Messung des Einspritzgesetzes und der Einspritz Regelmassigkeit von Diesel Einspritzpumpen MTZ Jahrg.22 Heft 9 september 1961

PART XIII ENGINEERING ACOUSTICS

Yoshio Yamasaki

Waseda University, Tokyo, Japan

73 Energy Saving Transducers and Thermoacoustics

Yoshio Yamasaki 1349

Introduction • Challenges to Improving Dynamic Speakers • Superconductive Loudspeakers • Electrostatic Loudspeakers • One-Bit Driving • Thermoacoustics • Summary

74 Sinusoidal Representation of Acoustic Signals

Masaaki Honda 1365

Introduction • Phase Vocoder • Sinusoidal Model • Sinusoidal Model for Rapidly Changing Signal • Time-Scale and Pitch Modification

75 Room Transfer Function

Mikio Tohyama 1381

Introduction • Modal Expansion of the TF • Distribution of Eigenfrequencies • Chaotic Properties of Ray Tracing and Eigenfrequency Distribution • Distribution of Zeros • Phase Response of Room TF • Control of Magnitude and Phase for Artificial Reverberation Rendering

76 Spatial Information of Sound Fields

Yasuhiro Oikawa 1403

Introduction • DOA Estimation • Spatial Information Measurement: Four Closely Located Point Microphones Method • Sound Field Measurement Using a Laser • Summary

73

Energy Saving Transducers and Thermoacoustics

Yoshio Yamasaki

Waseda University, Tokyo, Japan

1 • Introduction

There are many problems concerning energy and the environment. It is important to realize energy saving and to use a little energy gap. This section describes how to make energy saving transducers and how to use the sound for energy sources.

The popular cone type of dynamic loudspeaker has evolved over the years in terms of the materials used for their magnets and cones, but the basic structure has not changed since 1925 when C. E. Rice and E. W. Kellog created them. In fact, it is said that "excellent" transducers are available in 500 million different sizes in Japan. Yet, the efficiency of a transducer, including the amplifier, is less than 0.1%, meaning that more than 99.9% of the energy is wasted [1]. This section presents the test manufacture of edgeless, damperless, superconductive loudspeakers and high-efficiency electrostatic loudspeakers driven by 1-bit switching amplifiers [2–4]. The measured speaker and amplifier efficiencies were 30% and 80%; thus, the total efficiency was 24%.

Thermoacoustic phenomena have been researched for a long time [5–11]. For example, the Sondhauss tube and the Rijke tube, which illustrate fundamental thermoacoustical phenomena,

are described in "The Theory of Sound" by Lord Rayleigh [12]. Sound is a rarefactional wave in air. Therefore, noise can be considered to be a form of sound energy. According to thermoacoustic theory, noise can be exchanged for thermal energy and thus be controlled. In addition, thermal energy can be extracted from noise. There are many applications of thermoacoustic theory, such as Stirling engines. A heat supply system whereby noise is an energy source is a unique idea. Because noise is usually absorbed by some substance before it is finally exchanged for thermal energy, such a system never adds heat loads to the global environment. In contrast, noise is usually discharged to an external field. It is also very unusual that noise energy is actively consumed by its being exchanged. Noise is an unusual type of energy because its wastage can be encouraged. When noise, which overflows in present-day environments, is exchanged, its energy can be used as a heat supply for purposes such as air-conditioning. Hence, noise can be viewed as an alternative energy source.

2 • Challenges to Improving Dynamic Speakers

The structure of a cone dynamic speaker is shown in Figure 1. There are two types of resistance: the electrical resistance of the voice coil and the mechanical resistance of the edge and damper that holds the cone. Because of them, a large proportion of the energy supplied to the speaker is wasted. The efficiency (input electric power to output acoustic power), which actually reached a few percent during the "tube age", is now less than 1% because of the use of transistors and much smaller speakers. Moreover, although the efficiency of amplifiers was improved by the introduction of semiconductors, it is still only about 10%, meaning that the total efficiency is less than 0.1%.

The efficiency of cone dynamic speakers can be improved by minimizing the resistance, reinforcing the magnetic circuit, enlarging the diaphragm, and improving the algorithm. The edge

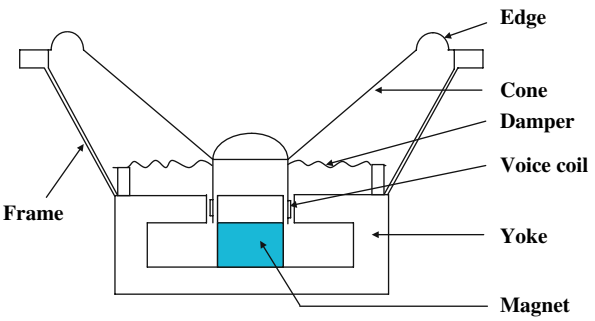


FIGURE 1 Dynamic speaker. (From Ref. [1].)

and the damper also cause distortion. In particular, the edge's movement causes the cone to deteriorate and this shortens the speaker's lifespan.

3 • Superconductive Loudspeakers

3.1 Peculiar Properties of the Superconductive State

Superconductivity has the following qualities:

1. zero electrical resistance and ceaseless current flow,
2. perfect diamagnetism (the Meissner effect),
3. flux quantization whereby the magnetic field is trapped and objects magnetically float (the pinning effect),
4. the Josephson effect whereby current flows with 0 V.

These special qualities have been used in many novel devices, such as single flux quantum (SFQ) devices, superconducting quantum interference devices (SQUIDs), and Josephson junctions.

The superconductive state was once thought to be limited to metals and alloys like mercury and Niobium cooled to liquid helium temperature (4 K = –269°), i.e., very close to absolute zero. However, experiments by great experimentalists since the 1980s have demonstrated that superconductive states can be achieved relatively easily at liquid nitrogen temperatures (77 K = –196°).

3.2 Composition of Superconductive Loudspeakers

By using superconductive materials, we can fabricate strong magnetic circuits of high efficiency through the removal of electrical resistance within the voice coil and edgeless, damperless speakers exploiting the pinning effect. A great energy saving can also be expected by connecting such speakers to a switching amplifier.

Figure 2 shows the composition of an edgeless, damperless inductive superconductive loudspeaker that we have trial manufactured. By using induction as opposed to directly connecting the voice coil, we could make the cone free of electrical wiring, edges, and dampers, and make the cone float in

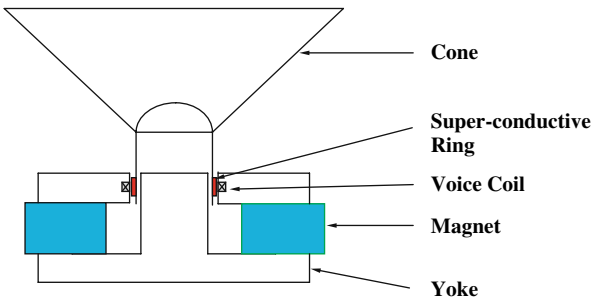


FIGURE 2 Superconductive speaker. (From Ref. [1].)

air while in use. We used a simple apparatus that holds it until it becomes superconductive and lets it go once it has cooled.

For the superconductors, the Y-Ba-Cu-O rare earth single crystal superconductive bulk material QMG (manufactured by Nippon Steal Corporation) was processed into a ring form with



FIGURE 3 Picture of superconductive speaker. (From Ref. [1].)

25.6 mm inner and 29.6 mm outer diameters and a height of 3.0 mm, and used it as one-turn voice coils.

Figure 3 shows a picture of the superconductive speaker.

3.3 Equivalent Circuit

The equivalent circuit of the inductive speaker in a superconductive state is shown in Figure 4. The speaker constants are as follows:

primary voice coil resistance: Rc1 = 0–5Ω
primary voice coil inductance: L1 = 0–11 mH
secondary coil inductance: L2 = 0.01 mH
coupling coefficient: k = 0.5–1
resonance frequency: F0 = 153.5 Hz
resonance Q factor: Qmo = 3.60
mass of movement: Md = 2.1 g
force coefficient: B1 = 0.08 Tm
radius of cone: a = 22.5 mm

As can be seen from the equivalent circuit, when the primary coil is also superconductive and the coupling coefficient k between the primary and the secondary is 1, all of the electrical input becomes a load; thus the output sound pressure level rises dramatically in comparison with conventional speakers, and it should keep on rising according to the frequency rise without concern for the movement requirement.

Figure 5(a) shows the superconductive speaker's frequency characteristics for the electrical impedance and output sound pressure level measured 1 m from the speaker for the input

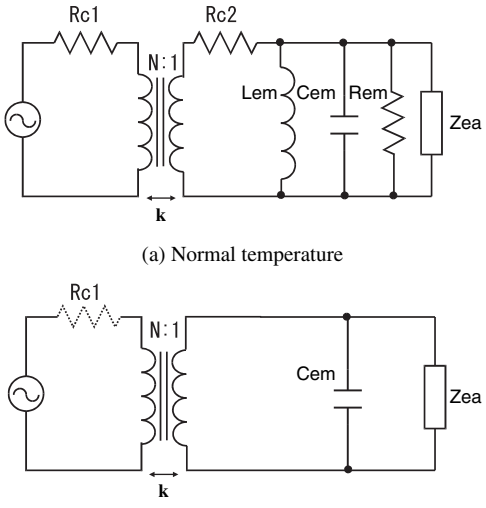


FIGURE 4 Equivalent circuit. N is the turns of the primary voice coil, L_{em} is compliance of movement, C_{em} is mass of movement, R_{em} is mechanical resistance and Z_{ea} is radiation impedance. (a) Normal temperature (b) Superconducting temperature. (From Ref. [1].)

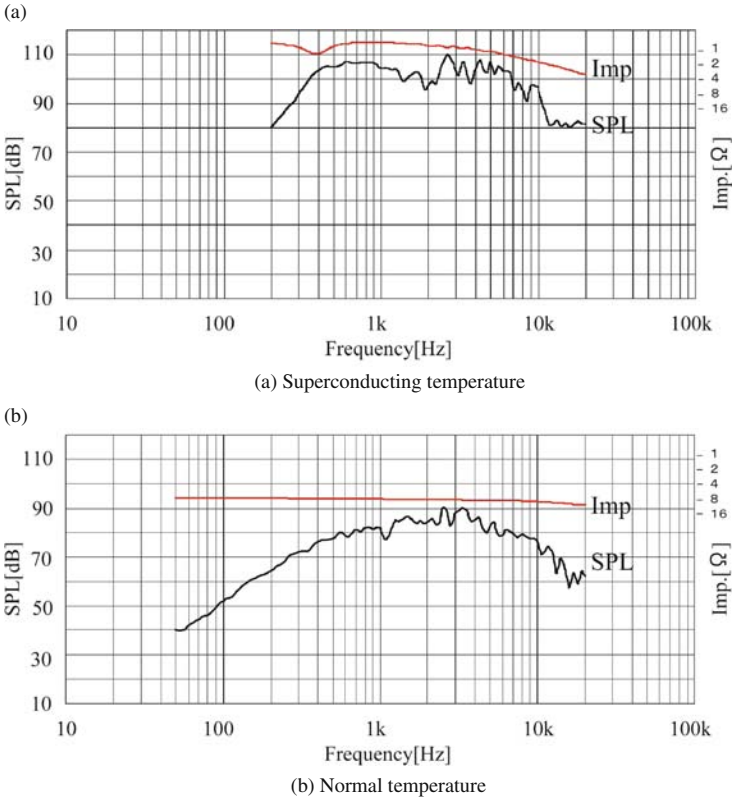


FIGURE 5 Frequency characteristics. (a) Superconducting temperature. (b) Normal temperature. (From Ref. [1].)

level of 1 W. Figure 5(b) shows the frequency characteristics at normal temperatures when the secondary voice coil is one turn of aluminum.

In the middle sound range under the above conditions, 110 dB of output sound pressure level was obtained. This is very high efficiency, more than 20 dB compared to a cone-type speaker of the same size.

The superconductive loudspeaker driven by a 1-bit switching amplifier reaches about 20% efficiency in the middle-range sound frequencies.

4 • Electrostatic Loudspeakers

4.1 Loudspeaker Composition

Electrostatic loudspeakers produce sound by applying a bias voltage between two parallel electrodes facing each other and use the input signal to change the aspiration force between the electrodes [13]. Figure 6 shows the structure of a single-

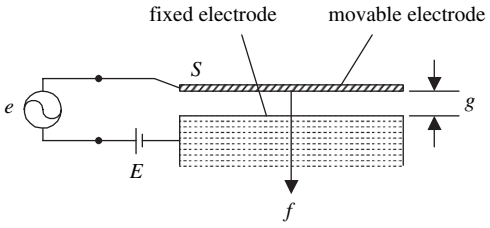


FIGURE 6 Single-ended electrostatic loudspeaker. (From Ref. [1].)

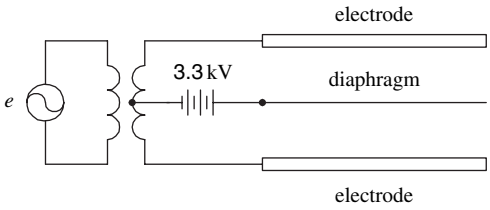


FIGURE 7 Push–pull type electrostatic loudspeaker. (From Ref. [1].)

ended electrostatic loudspeaker. The changes in aspiration force between the two electrodes f [N] are expressed as follows.

$$f \approx \frac{\epsilon_0 SE}{g^2} e = \frac{CE}{g} e = Ke \tag{1}$$

C is the capacitance between the electrodes and K is a force coefficient. Therefore, the aspiration force is proportional to the area of the electrodes S and bias voltage E, and is inversely proportional to the square of the distance between the electrodes g.

The electrostatic speaker we test manufactured is a push–pull type (Figure 7).

4.2 Transducer Efficiency

Electrostatic speakers are said to be inefficient transducers. However, as one can see from the equivalent circuit in Figure 8, there is no DC resistance in the voice coil, just as in the superconductive speaker. Thus, theoretically speaking, such as transducer has got to be efficient.

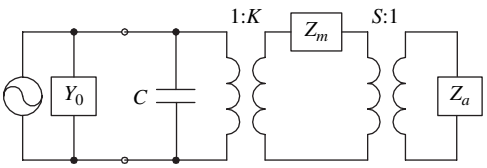


FIGURE 8 Equivalent circuit of electrostatic loudspeaker. Y_0 is internal admittance of power supply, Z_m is mechanical impedance, and Z_a is acoustic impedance. (From Ref. [1].)

Electrostatic speakers do not consume electrical power, and most of the power becomes reactive. We expect the speaker will be very efficient when it is connected to a negative-feedback constant-current 1-bit amplifier.

5 • One-Bit Driving

It is effective to induce the switching circuit to make amplifiers of high efficiency. The power loss of the switching amplifier is mainly because of the internal resistance and the switching loss because of the gate capacitance and response delay. Therefore, besides improving the power device, it is also possible to reduce power loss by reducing the number of switching times. We tried to make a high S/N amplifier using the usual switching device by using "Run Length Limited Coding" in the sigma-delta modulator [14,15] so that high resolution in time domain can be acquired in the conditional pulse width [16]. As a result, a simulation has shown that a switching amplifier with over 100 dB S/N in the audible band should be possible when the pulse width is more than 1μs.

6 • Thermoacoustics

Swift [17] believes it is very likely that practical uses will be found for thermoacoustic engines, because these engines have the advantages of reasonable efficiency and extreme simplicity. Another advantage is that these engines can transfer heat without using chlorofluorocarbons.

Swift pointed out that there are two classes of heat engine: prime movers and heat pumps. In a heat pump, work is absorbed by the engine, resulting in heat being pumped from a low-temperature region to a high-temperature region. Thus, a source that produces work such as a loudspeaker or a heater is required for a thermoacoustic resonator, a kind of thermoacoustic heat pump, to operate.

6.1 Simplified Model of Thermoacoustic Heat Pump

Sound waves propagating without obstacles cause adiabatic changes in the oscillating fluid, whereas sound waves progressing through narrow paths exchange heat between the solid walls and the fluid. The sound wave can transfer heat through such heat exchange with solid walls.

Figure 9 shows the heat flux in a thermoacoustic resonator [18]. The standing wave arises in the resonator through sound waves entering from the left end of the resonator. The resonance

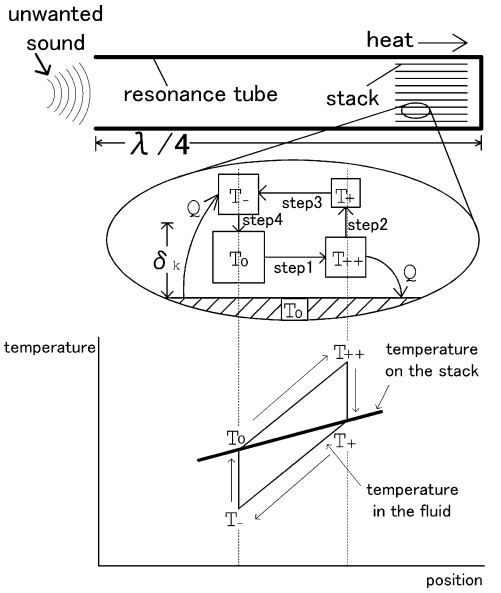


FIGURE 9 Schematic diagram of a quarter wavelength thermoacoustic refrigerator. (From Ref. [1].)

frequency, which corresponds to a quarter wavelength of the sound, equals the resonator length.

Near the closed end of the resonator, there are stacks whose spacing is a few thermal penetration depths. The thermal penetration depth δ_k represents the distance over which heat will diffuse through the fluid during a period $1/\omega$, and is defined by $\delta_k = \sqrt{2k/\omega}$, where $k = K/\rho_m c_p$ is the fluid's thermal diffusivity, ω is the angular frequency, K is its thermal conductivity, ρ_m is its mean density, and c_p is the isobaric specific heat per unit mass.

This length scale is crucial to understanding the performance of the thermoacoustic cycle because diffusive heat transport between the fluid and stack is only significant within this scale. Thus, the distance from the solid stack material is small enough that a substantial amount of thermal conduction can occur in an amount of time $1/\omega$.

Figure 9 shows a magnified portion in the middle of the stack, and a parcel of fluid undergoing an acoustic oscillation. The four steps in the oscillation cycle are presented in four boxes, which are shown as a rectangular path for clarity.

The temperature of the stack at the left-most position of the oscillating fluid parcel's excursion is therefore $T_\mathrm{m} - x_l \nabla T$, and $T_\mathrm{m} + x_l \nabla T$ at the right-most excursion, where T_m is the mean temperature on the stack, and ∇T is the mean temperature gradient. The fluid is transported along the stack a distance $2x_l$, and is heated from a temperature of $T_\mathrm{m} - x_l \nabla T$ to $T_\mathrm{m} - x_l \nabla T + 2T_l$ by

adiabatic compression (step 1), where T_l is the adiabatic temperature change of the fluid. The warmer fluid parcel transfers an amount of heat dQ_{hot} to the stack by thermal conduction at a constant pressure and its temperature decreases to that of the stack $T_{\text{m}} + x_l \nabla T$ (step 2). The oscillating fluid is transported back along the stack to position $-x_l$, and is cooled by adiabatic expansion to a temperature $T_{\text{m}} + x_l \nabla T - 2T_l$ (step 3). The fluid parcel absorbs an amount of heat dQ_{cold} from the stack. This raises the fluid's temperature back to its original value $T_{\text{m}} - x_l \nabla T$ (step 4).

Thus, this system forms complete cycles back to the original situation, and the overall heat pumping process is analogous to a "bucket brigade" in which each fluid parcel picks up heat from its neighbor on the left at a lower temperature and hands over the heat to its neighbor on the right at a higher temperature.

6.2 Experimental Apparatus

To optimize the constitution of the stack, the following experiments were performed [19]. The resonator was a 32-mm-diameter, 102-mm-long acrylic tube connected to a loudspeaker through a conoid tube. The right end of the resonator was covered by a plastic plate. The resonance occurred at 88 Hz, an operating frequency that corresponds to a resonator length of a quarter wavelength of the sound, and the sound pressure distribution (shown in Figure 10) was formed in the resonator by trailing a condenser type microphone (6-mm diameter) along the axis of the resonator.

The stack consisted of a bundle of 50-mm-long polyimide tubes. The temperature distribution in the resonator was measured by inserting a type K thermocouple probe (ϕ 1.5 mm) at 88 Hz. The measurements were performed for the three types of polyimide tube shown in Figure 11. The diameters of the tubes were 0.8, 1.0, and 1.5 mm. The lowest temperature shown in Figure 12 was measured when 1.0-mm-diameter tubes were set in the resonator.

6.3 Experiment on Railroad Noise

An experiment in which the sound source was railroad noise instead of a sinusoidal sound source was performed [20]. Figure 13 shows the frequency distribution of the railroad noise.

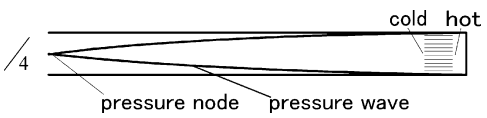
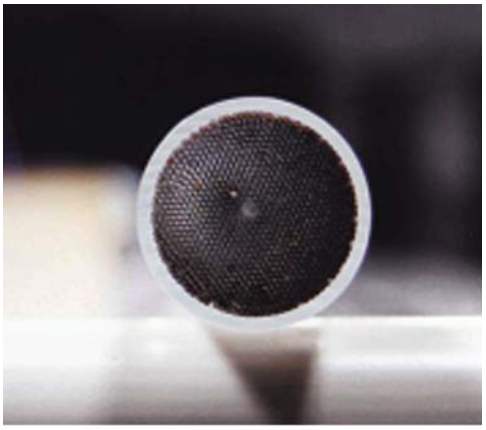
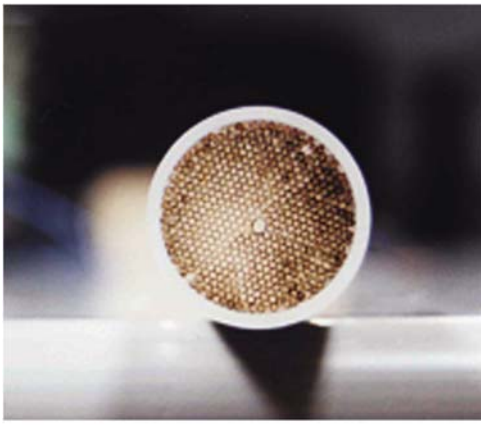


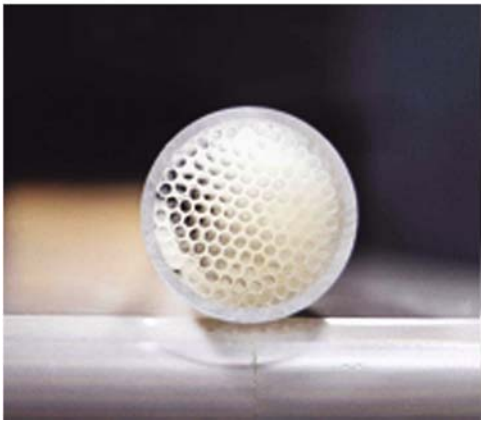
FIGURE 10 Sound pressure distribution in the thermoacoustic tube. (From Ref. [20].)



(a) ϕ 0.8 mm



(b) ϕ 1.0 mm



(c) ϕ 1.5 mm

FIGURE 11 Cross section of the stack. (From Ref. [20].)

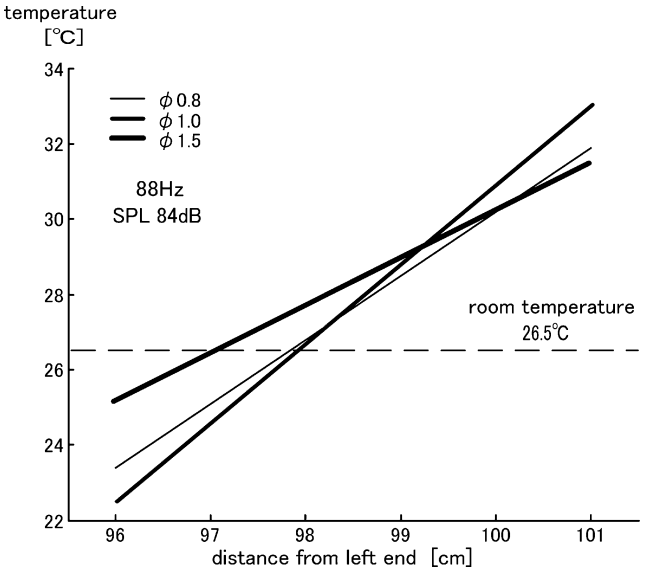


FIGURE 12 Temperature distribution around the stack. (From Ref. [20].)

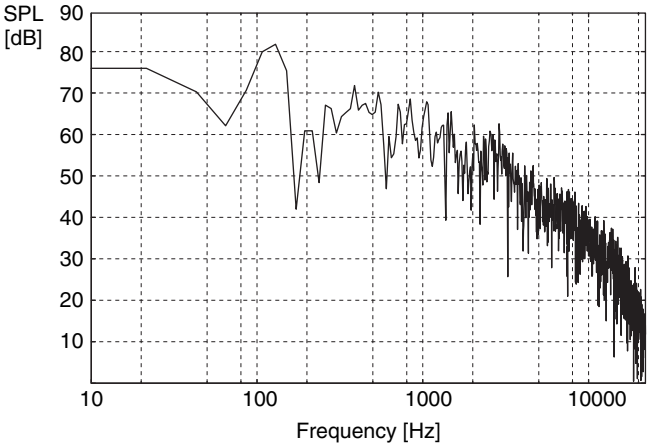


FIGURE 13 Frequency distribution of railroad noise used in the experiment. (From Ref. [20].)

The noise was recorded at the connection chamber between train cars. Note that there is a dominant component at 135 Hz in Figure 13. The experiment shown in Figure 14 was therefore performed with a 610-mm-long tube that corresponded to a

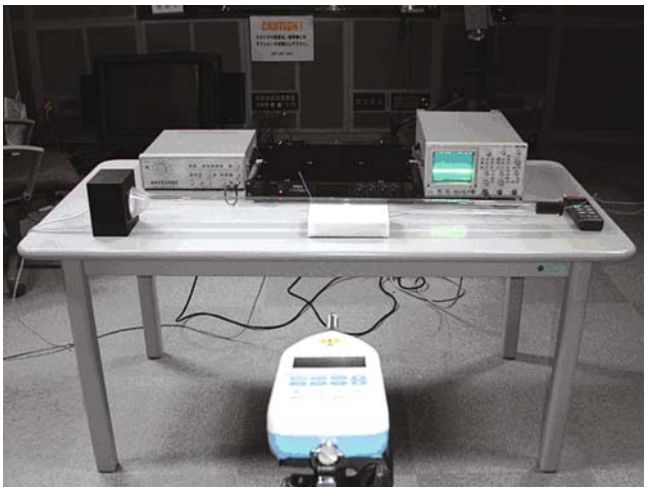


FIGURE 14 View of the experimental apparatus. (From Ref. [20].)

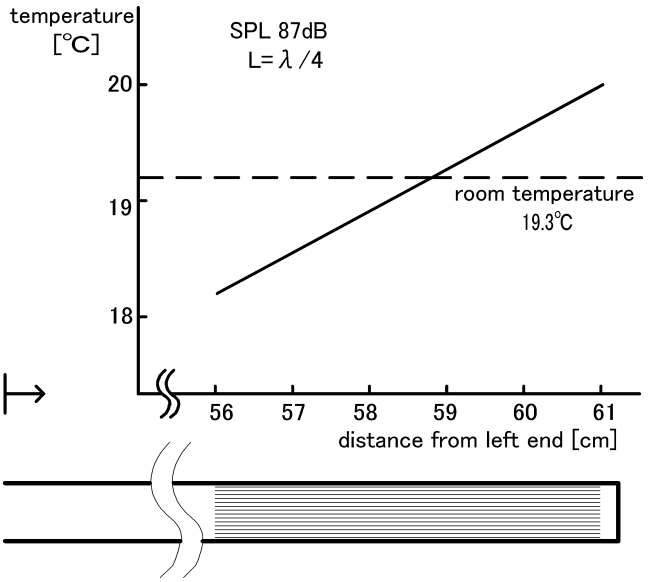


FIGURE 15 Temperature distribution around stack in thermoacoustic system using railroad noise. (From Ref. [20].)

quarter wavelength of sound at 135 Hz. Figure 15 shows the results. A temperature of 18.3°, lower than the room temperature of 19.3°, was measured.

The electric exchange system for environmental noise was thus shown to be possible, although the efficiency of exchange was severely limited.

7 • Summary

The electric-to-acoustic transducers of popular small cone dynamic loudspeakers have surprisingly low efficiencies; a pair of speakers, together with amplifiers, reaches only about 0.1% efficiency. In contrast, the test manufactured edgeless, damperless superconductive loudspeaker with a 1-bit switching amplifier reaches about 20% efficiency in the middle-range sound frequencies. Also, the electrostatic loudspeaker with a 1-bit switching amplifier reaches 30% efficiency. The electrostatic speaker runs on 8 A batteries and can withstand a continuous drive for 3 days.

This section also described aspects of thermosacoustics, measurements on three types of polyimide tube to optimize the constitution of the stack, and a thermoacoustic system whereby railroad noise can be used as an energy source. An electric exchange system from environmental noise was also constructed [21].

References

1. Y. Yamasaki, H. Nakajima, M. Okazaki and Y. Kato, "Transducers considering energy saving –super-conductive and electrostatic loudspeaker driven by 1 bit Amplifier–," *In Proc. Int. Cong. Acoust.*, vol. IV, pp. 2427–2432, 2004.
2. H. Nakajima and Y. Yamasaki, "Edgeless damper-less superconductive loudspeaker," *J. Acoust. Soc. Japan*, **59**(10), 634–638, 2003 (in Japanese).
3. Y. Yamasaki and H. Nakajima, "Prototype edgeless loudspeaker by pinning of superconductor," *In Proc. Spring Meeting Acoust. Soc. Japan*, pp. 517–518, 2003 (in Japanese).
4. Y. Yamasaki, H. Nakajima, I. Chatani and K. Kitazawa, "Superconductive speaker," *Documents of 50th Anniversary Seminar of Japan Audio Soc.*, pp. 1–8, 2001 (in Japanese).
5. A. Tominaga, *Fundamental Thermoacoustics* (Uchida Rokakuho Publishing Co., Ltd., Tokyo, 1998) (in Japanese).
6. T. Yazaki, A. Iwata, T. Maekawa and A. Tominaga, "Traveling wave thermoacoustic engine in a looped tube," *Phys. Rev. Lett.*, **81**, 3128–3131, 1998.
7. T. Yazaki, T. Biwa and A. Tominaga, "A pistonless Stirling cooler," *Appl. Phys. Lett.*, **80**, 157–159, 2002.

8. U. Ueda, T. Biwa, U. Mizutani and T. Yazaki, "Experimental studies of a thermoacoustic Stirling prime mover and its application to a cooler," *J. Acoust. Soc. Am.*, **115**, 1134–1141, 2004.
9. S. Backhaus and G. W. Swift, "A thermoacoustic Stirling heat engine," *Nature*, **399**, 1134–1141, 2004.
10. Y. Tashiro, T. Biwa and T. Yazaki, "Calibration of a thermocouple for measurement of oscillating temperature," *Rev. Sci. Instrum.*, **76**, 124901, 2005.
11. G. W. Swift, "Analysis and performance of a large thermoacoustic engines," *J. Acoust. Soc. Am.*, **92**, 1551–1563, 1992.
12. L. Rayleigh, *The Theory of Sound* (Dover Publications, Dover 1896).
13. M. Okazaki, Y. Oikawa and Y. Yamasaki, "Distance learning and conference system using electrostatic loudspeakers," *In Proc. Autumn Meeting Acoust. Soc. Japan*, pp. 571–572, 2003 (in Japanese).
14. J. Ohga, Y. Yamasaki and Y. Kaneda, *Acoustic System and Digital Processing* (Korona Publishing Co., Ltd., Tokyo, 1995).
15. Y. Oikawa, Y. Kimura, Y. Yamasaki and J. Fujimori, "Recording system of high-speed 1bit signal and HD video using IEEE1394," *In Proc. AES 12th Regional Convention Tokyo*, pp. 128–131, 2006.
16. S. Takeoka, M. Kurihara, M. Okazaki, Y. Oikawa, M. Nishikawa and Y. Yamasaki, "1-bit wave field recording/reproduction system using electrostatic microphone and loudspeaker," *Tech. Rep. IEICE*, **105**(136), 25–30, EA2005-24, 2005.
17. G. W. Swift, "Thermoacoustic engines," *J. Acoust. Soc. Am.*, **84**(4), 1145–1180, 1988.
18. S. L. Garrett and T. J. Hofler, "Thermoacoustic refrigeration," *ASHRAE J.*, **34**(12), 28–36, 1992.
19. R. Yoonsun, M. Okazaki, Y. Oikawa, Y. Tokita and Y. Yamasaki, "Thermoacoustic system driven by acoustic energy," *In Proc. Autumn Meeting Acoust. Soc. Japan*, pp. 569–570, 2003 (in Japanese).
20. Y. Tokita, R. Yoonsun, Y. Oikawa and Y. Yamasaki, "Construction of thermal and/or electric exchange system whereby environmental noise is considered to be an energy source," *In Proc. Int. Cong. Expo. Noise Control Eng.*, 2004.
21. Y. Tokita, R. Yoonsun, Y. Oikawa and Y. Yamasaki, "Construction of electric exchange system whereby environmental noise is considered to be an energy source," *In Proc. Int. Cong. Exp. Noise Control Eng.*, 2005.

74

Sinusoidal Representation of Acoustic Signals

Masaaki Honda

Waseda University, Tokyo, Japan

1 • Introduction

Sinusoidal representation of acoustic signals has been an important tool in speech and music processing like signal analysis, synthesis and time scale or pitch modifications. It can be applicable to arbitrary signals, which is an important advantage over other signal representations like physical modeling of acoustic signals. In sinusoidal representation, acoustic signals are composed as sums of sinusoid (sine wave) with different amplitudes, frequencies and phases, which is based on the time-dependent short-time Fourier transform (STFT). This article describes the principles of acoustic signal analysis/synthesis based on a sinusoid representation with focus on sine waves with rapidly varying frequency.

2 • Phase Vocoder

The phase vocoder [1] models signal in each channel of filter a bank as a sinusoidal component for analysis and synthesis, which is derived from the STFT of the signal $x(t)$

$$\begin{aligned} X(\omega,t) &= \int_{-\infty}^{\infty} x(\tau)w(\tau-t)\mathrm{e}^{-j\omega\tau}\mathrm{d}\tau \\ &= e^{-j\omega t}\int_{-\infty}^{\infty} x(t-\tau)w(\tau)\mathrm{e}^{-\mathrm{j}\omega\tau}\mathrm{d}\tau, \end{aligned} \tag{1}$$

where $w(t)$ is a window function [2]. The STFT is interpreted in two ways: a running Fourier transform of the windowed signal $x(\tau)w(\tau-t)$ and band-pass filtering with $h(\omega,t)=w(t)\mathrm{e}^{-j\omega t}$ followed by demodulation to the baseband. In the latter interpretation, the filter-bank output is represented by

$$Y(\omega,t)=\mathrm{e}^{j\omega t}X(\omega,t). \tag{2}$$

Then, the signal $x(t)$ is expressed as a linear combination of basis function $\xi(\omega,t)=h(\omega,t)e^{j\omega t}$ by

$$x(t)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Y(\omega,\tau)\xi(\omega,\tau-t)\ \mathrm{d}\omega\mathrm{d}\tau, \tag{3}$$

provided that $\int_{-\infty}^{\infty} w^2(t)\,\mathrm{d}t=1$. Therefore, $Y(\omega,t)$ can be regarded as the coefficients of the basis functions $\xi(\omega,t)$ [3].

For a discrete signal $x(n)$ and a discrete band of filters with equal frequency spacing, the band-pass filter function is represented as

$$h_k(n)=w(n)\mathrm{e}^{j\frac{2\pi kn}{R}}, \tag{4}$$

where $0\le n\le M$, M being the duration of the impulse response and R is the number of filters. Then, the signal $x(n)$ is reconstructed as a sum of the filter-bank outputs $Y_k(n)$ by

$$\begin{aligned} x(n) &= \sum_{k=0}^{R-1} Y_k(n) \\ &= \sum_{k=0}^{R-1} r_k(n)e^{j\phi_k(n)}, \end{aligned} \tag{5}$$

if the duration is less than the number of filters, $M<R$ [4]. Here, $r_k(n)$ and $\phi_k(n)$ are the amplitude and the phase of each filter output. This signal analysis and synthesis structure is referred to as the phase vocoder [1], which is shown in Figure 1.

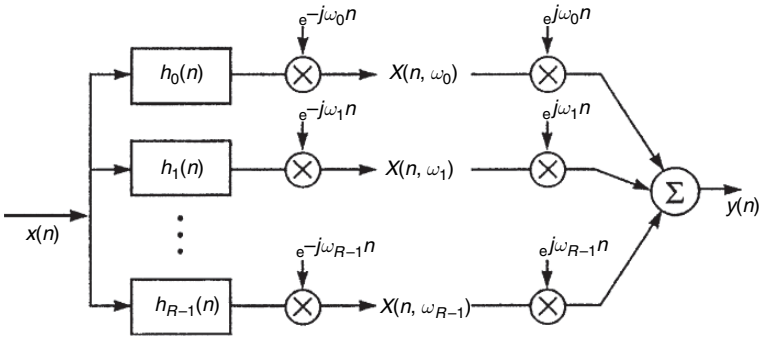


FIGURE 1 Phase vocoder based on short-time Fourier analysis and synthesis. (From T.F.Quatieri and R. J. McAulay, *Application of Digital Signal Processing to Audio and Acoustics*, edited by M. Kahrs and K. Brandenburg, Kluwer Academic Publishers, 1998.)

Denoting the magnitude and the phase of the discrete STFT $X_k(n) = e^{-j\omega_k n} Y_k(n)$ by $|X_k(n)|$ and $\theta_k(n)$, the reconstructed signal is written by

$$x(n) = \sum_{k=0}^{R-1} |X_k(n)| \, e^{j[\omega_k n + \theta_k(n)]}. \tag{6}$$

Also, the signal is written as a sum of real sine waves by

$$x(n) = \sum_{k} 2 \, |X_k(n)| \cos[\omega_k n + \phi_k(n)]. \tag{7}$$

3 • Sinusoidal Model

Sinusoidal models [5–7] represent the audio signal by harmonic components that are sine waves with time-varying magnitude and phase, which is given by

$$x(n) = \sum_{i=1}^{N} r_i(n) \cos(\theta_i(n)), \tag{8}$$

where $r_i(n)$ is a time-varying amplitude of the ith sinusoid and $\theta_i(n)$ is the phase. The phase is written by

$$\theta_i(n) = \int_0^{nT} \lambda_i(\tau) d\tau + \phi_i, \tag{9}$$

where $\lambda_i(n)$ is the instantaneous frequency (IF) and ϕ_i is the initial phase.

In the sinusoidal analysis, the amplitudes, frequencies and phases of the model are estimated from the input signal at every analysis frame, and in the synthesis, these parameters are interpolated so as to maintain the continuity on successive frames.

If the signal is quasi-periodic, the DFT spectrum shows multiple peaks at each harmonic frequency. The magnitude, the frequency and the phase of each sinusoid are determined directly by tracking the local peaks of the DFT magnitude spectrum [5]. Figure 2 shows an example STFT spectrum and

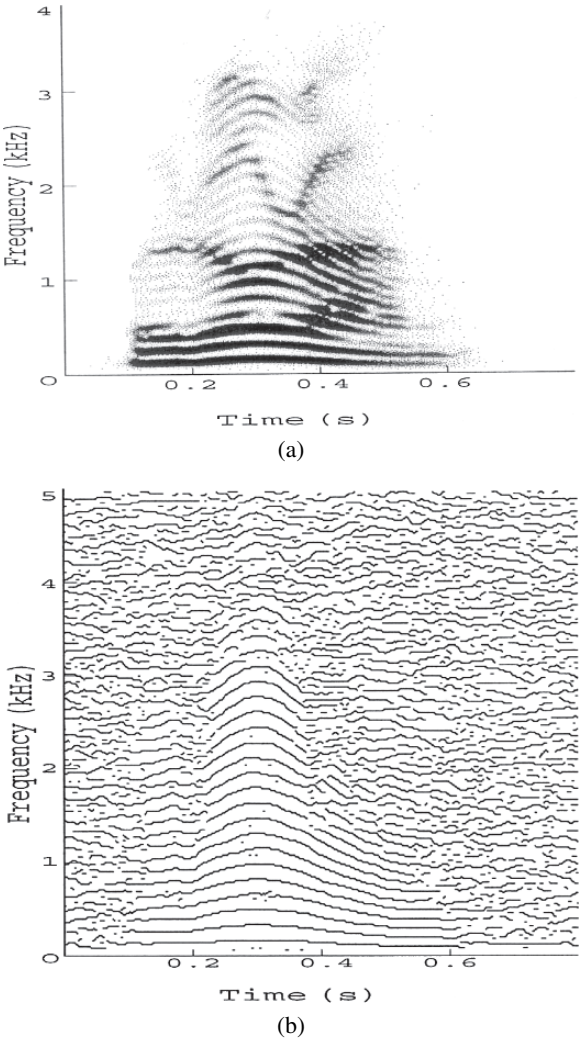


FIGURE 2 (**a**) STFT spectrogram and (**b**) trajectories of peaks on STFT spectrogram from a speech signal. (From [8].)

sinusoidal frequency for a rapidly varying speech signal. The location and number of the peaks, however, will change as the pitch of the signal rapidly changes. In order to obtain a continuous and slowly changing sequence of magnitude and frequency (sinusoidal trajectory), sinusoidal tracking based on a frequency-matching algorithm over successive frames is needed. A typical birth–death detection algorithm [5] finds the sinusoidal frequencies successively by using a "nearest neighbors" rule while allowing for the birth and the death of the sinusoidal trajectory.

In the synthesis stage, the magnitudes and phases of the sinusoidal trajectory are interpolated. Denoting the ith sinusoidal parameters for frame k by $(r_i^k, \lambda_i^k, \phi_i^k)$, the amplitude interpolation is simply given by

$$\hat{r}_i(n) = r_i^k + \frac{(r_i^{k+1} - r_i^k)}{(n/T)}, \tag{10}$$

where $n = 0, 1, \cdots, T-1$ is the time sample in the kth frame.

The phase is interpolated in a cubic polynomial form as

$$\hat{\theta}_i(n) = \theta_i^k + \lambda_i^k n + \alpha n^2 + \beta n^3 + 2\pi M, \tag{11}$$

where the term $2\pi M$, M an integer, is used to account for the phase unwrapping since the principal value of the phase is measured. The polynomial parameters and the unwrapping factor are determined so that the cubic function and its time derivative (IF) at $n = T$ are equal to the measure values at frame $k+1$. Thus, the cubic parameters α, β are determined as a solution of the linear equation,

$$\begin{aligned} \hat{\theta}_i(T) &= \theta_i^k + \lambda_i^k T + \alpha T^2 + \beta T^3 = \theta_i^{k+1} + 2\pi M \\ \hat{\dot{\theta}}_i(T) &= \lambda_i^k T + 2\alpha T + 3\beta T^2 = \lambda_i^{k+1}. \end{aligned} \tag{12}$$

The phase unwrapping parameter $M$is determined to make the unwrapped phase maximally smooth at $n = T$.

Finally, using the interpolated amplitude and unwrapped phase functions, the signal is synthesized by

$$\hat{x}(n) = \sum_{i=1}^{N} \hat{r}_i(n) \cos(\hat{\theta}_i(n)). \tag{13}$$

The sinusoidal analysis and synthesis system is shown in Figure 3.

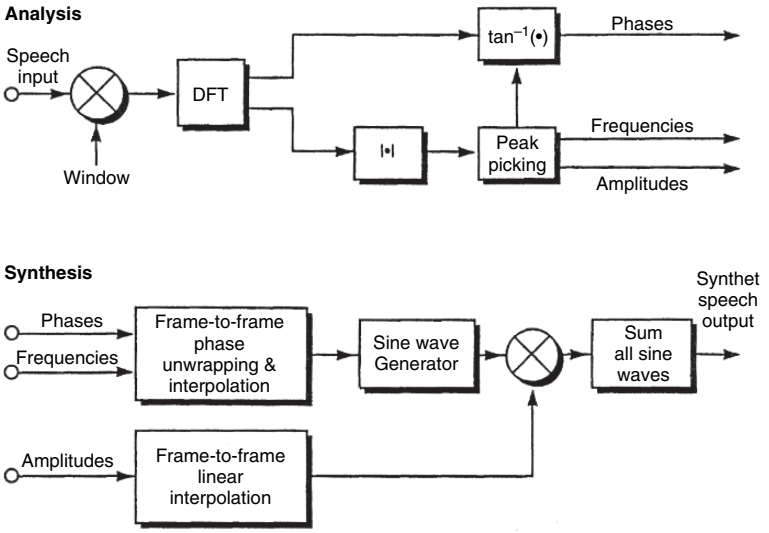


FIGURE 3 Block diagram of sinusoidal analysis and synthesis. (From [5].)

4 • Sinusoidal Model for Rapidly Changing Signal

Efficiency of the sinusoidal model depends on the estimation accuracy of the sinusoidal frequencies and phases. There are two causes that degrade the estimated sinusoidal frequencies. One is inaccurate frequency estimation for an extracted sinusoidal component. The other one is the interference from other components, which is due to insufficient resolution in time and frequency. In this section, two methods that improve the estimation accuracy of the sinusoidal frequencies are described.

4.1 Instantaneous Frequency Attractor

In the sinusoidal model, the IF is defined as the time derivative of the phase of the filter-bank output by

$$\lambda(\omega,t) = \frac{\partial}{\partial t}\arg[Y(\omega,t)]. \tag{14}$$

Avoiding a direct calculation of phase, the IF can be calculated by

$$\lambda(\omega,t) = \frac{a\dfrac{\partial b}{\partial t} - b\dfrac{\partial a}{\partial t}}{a^2+b^2}, \tag{15}$$

where $Y(\omega,t) = a + jb$ [1]. Here, Equation (15) has time derivative of the real and imaginary part of $Y(\omega,t)$.They are obtained from the STFT, and its time derivative is

$$\frac{\partial}{\partial t}Y(\omega,t) = \mathrm{e}^{j\omega t}\frac{\partial}{\partial t}X(\omega,t) + j\omega\mathrm{e}^{j\omega t}X(\omega,t), \tag{16}$$

where the time derivative of the STFT is obtained from the time derivative of the window function $\partial w(t)/\partial t$ by

$$\frac{\partial}{\partial t}X(\omega,t) = \int_{-\infty}^{\infty} x(\tau)\left(-\frac{\partial w(\tau - t)}{\partial t}\right)\mathrm{e}^{-j\omega\tau}. \tag{16}$$

Thus, by using the window function and its time derivative, the IF is directly calculated. This computation method has advantage of not needing approximations such as using the phase difference between successive samples.

The IF attracters correspond to harmonic frequency trajectories. They are defined as the points which satisfy

$$\mu(\omega,t) = 0 \text{ and } \frac{\partial \mu}{\partial \omega} < 0, \tag{17}$$

where $\mu(\omega,t) = \lambda(\omega,t) - \omega$ [8]. This is equivalent to

$$\lambda(\omega,t) = \omega \quad \text{and} \quad \frac{\partial \lambda}{\partial \omega} < 1. \tag{18}$$

The IF attractors are fixed points through the mapping ω to the IF λ. The inequality chooses only cross-points at which the slope of λ is less than 1. Figure 4 shows an example of the relationship between $\lambda(\omega,t_0)$ and $\mu(\omega,t_0)$ when multiple sinusoids exist at the location indicated by the arrows. Figure 5 shows an example of IF attractors for the speech sample in Figure 2. In practical implementation, the IF attractors are determined by finding zero crossings

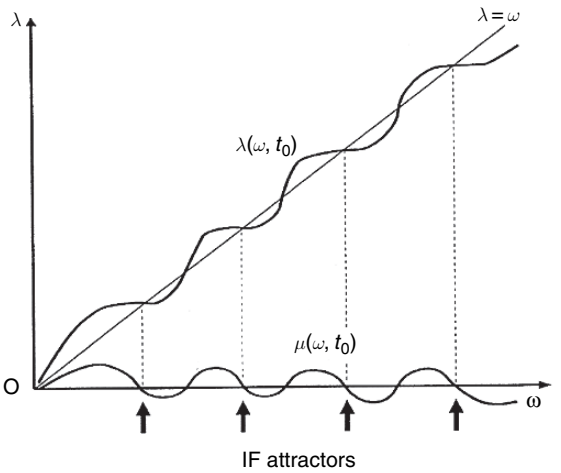


FIGURE 4 Definition of IF attractors. (From [8].)

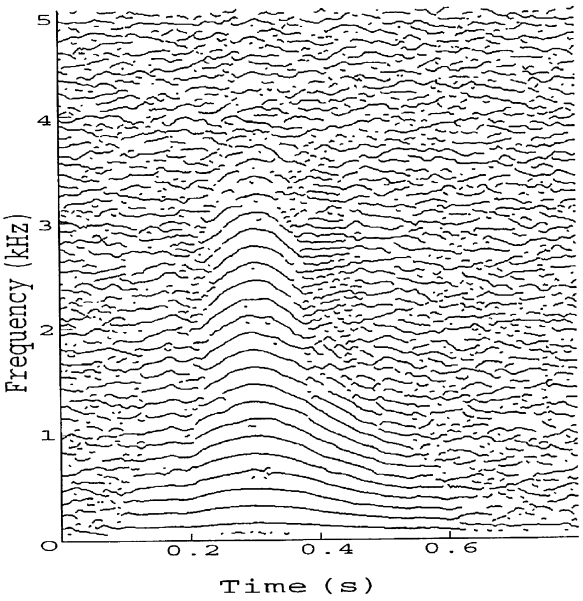


FIGURE 5 Instantaneous frequency (IF) attractors from a speech signal. (From [8].)

of $\mu(\omega,t)$ in Equation (17) by using simple linear interpolation. The IF attractors correspond well to harmonic frequencies for the interval where they are not rapidly changing. When the IF attractors in Figure 5 are compared with STFT spectrum's peaks in Figure 2 for the same speech sample, the STFT peak trajectories look like stepwise contour due to the limited frequency resolution in the STFT. In contrast, the contours of the IF attractors are smooth because the IF takes continuous values.

The IF attractor gives an accurate estimate of a harmonic frequency. The IF attractor, however, will not give a harmonic frequency of either components if the harmonic components are not separated by the filter bank of Equation (2). This happens for rather high-order harmonic frequencies when the sinusoidal frequency changes rapidly. Thus, in Figure 5, the IF attractors do not match harmonic frequencies in the interval where the harmonic frequencies changes rapidly.

4.2 IF Attractor Analysis with Time Warping

When harmonic frequencies change rapidly, it is difficult to extract the higher order harmonic frequencies because the frequency modulation broadens the bandwidth of each harmonic component so that adjacent components interfere. Accuracy of the IF attractors for rapidly changing harmonic frequencies can be improved by introducing a time warping which cancels out

the change in the harmonic frequencies [9]. Let $x(t)$ denote a sinusoid modulated in amplitude and frequency as

$$x(t) = r(t)\cos(\int^{t} \phi(\tau)\mathrm{d}\tau + \phi_0), \tag{19}$$

where $\theta(t) = \int^{t} \phi(\tau)d\tau + \phi_0$ is the signal phase. Here, let u denote a warped time axis, which is mapped from the original axis t by the warping function $u = p(t)$. Also, assume that the phase on the time axis u is represented by constant ψ and ψ_0 as $\theta_u(u) = \psi u + \psi_0$ and the phase on u is equal to that on t, $\theta_u(u) = \theta(t)$. Then, the signal on the time axis u becomes

$$x_u(u) = r_u(u)\ \cos(\psi u + \psi_0), \tag{20}$$

whose frequency is constant. Therefore, an analysis on u suppresses the sideband generated by the frequency modulation. Such a time warping is shown in Figure 6.

The problem is to find the time warping function $u = p(t)$ so that it cancels out the change in the frequency. Since the phases at t and u, corresponding to the same point in time, are identical

$$\theta(t) = \int^{t} \phi(\tau)d\tau + \phi_0 = \psi u + \psi_0 = \theta_u(u), \tag{21}$$

and differentiating Equation (21) with respect to u gives

$$\phi(t)\frac{\mathrm{d}t}{\mathrm{d}u} = \psi. \tag{22}$$

Here, $\phi(t)$ is the IF of the signal. By differentiating again with respect to u, and after some calculus, the following differential

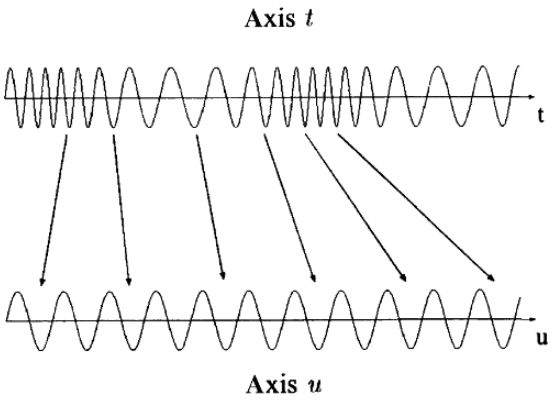


FIGURE 6
Cancellation of frequency variation by time warping. (From [8].)

equation is obtained, which the warping function $u = p(t)$ has to satisfy for a given phase $\theta(t)$ [8].

$$\frac{\mathrm{d}^2 t}{\mathrm{d}u^2}\frac{\partial \theta}{\partial t} + \left(\frac{\mathrm{d}t}{\mathrm{d}u}\right)^2 \frac{\partial^2 \theta}{\partial t^2} = 0. \tag{23}$$

After some calculus, the equation can be written by

$$\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} = q(t)\frac{\mathrm{d}u}{\mathrm{d}t} \tag{24}$$

defining

$$q(t) = \left(\frac{\partial \theta}{\partial t}\right)^{-1} \frac{\partial^2 \theta}{\partial t^2}. \tag{25}$$

This means that if $q(t)$ is given, the $\mathrm{d}^2u/\mathrm{d}t^2$ is also determined for the warped time axis.

So far, only a single harmonic component is considered. There are, however, many harmonic components in the signal. Denoting the phase of other harmonic components of the signal by $\hat{\theta}$, the IF is proportional to that of the fundamental frequency as $\partial\hat{\theta}/\partial t = c\partial\theta/\partial t$, c being constant. Therefore, the phase of other harmonics satisfies the differential Equations (23) and (24), that means the one time warping function can convert the signal with many harmonic components to the signal with constant harmonic frequencies on the warped time axis.

In a practical analysis, a signal with rapidly changing IF is assumed to be a chirp signal. The frequency of the chirp signal is a first-order polynomial of time and the time warping function becomes a second-order polynomial of time. The warped time axis u is reset to the current time t_0, which is the center of the window function for the IF analysis. Denoting the u axis at t_0 by $u_{t_o} = p_{t_0}(t)$ and assuming $p_{t_0}(t_0) = 0$, the time warping function is written as

$$p_{t_0}(t) = \frac{1}{2}a_{t_0}(t - t_0)^2 + b_{t_0}(t - t_0), \tag{26}$$

around t_0 with constants a_{t_0} and b_{t_0}. Also, assuming that the time scale of the warped time axis at the origin $u_{t_o} = 0$ is equal to that of the original time axis t, namely $\mathrm{d}u_{t_0}/\mathrm{d}t = 1$, the constants of the time warping function are given as

$$b_{t_0} = \left.\frac{\mathrm{d}u_{t_0}}{\mathrm{d}t}\right|_{u_{t_0}=0} = 1 \quad \text{and} \quad a_{t_0} = \left.\frac{\mathrm{d}^2 u_{t_0}}{\mathrm{d}t^2}\right|_{u_{t_0}=0} = q(t_0). \tag{27}$$

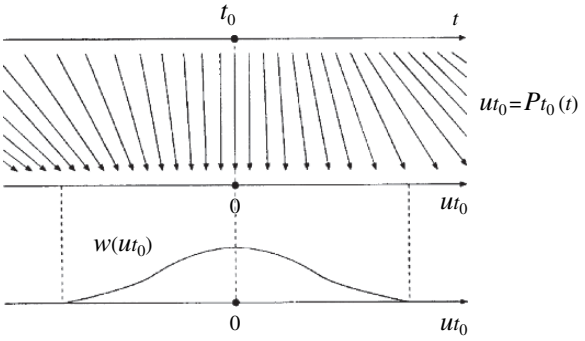


FIGURE 7 Practical implementation of time warping applied frame by frame. (From [8].)

Figures 7 and 8 show the relationships between the original and the warped time axis and the time warping function.

To achieve time warping, $q(t)$ in Equation (25) or (27), which reflects the change in the harmonic frequencies, is estimated from the IF attractors. Denoting one of the attractors as a function of time by $\lambda_a(t)$, the time derivative is written from Equation (17) as

$$\frac{\mathrm{d}\lambda_a}{\mathrm{d}t}\frac{\partial\mu}{\partial\omega} + \frac{\partial\mu}{\partial t} = 0. \tag{28}$$

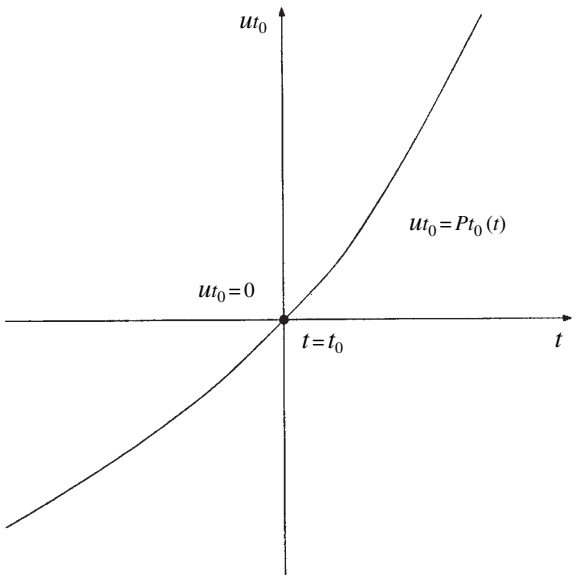


FIGURE 8 Relationship of the original and warped time axes. (From [8].)

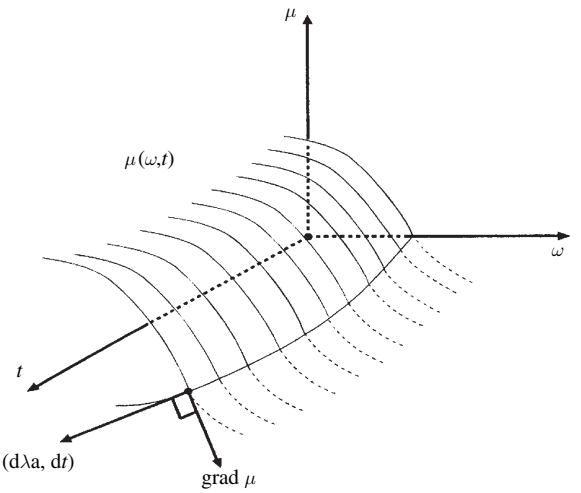


FIGURE 9 Property of tangent vector on trajectory of IF attractor. (From [9].)

That means that the tangent vector $[\mathrm{d}\lambda_a/\mathrm{d}t, 1]$ on the IF attractor's trajectory is orthogonal to the gradient of $\mu(\omega,t)$ as shown in Figure 9. Then, the estimate of $q(t)$ in Equation (25) for the IF attractor, which is the estimate of the harmonic frequency, is given by

$$q_F(\omega,t) = \lambda^{-1}(\omega,t)\frac{\partial\lambda_a(\omega,t)}{\partial t}, \tag{29}$$

where $\mathrm{d}\lambda_a(\omega,t)/\mathrm{d}t = -(\partial\mu/\partial t)/(\partial\mu/\partial\omega)$. For multiple harmonic frequencies, the estimate of $q(t)$ is given by a weighted average of the IF attractors as

$$q(t) = \frac{\int_{\omega_0}^{\omega_1} \rho(\omega,t) q_F(\omega,t)\mathrm{d}\omega}{\int_{\omega_0}^{\omega_1} \rho(\omega,t)\mathrm{d}\omega}. \tag{30}$$

Here, $\rho(\omega,t)$ is the weighting factor given by $\rho(\omega,t) = \gamma(\omega,t)\alpha(\omega,t)\,|Y(\omega,t)|$, where $\gamma(\omega,t) = 1$ at the attractors and $\gamma(\omega,t) = 0$ otherwise, $\alpha(\omega,t)$ is an enhancing factor for the attractors corresponding to harmonic components [10] and $|Y(\omega,t)|$ is the magnitude of the filter-bank output.

When the time warping with the second-order polynomial is applied to the signal, the signal reconstruction from the filter-bank outputs in Equation (3) is rewritten as

$$x(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Y(\omega, \tau) \xi_{t_0}(\omega, \tau - t) \, \mathrm{d}\omega \mathrm{d}\tau, \tag{31}$$

where the basis function is given by

$$\xi_{t_0}(\omega, t) = (1 + a_{t_0} t) \xi(\omega, \frac{a_{t_0}}{2} t^2 + t). \tag{32}$$

This means that the kernel is made from a chirp signal $e^{j\omega(\frac{a_{t_0}}{2}t^2+t)}$.

In practical implementation, the whole procedure has two steps. First, the IF analysis without time warping is performed and the time warping $p(t)$ in Equation (26) is determined by calculating $q(t)$ in Equation (30). Then, the IF attractors are obtained again for the time-warped signal [8].

The signal synthesis from IF attractors is the same as conventional sinusoidal synthesis in Equation (13). The phase of the sinusoid is obtained by integrating the IF attractor $\lambda_i(t)$ and the amplitude is given by the magnitude of the filter-bank output $|Y(\lambda_i(t), t)|$. The block diagram of IF attractor-based sinusoidal

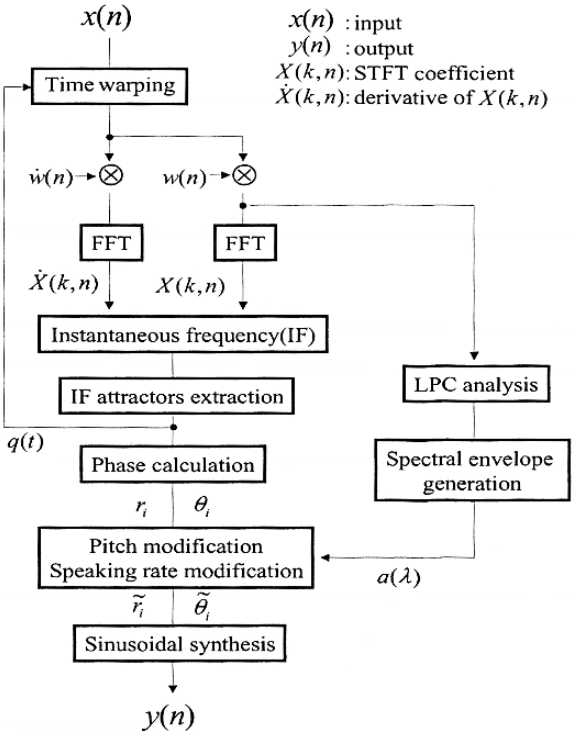


FIGURE 10 Instantaneous frequency (IF) attractors with time warping for a speech signal. (From [8].)

analysis/synthesis system is shown in Figure 10. Figure 11 shows the IF attractors obtained by the two-step time warping analysis for the speech sample in Figure 2 and Figure 5. These IF attractors are tracking the harmonic frequencies well, even if they are rapidly changing. The analysis is performed every 2 ms and the

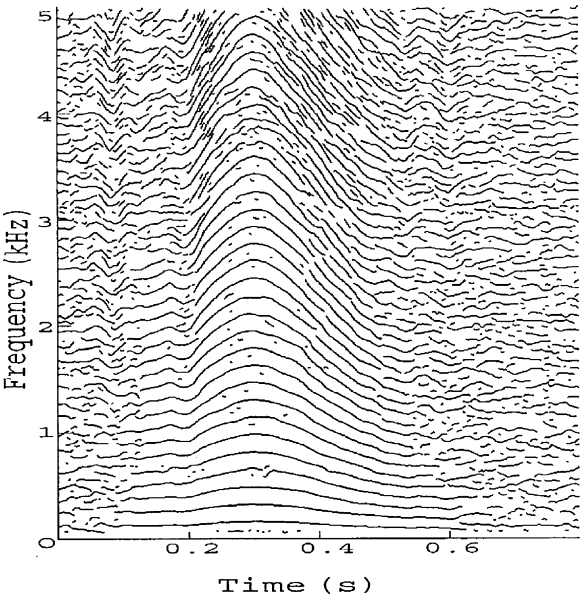


FIGURE 11 Block diagram of IF attractor-based sinusoidal analysis-synthesis. (From [8].)

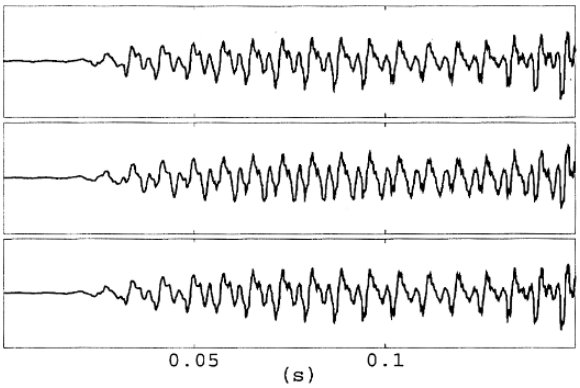


FIGURE 12 Waveform examples of sinusoidal model. Original signal (*top*), synthesized signal without phase adjustment (*middle*) and with phase adjustment (*bottom*) for a speech signal. (From [8].)

continuity of IF attractors is detected by simple nearest search. In the synthesis, the IF attractors are linearly interpolated so that the phase, computed by accumulating the estimated IF, coincides with the observed phase at the beginning of the analysis frame [8]. This phase adjustment with linear interpolation instead of cubic interpolation of phase described in Section 3 eliminates the accumulated phase error in the phase computation from the estimated IF. Figure 12 shows an example of speech waveform synthesized by the sinusoidal analysis-synthesis system.

5 • Time-Scale and Pitch Modification

The sinusoidal model is an important tool for time-scale and pitch modification of signals, as well as analysis and synthesis [11–12]. In sinusoidal synthesis, the time scale can be altered by modifying the IF and the magnitude as $\tilde{\lambda}_i(n) = \lambda_i(c_t n)$ and $\tilde{r}_i(n) = r_i(c_t n)$, where c_t is a scaling factor. Then, the phase is also computed by accumulating the modified IF. In the pitch modification, the IF is modified as $\tilde{\lambda}_i(n) = c_p \lambda_i(n)$, being a scaling factor c_p. To maintain the spectral envelope after pitch modification, the magnitude of the sinusoid is adjusted according to $\tilde{\lambda}_i(n)$ using the LPC spectral envelope $a(\lambda)$ by $\tilde{r}_i(n) = r_i(n) \cdot a(\tilde{\lambda}_i(n))/a(\lambda_i(n))$ as shown in Figure 11 [8].

Time-scaled or pitch-scaled signal $y(n)$ is synthesized by using the modified phases and amplitudes as

$$y(n) = \sum_{i=1}^{N} \tilde{r}_i(n) \cos(\tilde{\theta}_i(n)). \tag{33}$$

References

1. J. L. Flanagan and R. M. Goden, "Phase-vocoder," *Bell Syst. Tech.*, vol. **45**, pp. 1492–1509, 1966.
2. L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1978.
3. P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993.
4. H. Nawab and T. Quatieri, "Short-time fourier transform," In J. S. Lim and A. V. Oppenheim, editors, *Advanced Topics in Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1988.
5. R. McAulay and T. Quatieri, "Speech analysis-synthesis based on sinusoidal representation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-**34,** no. 4, pp. 744–754, 1986.
6. J. O. Smith and X. serra, "PARSHL: An analysis/synthesis program for nonharmonic sounds based on a sinusoidal representation," In *Proc. Int. Computer Music Conf.*, pp. 290–297, 1987.

7. E. B. George and M. J. smith, "Speech analysis/syntesis and modification using an analysis-by-synthesis/overlap-add sinusoidal model," *IEEE Trans. Speech and Audio Process.*, vol. **5**, no. 5, pp. 389–406, 1997.
8. T. Abe and M. Honda, "Sinusoidal model based on instantaneous frequency attractors," *IEEE Trans. on Audio, Speech and Language Processing,* July 2006; **14(4)**: 1292–300.
9. R. Kumaresan and C. S. Ramalingam, "On separating voiced-speech into its components," In *Proc. 27th Asilomar Conf. Signals, Systems, Computers,* vol. **1–2**, Pacific Grove, CA, 1993, pp. 1041–1046.
10. T. Abe, T. Kobayashi and S. Imai, "Robust pitch estimation with harmonic enhancement in noisy environments based on instantaneous frequency," In *Proc. ICSLP*, Philadelphia, PA, pp. 1277–1280, 1996.
11. T. Quatieri and R. McAulay, "Speech transformation basen on a sinusoidal representation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-**34**, no. 6, pp. 1449–1464, 1986.
12. T. Quatieri and R. McAulay, " Shape-invariant time-scale and pitch modification of speech," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-**40**, no. 3, pp. 497–510, 1992.

75

Room Transfer Function

Mikio Tohyama

Waseda University, Tokyo, Japan

1 • Introduction

Sound, which is an important carrier of information for human communication, propagates in elastic media as waves. A sound wave arriving from a sound source, however, deforms as it travels in a space due to the elastic and geometrical conditions of the medium. A sound wave propagating in a space can be represented by a wave equation and is characterized by the transfer function (TF) between the sound source and receiving positions. This section describes the magnitude and phase characteristics, which represent the sound wave propagating from the sound source, in terms of the poles and zeros of the TF. Although the frequency characteristics in a reverberant space appear too irregular, spatial sound, which is a general term for sound propagating in a space, is necessary for daily communication. Room TF studies are important in processing spatial sound signals to obtain informative, high-quality sound in communication spaces.

2 • Modal Expansion of the TF

The modal expansion form of a TF with a source at $\mathbf{x}_s$ and an observer, $\mathbf{x}_o$, in a room can be represented by [1]

$$H(\omega) = \sum_m \frac{\psi_m(\mathbf{x}_s)\psi_m(\mathbf{x}_o)}{\omega^2 - \omega_m^2 - j\omega\omega_m\eta_m}, \tag{1}$$

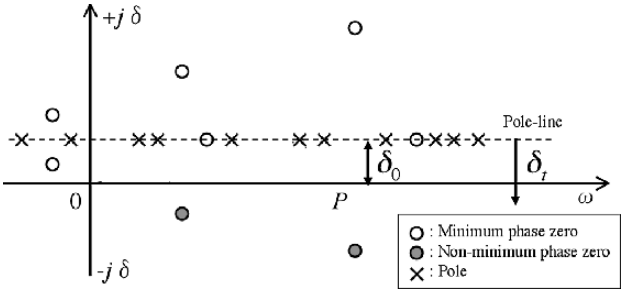


FIGURE 1 Schematic illustration for poles and zeros in the complex frequency plane for the transfer function.

where ψ_{m} is the mode shape, ω_{m} is the undamped resonance (angular) frequency, and η_{m} is the loss factor. The poles of this function are at

$$\omega_{\mathrm{p}} = \pm\omega_{\mathrm{m}} + j\frac{1}{2\omega_{\mathrm{m}}\eta_{\mathrm{m}}}. \tag{2}$$

If we assume that the reverberation time, T_{R}, which can be defined as

$$T_{\mathrm{R}} = \frac{3\log 10}{\omega_{\mathrm{m}}\eta_{\mathrm{m}}/2}, \tag{3}$$

is the same for all modes, then the poles lie along a straight line at the same distance above the real frequency axis in the complex frequency plane, as we can see from Figure 1. If the reverberation time depends on the frequency, then the pole line will have a different locus.

3 • Distribution of Eigenfrequencies

The poles are determined by eigenfrequencies. The eigenfrequencies of a rectangular room of dimensions (L_x, L_y, L_z) with rigid boundaries are given by

$$\omega_{lmn} = \frac{c\pi}{L}\sqrt{(al)^2 + (bm)^2 + (cn)^2}, \qquad \text{(rad/s)} \tag{4}$$

where $l, m,$ and n are integer numbers; the lengths of the sides (in meters) are $L = aL_x = bL_y = cL_z$, and $a, b,$and c are the ratios of the lengths.

The distribution of the eigenfrequencies appears complicated even for rectangular rooms. However, the process by which the eigenfrequencies occur can be characterized by a Poisson process that results in an exponential distribution of eigenfrequency spacing [1,2]. Figure 2 shows a numerical example of the (modal) spacing distribution for oblique modes in a rectangular room. The oblique modes denote eigenmodes whose eigenfrequencies have no zero entries for $l, m,$ or n. We can see that pairs of eigenfrequecies with very narrow spacing are quite likely to occur.

The Poisson model is not adequate for practical cases, however. Groups of eigenmodes that have degenerate eigenfrequencies will suffer from the eigenfrequencies splitting due to perturbations in practical systems. The effects of these perturbations on spacing statistics can be illustrated by coupled mechanical oscillations [3]. Referring to Figure 3, the masses of the two oscillators are M_1 and M_2, and their spring constants are K_1 and K_2, and the

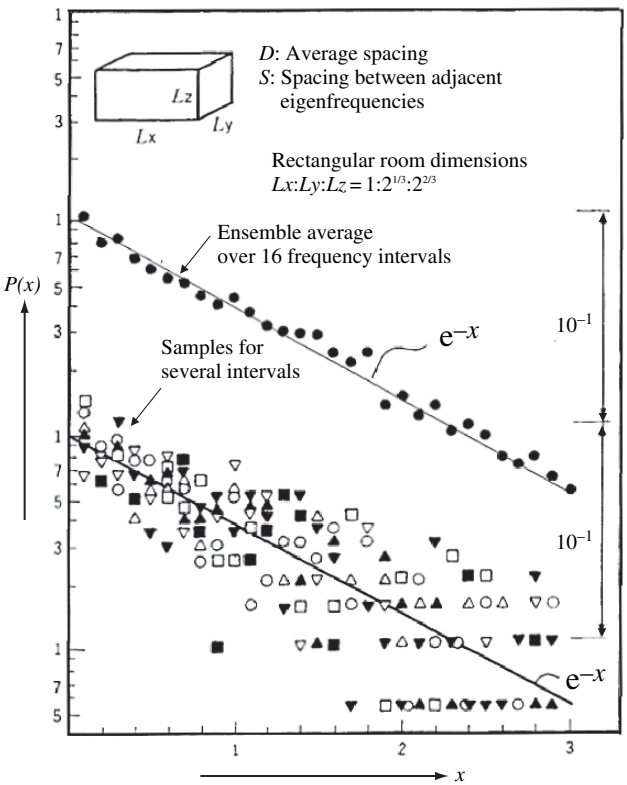


FIGURE 2 An example of numerical analysis for modal spacing statistics of a rectangular room with rigid walls. (From Ref. [3].)

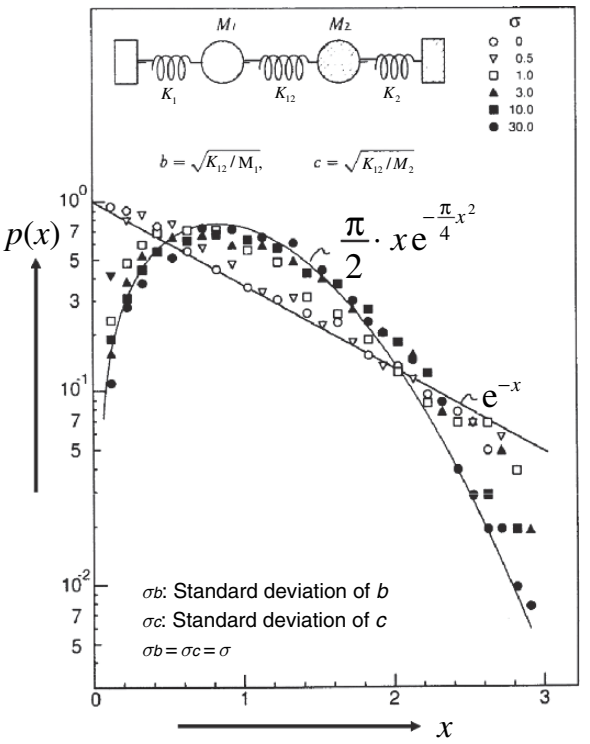


FIGURE 3 Samples of eigenfrequency spacing histograms for a coupled oscillator. (From Ref. [3].)

coupling spring constant is K_{12}. The eigenfrequencies are given by the eigenvalues for the matrix

$$E = E_{\mathrm{o}} + E_{\mathrm{p}} = \begin{pmatrix} \omega_1^2 & 0 \\ 0 & \omega_2^2 \end{pmatrix} + \begin{pmatrix} b^2 & -b^2 \\ -c^2 & c^2 \end{pmatrix}$$

where

$$\omega_1^2 = \frac{K_1}{M_1}, \quad \omega_2^2 = \frac{K_2}{M_2}, \qquad b^2 = \frac{K_{12}}{M_1}, \quad c^2 = \frac{K_{12}}{M_2}. \tag{5}$$

Matrix E_0 in the equation above has the eigenvalues of two independent oscillators and matrix E_{p} represents the coupling effect on the eigenvalues.

The perturbation caused in an actual acoustic space like that in an irregularly shaped room results in random coupling between adjacent modes [4]. Assume that the coupling parameters b and c of matrix E_{p} are mutually independent Gaussian variables with a zero mean value and a standard deviation of σ. The spacing of the two eigenfrequencies $\Delta\omega$ can thus be expressed by

$$(\Delta\omega)^2 \cong \begin{cases} b^2 + c^2 & b^2, c^2 \rightarrow \text{large} \\ (\omega_2 - \omega_1)^2 & b^2, c^2 \rightarrow 0, \end{cases} \tag{6}$$

where

$$(\Delta\omega)^2 = \Omega^2 + P^2, \tag{7}$$

$$\Omega^2 = \omega_1^2 + \omega_2^2, \tag{8}$$

$$P^2 = b^2 + c^2 - 2\sqrt{D^2}, \tag{9}$$

and

$$D^2 = (\omega_1^2 + b^2)(\omega_2^2 + c^2) - b^2c^2. \tag{10}$$

The random variable $\Delta\omega$ is the positive square root of the squared sum of the two independent Gaussian variables. The spacing for the two eigenfrequencies therefore follows a Rayleigh distribution as standard deviation σ (perturbation) increases.

Figure 3 illustrates the transition in spacing histograms from a Poisson (an exponential) to a Rayleigh distribution. Five thousand pairs of eigenfrequencies with a spacing that followed a Poisson distribution with a zero coupling effect were used for this calculation. We can surmise that the Rayleigh distribution of the eigenfrequency spacing is produced by randomly coupled wave modes [1–2].

The results stated above can be generalized by studies about the eigenvalues for an ensemble of real symmetric random matrices such that the spacing of successive eigenvalues has a probability density, $p(x)$, well approximated by the Rayleigh distribution,

$$p(x) \cong \frac{\pi x}{2} \exp(-\frac{\pi x^2}{4}), \tag{11}$$

when spacing x is normalized by the mean spacing and the first moment of density is unity [5].

However, as Lyon [1] suggested, some experimental data indicate that the spacing distribution could be well fitted by [6]

$$p(x) = 4x \exp(-2x). \tag{12}$$

The equation above is easier to handle analytically than that in Eq. 11. If we approximate the eigenfrequency distributions by using a family of Γ functions including Eq. 12, we can find the transition from the exponential distribution using the degree of freedom of the Γ distribution [7] as we will see in the next section.

4 • Chaotic Properties of Ray Tracing and Eigenfrequency Distribution

The transition from an exponential distribution could be confirmed in two-dimensional vibration systems surrounded by semi-stadium boundaries, which change from regular (such as rectangular) to irregular (represented by a stadium) shapes as defined in Figure 4. The shape of a semi-stadium is changed from a stadium to a rectangle as i is increased.

The reason we call boundaries regular or irregular is due to the chaotic property of ray tracing in the surrounding region. Suppose that we have a closely located pair of sound rays at the initial condition shown in Figure 5. Figure 6 shows how much a slight difference (error) in the initial condition of a sound ray causes changes in the ray trajectories [8]. The distance between the pair of rays at reflection points on the boundary, however, exponentially increases in the stadium system, as shown in Figure 6(a), while it linearly increases under the rectangular boundary (Figure 6(b)), as the rays continue to propagate.

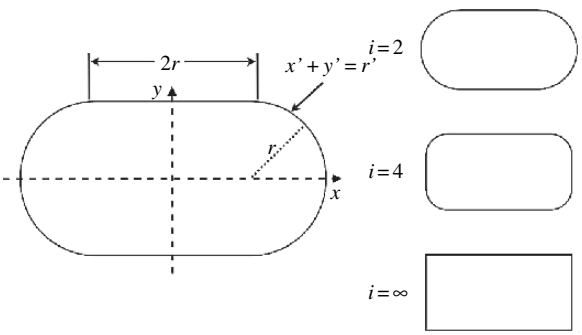


FIGURE 4
Semi-stadium boundaries.
(From Ref. [7].)

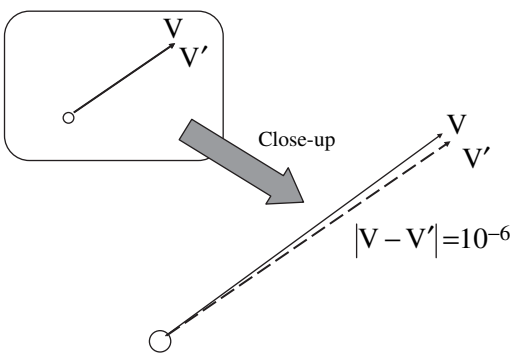


FIGURE 5 A display of the initial condition for a closely located pair of sound rays.

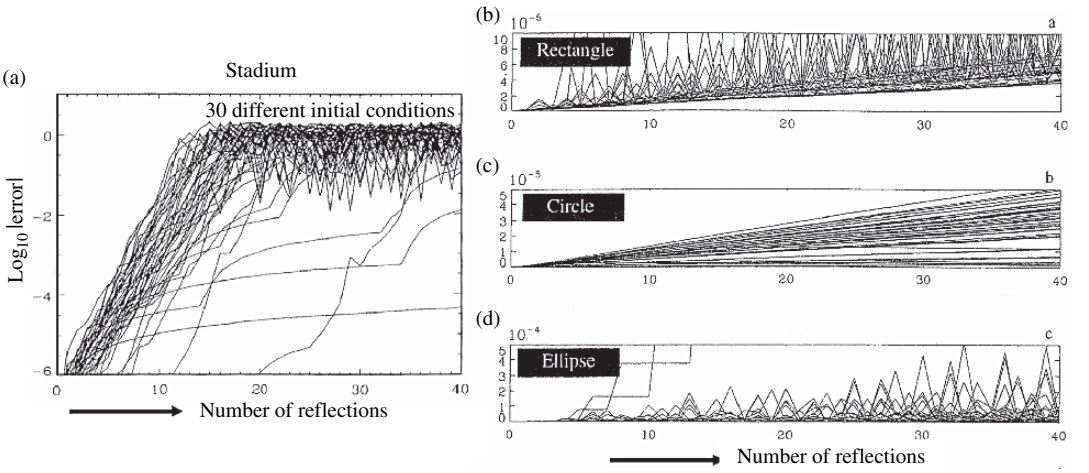


FIGURE 6 Error propagation analysis for sound-ray tracing simulation. (From Ref. [8].)

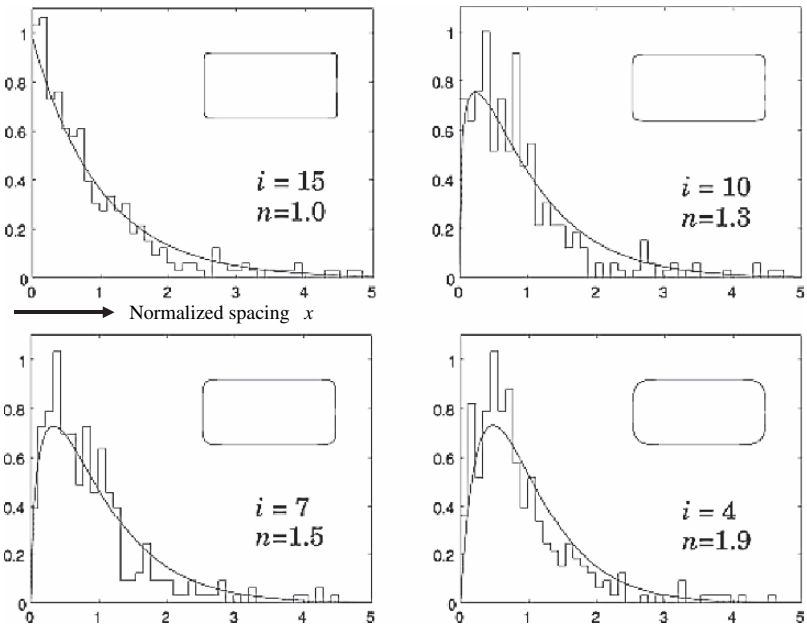


FIGURE 7 Modal spacing statistics for semi-stadium fields. (From Ref. [7].)

A system such as that in the rectangular (circle or ellipse) case is regular but that in the stadium case is irregular or chaotic.

The eigenvalues of the semi-stadium fields can be analyzed with the finite element method. Figure 7 shows the spacing distributions for eigenfrequencies when $i = 4$, 7, 10, and 15. We see that a family of Γ distributions such as

$$p(x, n) = \frac{n^n}{\Gamma(n)} x^{n-1} \exp(-nx), \tag{13}$$

where spacing x is normalized by the mean spacing, could be fitted to the distribution with a non-integer degree of freedom n, 1.9, 1.5, 1.3, and 1.0 for boundaries [7]. Here, the degree of freedom is unity for the exponential distribution, while the distribution expressed by Eq. 12 has a degree of two. This transition in the degree of freedom might be interpreted as indicating the breakdown of regularity.

5 • Distribution of Zeros

5.1 Zeros Under Low Modal Overlap Conditions

The poles are located on the pole line, but the zeros can be distributed on a complex-frequency plane. If the mode-shaped functions are real and the pole is as it appears in Figure 1, then the zeros are evenly distributed above and below the pole line, because the TF is symmetrical with respect to the pole line [9].

The TF is approximately written as

$$H(\omega) = \frac{A}{\omega - \omega_1} + \frac{B}{\omega - \omega_2} + R \qquad (\omega_1 < \omega < \omega_2) \tag{14}$$

on the pole line [10]. The primary part will dominate TF in modal expansion, with rest of the terms producing a small remainder function, R. A minimum-phase zero is produced between the two adjacent poles when they have residues of the same sign. Such a zero is located on the line connecting the poles (the pole line), as shown in Figure 8(a) and (b). However, if the poles have residues of opposite signs, the occurrence of zeros will depend strongly on the remainder function [11].

Two cases of zeros can be produced from the residues of opposite signs and the remainder function, as shown in Figure 8(c) and (d). The first is a double zero between two adjacent poles [11]. Both of these double zeros are minimum phase because they are located on the pole line. The second case produces a conjugate pair of zeros in the complex frequency plane [9]. These zeros are located symmetrically with respect to each other at equal distances from the pole line. One of these zeros can be non-minimum phase.

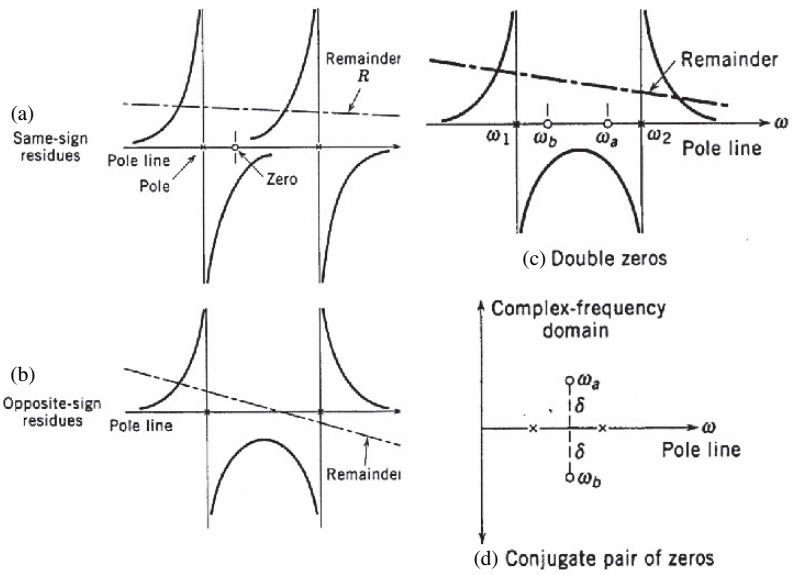


FIGURE 8 Zero locations between a pair of adjacent poles. (From Ref. [9–11].)

5.2 Zeros Under High Modal Overlap Conditions

We again take TF written in modal expansion form using Eq. 1. When the source and the observer are in the same location, all the residues are positive. The poles thus interlace with zeros and the number of poles and zeros are equal. As the source and receiver move apart, the zeros migrate (poles do not move). Some move above the line, an equal number moves symmetrically below the line, and the remainder stays on the line. This migration of zeros produces the propagation phase, which will be explained later.

If the observer is located in the incoherent region away from the source, we can then assume the probability that the residue sign will change is 1/2. The possible number of zeros below the real frequency axis (non-minimum-phase zeros) could thus be estimated using $1/4\, N_{\mathrm{p}}$, where N_{p} denotes the number of poles and the case for double zeros has been neglected [10–11]. However, this possible number of zeros is reduced as $\delta_0 \cong 6.9/T_R$ increases [12].

Statistical approaches to room acoustic TF analysis under high modal overlap conditions have assumed the TF for a reverberant field (incoherent field) to be a random process in the complex frequency domain [13]. Following random process theory [14], Tohyama et al. [12] postulated that the zeros are distributed around the pole line following a Cauchy distribution

$$p(\beta) = \frac{4/\pi}{1+4\beta^2}, \tag{15}$$

where $\beta = \delta_t/\overline{\Delta\omega_p}$, δ_t denotes the distance from the pole line as shown in Figure 1, with $\overline{\delta\omega_p} = 1/n_p(\omega)$ being the average pole spacing. For a high modal overlap, the Cauchy distribution reduces to

$$p(\beta) = \frac{1}{\pi\beta^2}. \tag{16}$$

Thus, the number of non-minimum phase zeros in frequency band $\Delta\omega$ is

$$N_z^- = \frac{N_p(\Delta\omega)}{4}\int_{\beta_0}^{\infty}\frac{1}{\pi\beta^2}d\beta = \frac{\Delta\omega/4}{\pi\delta_0}, \tag{17}$$

where $N_p(\Delta\omega)$ denotes the number of poles in frequency band $\Delta\omega$, i.e., $N_p(\omega) = n_p(\omega)\Delta\omega$, $n_p(\omega)$ is the density of the poles, and $\beta_0 = \pi\delta_0 \cong \pi 6.9/T_R$. The density of the non-minimum-phase zeros is independent of the frequency, while the number of poles increases as the frequency increases.

6 • Phase Response of Room TF

Figure 9 has samples of the magnitude and phase response for a room TF [15]. We can see dips and many peaks in the magnitude response, in addition to abrupt phase changes. We can normally

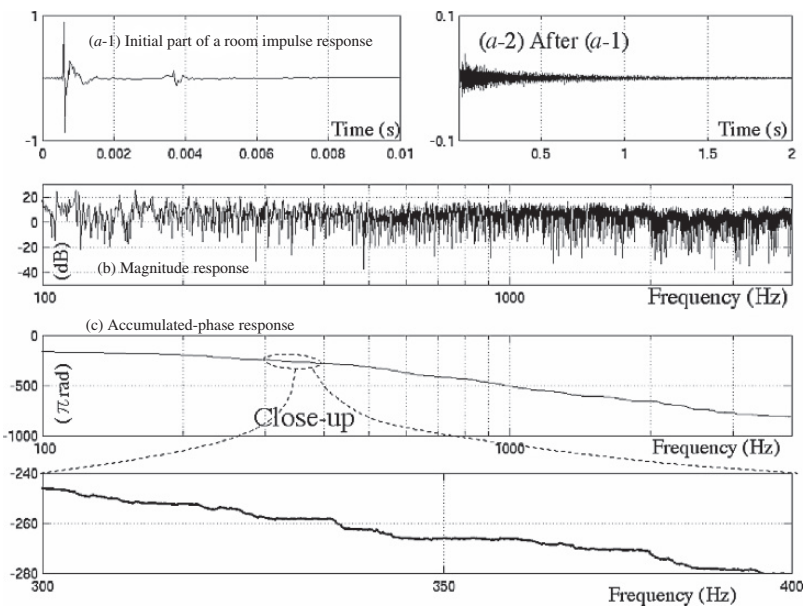


FIGURE 9 A sample of magnitude and phase response for the room transfer function from an impulse response. (From Ref. [15].)

expect that steep magnitude peaks will become less as damping increases. The statistics on magnitude response have been extensively described in terms of the eigenfrequency distribution in the literature [1,2,16].

However, the dips depend on the location of the zeros. If we assume that the density of zeros has decreased following a Cauchy distribution in the complex frequency plane, then isolated dips should be seen in the magnitude response under short reverberation conditions, because the zero density becomes sparse as the reverberation time decreases. The phase response is estimated using the distributions for both the eigenfrequencies (poles) and the zeros.

6.1 Propagation Phase and Coherent Region

The increase in phase between the source and the observation locations in a one-dimensional system can be simply understood [10]. Assume that at the nth eigenfrequency, the x_0 and x_s are separated by p nodes so that $\mathrm{Int}[(k_n/\pi)(x_0 - x_s)] = p$, where k_n is the wavenumber (in reciprocal meters) corresponding to the nth eigenfrequency. For a higher eigenfrequency whose wavenumber is k_m, these source and receiver locations stay in phase until p increases by unity. Then, $\mathrm{Int}[(k_m/\pi)(x_0 - x_s)] = p + 1$ and the phase between observation and source locations has increased by π. The phase trend in the one-dimensional case where the number of nodes increases in an orderly fashion between the source and the receiver is

$$\Phi = k(x_0 - x_s) \qquad \text{(rad)}. \tag{18}$$

The accumulated phase of a one-dimensional system is just the same as the phase delay due to direct wave propagation from the source to the observer. We call this the propagation phase [10].

We can find the trend for the propagation phase even in a three-dimensional system like reverberant space [17]. Figure 10 shows the measured impulse responses. We can confirm that the arrival time is delayed and the energy of direct sound is decreased when the sound source distance (SSD) increases. Figure 11 shows examples of minimum-phase characteristics of the impulse responses. Because the minimum-phase characteristics are such that the accumulation phase becomes 0 as shown in the figure, we were unable to find the phase propagation according to SSD. However, we could extract the propagation phase by using the linear regression analysis of narrow-band phase frequency characteristics [17].

Figure 12 shows the linear regression analysis of (normalized) minimum-phase characteristics against wavenumber k, where $k_{200} = 2\pi \cdot 200/c = 3.70$ (c is the speed of sound). We can see the increasing gradient of the regression line (phase trend) with

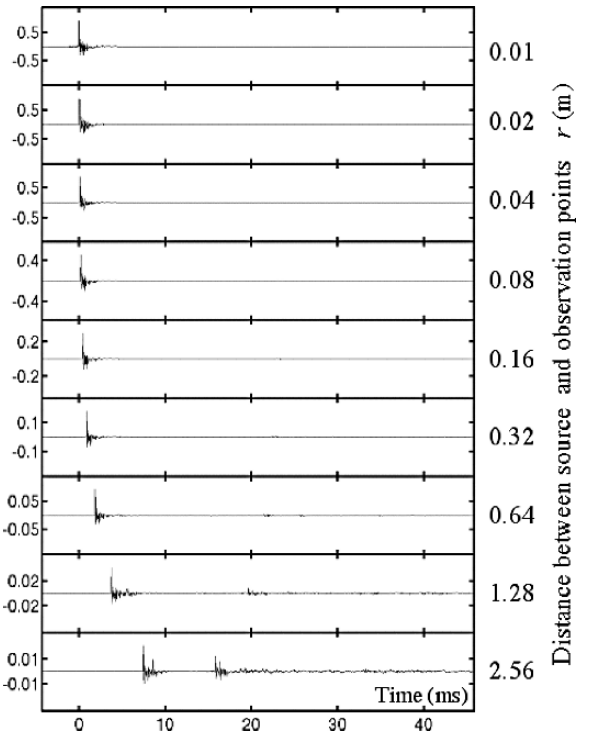


FIGURE 10 A display of impulse responses measured in a reverberant space (From Ref. [17].)

increasing r. Figure 13 plots the results from evaluating the gradients of the regression lines with r. The dotted line in this figure represents the propagation phase, i.e., r. The limit that can be considered to be SSD in the gradient of the phase regression line in the minimum-phase component is around $r - r_0 \cong 0.7$ (m).

The coherent field described by Morse and Bolt [4] gives

$$r_{\text{ph}} = \sqrt{\frac{A}{64}} \quad \text{(m)}, \tag{19}$$

where $A \cong 0.161V/T_R$ (m^2), V is room volume (m^3), and T_R denotes the reverberation time(s). Here, we use a reverberation time of 1 s and a room volume, V, of 183 m^3; thus,

$$r_{\text{ph}} \cong 0.68 \quad \text{(m)}, \tag{20}$$

indicated by the dash-dotted line in Figure 13.

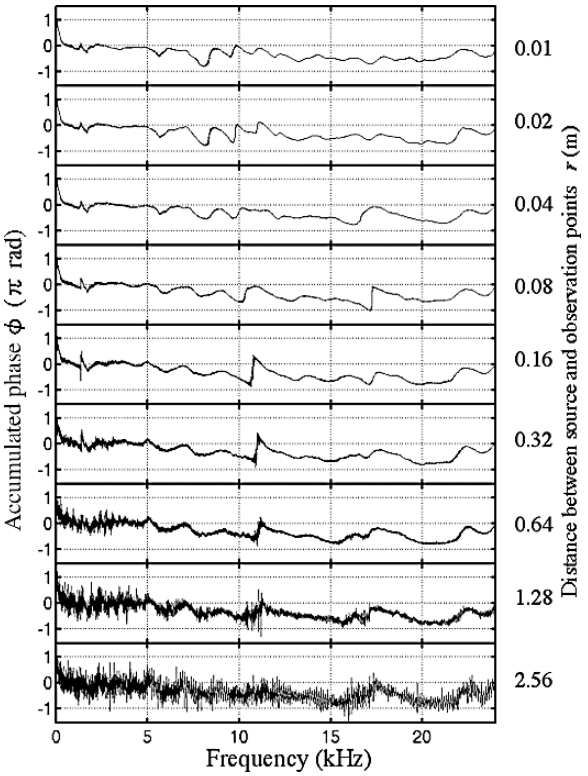


FIGURE 11
Minimum-phase components for the impulse responses shown in Figure.10. (From Ref. [17].)

The critical distance, r_c [17], based on the energy rate between direct and reverberant sound is defined by

$$\frac{I_0}{I_{rev}} = \frac{\bar{\alpha}S}{16\pi {r_c}^2} = 1, \tag{21}$$

where I_0 means the intensity (W/m^2) of direct sound and I_{rev} means that of reverberant sound. Here, $\bar{\alpha}$ means the averaged absorption coefficient and S means the surface area of the room (m^2). Consequently, the critical distance yields

$$r_c = \sqrt{\frac{\bar{\alpha}S}{16\pi}} \cong 0.77 \ \text{(m)}. \tag{22}$$

Figure 13 shows r_c as a broken line. The direct sound field (coherent field) presumed from the propagation phase of minimum-phase characteristics corresponds to the coherent field within the critical distance.

In Figure 12, we can not only see the increasing gradient for the regression line but also the variance in deviation from the

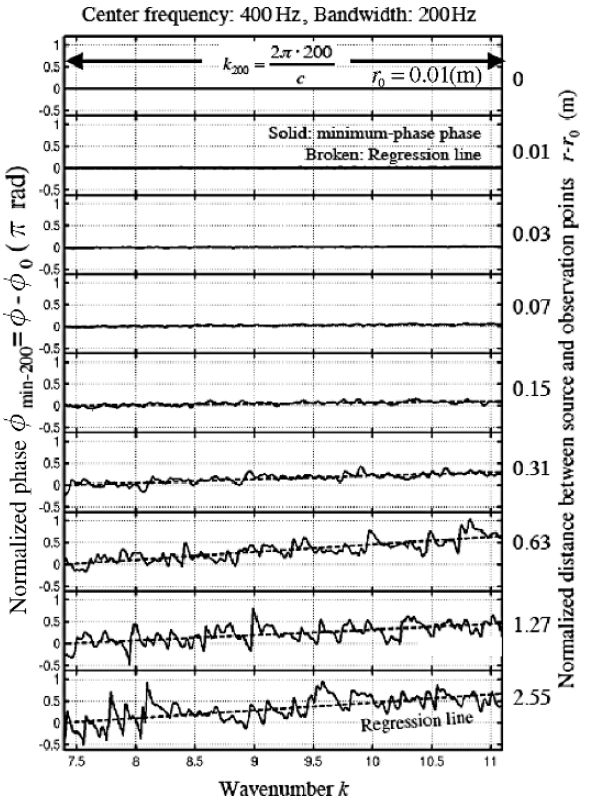


FIGURE 12 Linear regression analysis for the narrow-band minimum-phase characteristics normalized by the accumulated phase at $r_0 = 0.01$(m) in Figure.11. (From Ref. [17].)

propagation phase with increasing r. Figure 14 plots our evaluation of the variance in deviation from the propagation phase of minimum-phase characteristics by using group delay. The variance in group delay does not depend on the frequency band and is almost constant [19].

6.2 Reverberation Phase and Non-Minimum-Phase Zeros

The phase characteristics in the two- and three-dimensional cases are not as orderly as those in the one-dimensional case, however. When the observer is in the reverberant (incoherent) field, the number of zeros remaining on the pole line is about half the number of poles. The remaining zeros are evenly distributed above and below the pole line due to the symmetry of TF, which means a quarter of them are below the pole line. The number of these zeros below the pole line that are also below the real frequency axis are non-minimum-phase zeros, and these zeros generate excessive phase that is independent of the

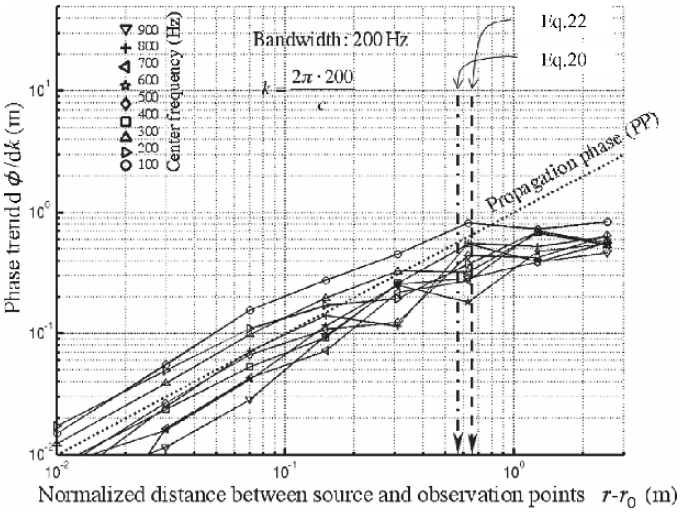


FIGURE 13 Slopes of the regression lines (phase trends) for the narrow-band phase as shown in Figure.12. (From Ref. [17].)

source-receiver response between the source and the observer. Thus, in the incoherent region outside the coherent field, the accumulated phase increases to become almost stationary. We call this stationary phase the reverberation phase [10].

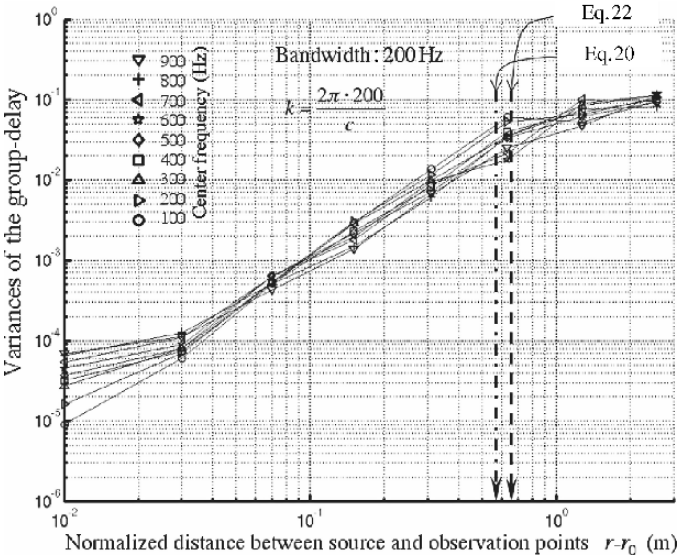


FIGURE 14 Group- delay variances from the propagation phase.

The accumulated phase can be determined by both the poles and the zeros. If there are N_p poles and if there are N_z^+ and N_z^- zeros in the lower and upper halves of the planes, then the accumulated phase is

$$\Phi = -\pi(N_p + N_z^+ - N_z^-) \qquad \text{(rad)}. \tag{23}$$

Impulse response records, which were measured in a reverberant space and were used in our method of processing digital signals, are of finite length. Therefore, the TF is characterized by the finite number of zeros in terms of mathematics. However, the magnitude and phase characteristics of the TF can be virtually represented (or interpreted) in terms of modal acoustics following Eq. 1.

Tohyama et al. [12] confirmed the density of non-minimum-phase zeros on the complex frequency plane decreased and was inversely proportional to the square of the distance from the pole line, using the measured impulse response records with adequate record lengths [20]. Thus, if we set $N_p = N_z^+ + N_z^-$, then the trend

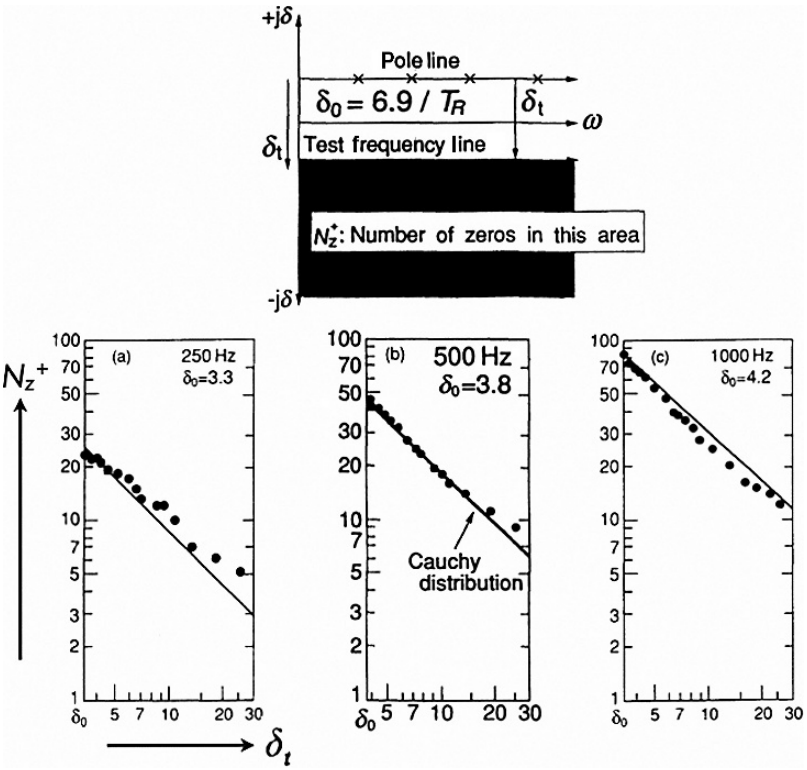


FIGURE 15 Distribution samples of the non-minimum phase zeros. (From Ref. [3,12].)

in accumulated phase could be estimated from the number of non-minimum-phase zeros, N_z^+, following

$$\Phi = -2\pi N_z^+. \tag{24}$$

Figure 15 is an example distribution for the non-minimum-phase zeros, where the average number of zeros below the test frequency line in the lower half of the plane is illustrated. Four measured TFs have been analyzed in a reverberation space. This demonstrates that the distribution of zeros followed the Cauchy distribution using Eq. 15. Figure 16 plots samples of phase responses as SSD increases using narrow-band analysis. We can see that the phase reaches the reverberation phase as SSD increases, thus confirming that the phase can be estimated from

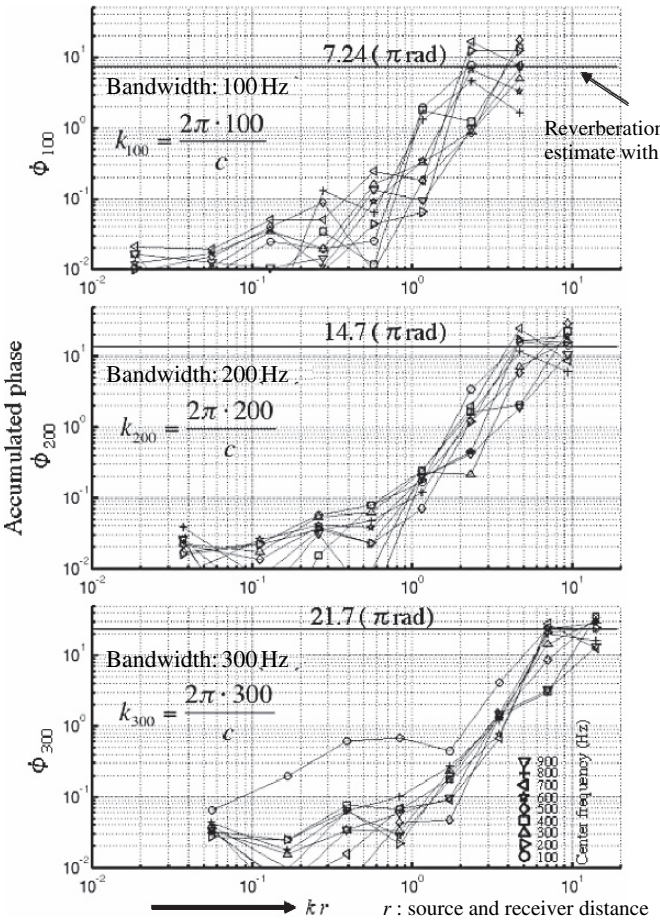


FIGURE 16 Samples of accumulated phase responses in a reverberant space. (From Ref. [17].)

the number of non-minimum-phase zeros. Estimates of reverberation phases are indicated by the solid lines in the panels and these are proportional to the bandwidth.

7 • Control of Magnitude and Phase for Artificial Reverberation Rendering

Reverberation is important for creating three-dimensional effects as well as for localizing sound images. Thus, rendering three-dimensional reverberation effects is a key component in modeling acoustic events. The modeling of room acoustics, including reverberation, has been investigated through various means, including scale models and wave- and ray-based methods [21]. These have been used to analyze sound fields and require detailed room-boundary conditions. Therefore, they have not always been an effective approach to modeling a sound field to render virtual-room reverberations.

Convolution-based reverberators that use impulse-response records measured in sound fields, such as those in concert halls, are effective tools for producing three-dimensional reverberation sounds. However, their use is restricted to impulse-response record databases. In other words, they cannot be adjusted to match room-reverberation conditions such as the SSD and frequency characteristics of reverberation.

7.1 Reverberation Rendering and SSD

The rendering perception of SSD is an important issue in virtual acoustics. The SSD-dependent frequency characteristics of TF could be one of these cues instead of energetic cues such as the direct to reverberant energy ratio (D/R) or others [22].

As previously stated, we confirmed that the coherent region could be estimated based on propagation-phase properties in the minimum-phase component of TF and deviations from the phase corresponding to SSD. The modeling of minimum-phase impulse responses based on the propagation phase with variances could render TF and the perception of SSD [23], if we changed the variance in the group delay.

7.2 Control of Frequency Characteristics for Artificial Reverberation

Frequency characteristics of reverberation are important for rendering reverberations that sound natural. In particular, it is important to maintain the tonal quality to control virtual reverberations. The reverberation-time characteristics can be artificially controlled while maintaining the locations of peaks and dips in the frequency response to the original impulse response [15].

Reverberation control, with respect to magnitude response, can be interpreted as the exponential time-windowing of the minimum-phase component [24]. Figure 17 plots (a) the magnitude and (b) phase response for the minimum-phase component, with and without windowing. Windowing with a negative exponent moves the zeros radially away from the unit circle in the z-plane, whereas with a positive exponent, it moves them closer to the unit circle. Plot A was modified with the negative exponent, while B was obtained with the positive one.

Reverberation control, i.e., moving the poles and zeros of the all-pass response to correspond to the exponential windowing for the minimum-phase part, can be done in the cepstral domain. Figure 18(a) plots the all-pass impulse response, and panel (b) illustrates its cepstrum. Figure 18(c) is a newly constructed response from the causal part of the cepstrum in Figure 18(b). This time sequence in Figure 18(c) could be interpreted as the response due to all-pass poles. Figure 18(d) shows the modified pole response obtained by exponentially windowing the sequence in Figure 18(c) and panel (e) shows its cepstrum. We can obtain

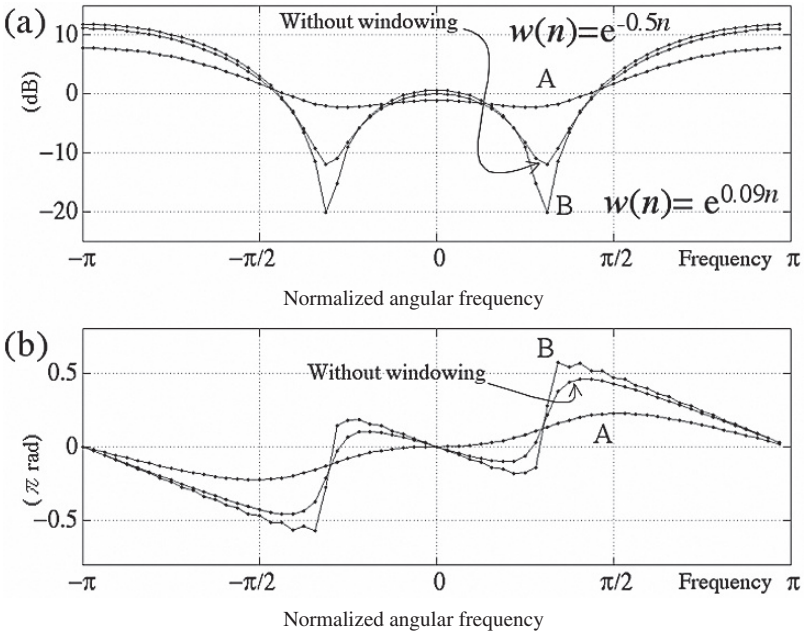


FIGURE 17 An example of magnitude and phase control for a minimum-phase impulse response by the exponential time windowing. (From Ref. [15].)

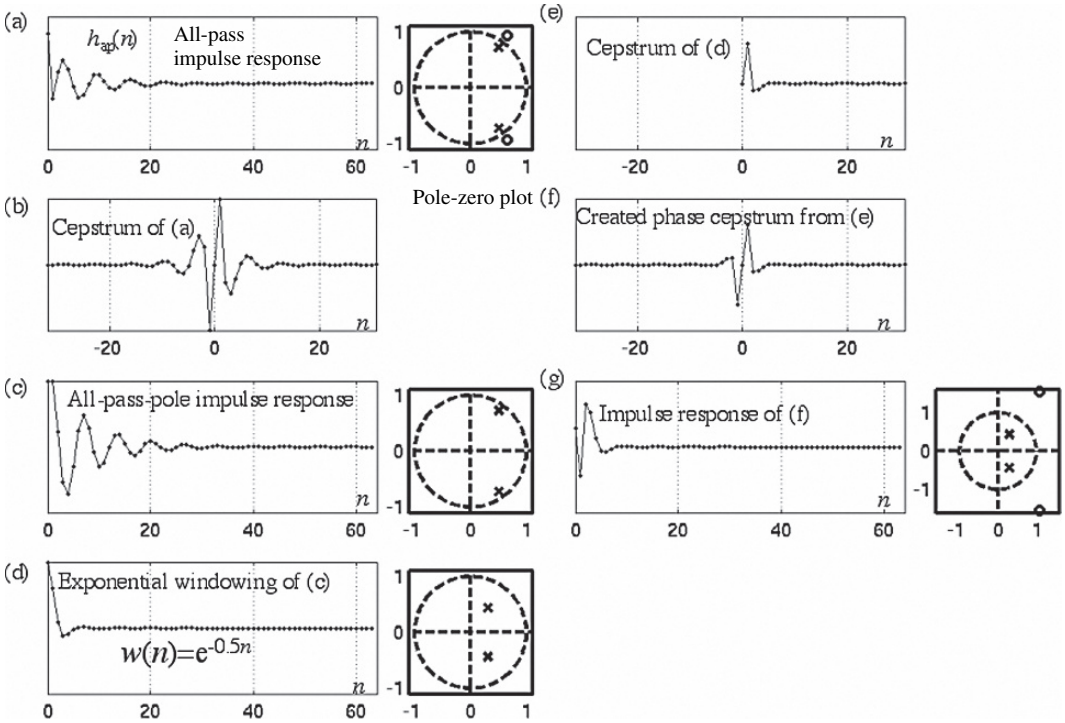


FIGURE 18 Reverberation control in the cepstral domain.

the all-pass phase cepstrum in Figure 18(f) by adding its non-causal part to the causal cepstrum shown in Figure 18(e). The all-pass impulse response can be constructed from the all-pass phase cepstrum in Figure 18(g).

Figure 19 shows examples of reverberation control. Plot O in Figure 19(a) denotes the original impulse response record and O in (b) plots the reverberation energy-decay curve [25]. Here, $h(n)$ denotes the impulse response record with a length of N. Plot A(B) is for a modified example using exponential windowing with a negative(positive) exponent. Figure 19(c) is a close-up of the magnitude and (d) is a close-up of all-pass phase responses. We can similarly confirm what effect reverberation control has shown in plots A and B.

We can see reverberation control is possible within the global range of reverberation time. Smoothing effects can be obtained for both magnitude and phase responses without pole/zero modeling while maintaining the global trend (frequency locations of peaks and dips) of the responses.

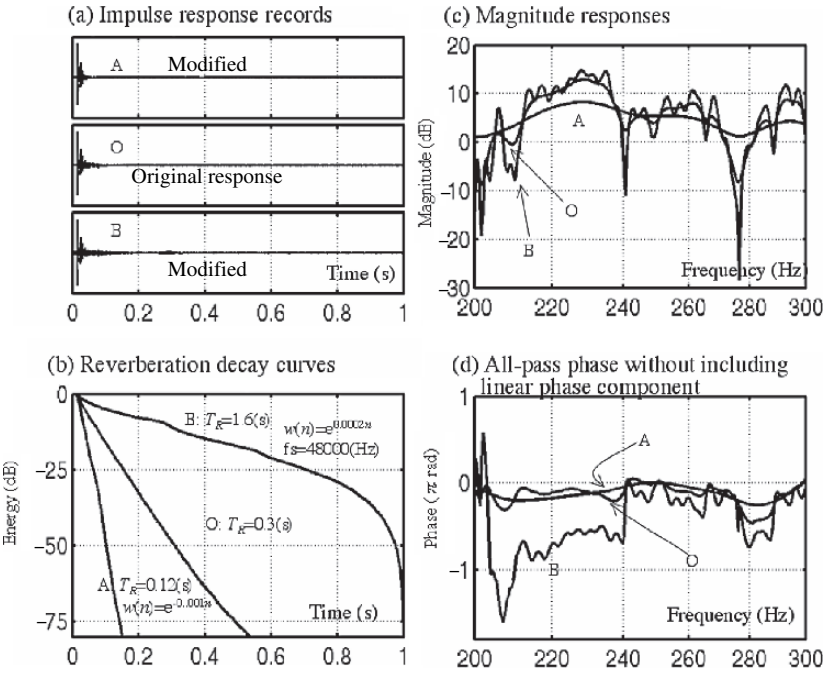


FIGURE 19 Examples of reverberation control.

References

1. R. H. Lyon, "Statistical analysis of power injection and response in structures and rooms," J. Acoust. Soc. Am. 45 545–565 (1969)
2. M. R. Schroeder, "Statistical parameters of the frequency response curves in large rooms," J. Audio Eng. Soc. 35(5) 299–306 (1987)
3. M. Tohyama, H. Suzuki and Y. Ando, Nature and Technology of Acoustic Space, Academic press, London UK (1995)
4. P. Morse and R. Bolt, "Sound waves in rooms," Rev. Mod. Phys. 16 69–150 (1944)
5. R. L. Weaver, "Spectral statistics in elastodynamics," J. Acoust. Soc. Am. 85(3) 1005–1013 (1989)
6. I. I. Gurevich and M. I. Pevsner, "Repulsion of nuclear levels," Nuc. Phys. 2 575 (1957)
7. Y. Fujisaka and M.Tohyama, "Eigenfrequency spacing analysis and eigenmode breakdown for semi-stadium type 2-D fields," J. Sound & Vib. 267 867–878 (2003)
8. M. Tohyama and Y. Fujisaka, "Room acoustics and chaos (in Japanese)," J. Acoust. Soc. Japan 53(2) 154–150 (1997)
9. M. Tohyama and R. Lyon, "Zeros of a transfer function in a multi-degree-of-freedom vibrating system," J. Acoust. Soc. Am. 86(5) 1854–1863 (1989)

10. R. Lyon, "Progressive phase trend in multi-degree-of-freedom systems," J. Acoust. Soc. Am. 73(4) 1223–1228 (1983)
11. R. Lyon, "Range and frequency dependence of transfer function phase," J. Acoust. Soc. Am. 76(5) 1433–1437 (1984)
12. M. Tohyama, R. Lyon and T. Koike, "Reverberant phase in a room and zeros in the complex frequency plane," J. Acoust. Soc. Am. 89(4) 1701–1707 (1991)
13. M. R. Schroeder, "Frequency correlation functions of frequency responses in rooms," J. Acoust. Soc. Am. 34 1819–1823 (1962)
14. S. O. Rice, "Mathematical analysis of random noise," Bell Syst. Tech. J. 24 46–108 (1945)
15. Y. Takahashi and M. Tohyama, "Artificial reverberation control using cepstrum decomposition while maintaining locations of peaks and dips on the frequency responses," J. Audio Eng. Soc. 53(12) 1142–1151
16. K. J. Ebeling, "Statistical properties of random wave fields," in W. P. Mason and R. V. Thurston (eds) Physical Acoustics, Academic Press, New York 223–310 (1984)
17. Y. Takahashi, M. Tohyama and Y. Yamasaki, "Phase responses of transfer functions and coherent field in a reverberation room (in Japanese with English abstract and figures), "Trans. on the Inst. of Elect. Inf. and Comm. Eng. J89-A(4) 291–297 2006
18. H. Kuttruff, Room Acoustics (3rd ed), Elserier Science, New York 114–120 (1991)
19. M. Tohyama and R. Lyon, "Phase variabilities and zeros in a reverberant transfer function," J. Acoust. Soc. Am. 95(1) 286–323 (1994)
20. M. Tohyama and R. Lyon, "gTransfer function phase and truncated impulse response," J. Acoust. Soc. Am. 86(5) 2025–2029 (1989)
21. L. Savioja, J. Huopaniemi, T. Lokki and R. V*ää*n*ä*nen, "Creating interactive virtual acoustics environments," J. Audio Eng. Soc. 47 675–705 (1999)
22. P. Zahorik, "Direct-to-reverberant energy ratio sensitivity," J. Aoucst. Soc. Am. 112(5) 2110–2117 (2002)
23. Y. Takahashi, M. Tohyama and Y. Yamasaki, "Rendering Spatial Reverberation and Perception of source Distance According to the Minimum-phase characteristics," to be presented at 28th Int. Conf. AES (July 2006)
24. M. Tohyama, R.Lyon and T. Koike, "Pulse waveform recovery in a reverberant condition," J. Acoust. Soc. Am. 91(5) 2805–2812 (1992)
25. M. R. Schroeder, "A new method of measuring reverberation time," J. Acoust. Soc. Am. 37 409–412 (1965)

76

Spatial Information of Sound Fields

Yasuhiro Oikawa

Waseda University, Tokyo, Japan

1 • Introduction

The nature of a sound field can only be fully understood when we have spatial information about it. Being able to visualize a sound field is a very useful way of understanding it. This section describes how to estimate many direction of arrivals (DOAs) by using two microphones, how to get spatial information by using four closely located microphones, and how to visualize and measure sound fields by using a laser Doppler vibrometer (LVB).

Many topics on microphone array processing, e.g., signal enhancement, noise reduction, and DOA estimation, have been actively researched, and in particular, many DOA estimation methods have been suggested. The number of DOAs that can be estimated is limited by the number of microphones, and there is a spatial aliasing problem that depends on the microphone distance and wavelength [1]. The DOA estimation method described below uses a matching pursuit algorithm to estimate more DOAs than there are microphones [2].

The estimation of a sound field in a room requires spatial information, especially the early reflection periods. This information can be gotten from impulse responses measured at four closely located points, the origin and three points along lines (3–5 cm) from the origin on cartesian coordinate axes. From these four impulse responses, the coordinates and powers of virtual image

sources can be calculated with a correlation technique or an intensity technique. Such techniques are used to estimate the sound fields in concert halls, opera theaters, and many other sound fields.

At the end of this section, we describe new sound field measurement methods that use a LDV. By irradiating the reflection wall with a laser, we can observe the light velocity's change that is caused by the refractive index change from the change in air density. This means that it is possible to observe the change in sound pressure. We measured a sound field projection on a 2-D plane by using a scanning laser Doppler vibrometer (SVM), which can visualize a sound field. We also made a 3-D sound field reconstruction from 2-D laser projections based on computed tomography (CT).

2 • DOA Estimation

Our DOA estimation consists of the following steps. First, we separately calculate the normalized power of the array output, $P(\theta)$, for each frequency bin by using the delay-and-sum method [3]. Next, we average $P(\theta)$ over all frequency bins. Finally, we perform peak picking by using a matching pursuit algorithm to estimate the DOA over all frequency bands. The matching pursuit algorithm includes, after each iteration step, a reoptimization of all DOAs found thus far. Its main characteristic is that it is possible to find true DOAs even when the number of sources exceeds the number of microphones. We discuss these steps in detail in the following subsections.

2.1 Calculation of Power of Array Output

The power of the delay-and-sum array output is calculated as

$$P(\theta,\omega) = \mathbf{d}(\theta,\omega)^H \mathbf{R}(\omega)\mathbf{d}(\theta,\omega), \tag{1}$$

where $\mathbf{d}(\theta)$ is the steering vector:

$$\mathbf{d}(\theta,\omega) = [1, \exp(-j\omega\tau), \cdots, \exp(-j\omega(M-1)\tau)]^T. \tag{2}$$

Here, ω is the angular frequency, M is the number of microphones, $\tau = d\sin\theta/c$, d is the distance between microphones, c is the velocity of sound, and $\mathbf{R}$ is the covariance matrix of the array outputs $\mathbf{x}(t)$, i.e.:

$$\mathbf{R}(\omega) = E[\mathbf{X}(\omega,t)\mathbf{X}(\omega,t)^H]. \tag{3}$$

We only need to consider the θ-dependent component of Equation (1), $\hat{P}(\theta, \omega)$. The average of $\hat{P}(\theta, \omega)$ over the frequency bins is

$$\hat{P}_{\text{avrg}}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\hat{P}(\theta, \omega_i) \approx \sum_{k=1}^{K}\hat{P}_{\text{avrg}}(\theta|\theta_k), \tag{4}$$

where N is the number of frequency bins and $\hat{P}_{\text{avrg}}(\theta|\theta_k)$ is the frequency average of the θ-dependent component of the array output power from the kth sound, i.e., the θ-dependent component of the delay-and-sum array output is approximately the sum of that for each source [4].

2.2 Matching Pursuit to Estimate DOA

A matching pursuit algorithm was developed to decompose any signal into a linear expansion of waveforms [5]. We used a modified matching pursuit algorithm that includes a reoptimization step [6] to decompose the signal into a set of direct and reflected sounds [7]. We define the vector of the angles of i DOAs that is estimated during i iterations as

$$\Theta_i = [\hat{\theta}_1, \cdots, \hat{\theta}_i]^T, \tag{5}$$

where Θ_0 is a vector without any elements. The matching pursuit algorithm for DOA estimation consists of the following steps.

Step (1) Define a dictionary as

$$\mathcal{D} = \{\hat{P}_{\text{avrgn}}(\theta|\theta_k)\}_{-\pi/2<\theta_k<\pi/2}. \tag{6}$$

Step (2) Initialization:

$$e_0(\theta) = \hat{P}_{\text{observed}}(\theta) \text{ and} \tag{7}$$

$$i = 1. \tag{8}$$

Step (3) Calculate the residual for all θ_k:

$$e_i(\theta|\theta_k) = e_{i-1}(\theta) - a_{i-1}(\theta_k)\hat{P}_{\text{avrgn}}(\theta|\theta_k), \tag{9}$$

where $a_{i-1}(\theta_k)$ denotes the inner product of $e_{i-1}(\theta)$ and $\hat{P}_{\text{avrgn}}(\theta|\theta_k)$.

Step (4) Select θ_k (estimate DOA $\hat{\theta}_i$):

$$\hat{\theta}_i = \underset{\theta_k}{\text{argmin}} \sum |e_i(\theta|\theta_k)|^2. \tag{10}$$

Step (5) Reoptimize Θ_i (all DOAs) and calculate the residual $e_i(\theta)$:

$$e_i(\theta) = e_0(\theta) - \sum_{l=1}^{i} \hat{a}(\hat{\theta}_l)\hat{P}_{\text{avrgn}}(\theta|\hat{\theta}_l), \tag{11}$$

where $\hat{a}(\hat{\theta}_l)$ is computed using Equation (18).

Step (6) If

$$10\log\frac{\int e_0^2(\theta)d\theta}{\int e_i^2(\theta)d\theta} < \delta, \tag{12}$$

where δ is the stopping criterion, and

$$i = i + 1, \tag{13}$$

return to (Step 3), or else end the procedure.

2.3 Reoptimization of DOAs

The optimization techniques of DOAs are based on the squared difference between the array output and the reconstruction of components for the estimated DOAs [8].

We define the vector of the angles of L DOAs as

$$\Theta = [\theta_1, \cdots, \theta_L]^T. \tag{14}$$

The basis vectors and the observed vector are defined as

$$\hat{\mathbf{p}}_{\text{avrgn}}(\theta_k) = [\hat{P}_{\text{avrgn}}(-\frac{\pi}{2}|\theta_k), \cdots, \hat{P}_{\text{avrgn}}(\frac{\pi}{2}|\theta_k)]^T \text{ and} \tag{15}$$

$$\mathbf{e}_0 = [e_0(-\frac{\pi}{2}), \cdots, e_0(\frac{\pi}{2})]^T, \tag{16}$$

where we have discretized the normalized frequency average of the power of the array output of the kth sound as a function of the continuous steering direction variable θ.

For a given set of DOAs, the analysis matrix containing the basis vectors is constructed according to

$$\hat{\mathbf{P}}_{\text{avrgn}\Theta} = [\hat{\mathbf{p}}_{\text{avrgn}}(\theta_1), \cdots, \hat{\mathbf{p}}_{\text{avrgn}}(\theta_L)]. \tag{17}$$

The projection of $\mathbf{e}_0$ onto a space that is defined by bases $\hat{\mathbf{p}}_{\text{avrgn}}(\theta_1), \cdots$, and $\hat{\mathbf{p}}_{\text{avrgn}}(\theta_L)$ is

$$\hat{\mathbf{a}} = (\hat{\mathbf{P}}^T_{\text{avrgn}\Theta} \cdot \hat{\mathbf{P}}_{\text{avrgn}\Theta})^{-1} \cdot \hat{\mathbf{P}}^T_{\text{avrgn}\Theta} \cdot \mathbf{e}_0, \tag{18}$$

which is from the least-squares residual. Optimum DOAs are those for which the energy of the projection of $\mathbf{e}_0$ onto the column space of $\hat{\mathbf{P}}_{\text{avrgn}\Theta}$ is maximized:

$$\underset{\Theta}{\text{argmax}}\{\hat{\mathbf{P}}_{\text{avrgn}\Theta} \cdot \hat{\mathbf{a}}\}^T \cdot \{\hat{\mathbf{P}}_{\text{avrgn}\Theta} \cdot \hat{\mathbf{a}}\} = \underset{\Theta}{\text{argmax}}\, \mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0.$$

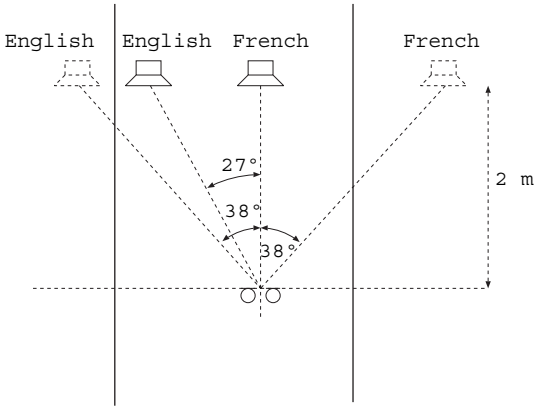


FIGURE 1 Artificial mixture conditions. (From Ref. [8]).

A gradient search based on Newton's method can be used to find the local maximum of $\mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0$ in the neighborhood of a set of initial DOAs.

2.4 Experiments

We estimated DOAs under artificial experimental conditions and compared their estimation performances with and without reoptimization for artificial data. The experiment conditions are shown in Figure 1. Two speech sources and two walls were arranged as shown in Figure 1. The speech sounds were assumed to arrive at microphones from each direct and each first-reflected sound direction. This is equivalent to a mixture with one true source and one image source for each independent signal.

We used 64 mixtures with 8 English and 8 French speeches, which were spoken by one female and one male for each language. The average speech length was 5 s. We estimated DOAs for each sentence and averaged the 64 results. The distance between microphones was 5 cm. We used a data sampling frequency of 44.1 kHz, a frame length of 46 ms, and a frame update of 23 ms.

The DOA estimation results were much better with than without reoptimization (Table 1).

TABLE 1 DOA estimation

Method	DOA [degree]
True	–38.0, –27.0, 0.0, 38.0
Without reoptimization	–34.0, –23.0, –1.0, 39.0
With reoptimization	–38.5, –27.7, –0.3, 38.2

From Ref. [8].

3 • Spatial Information Measurement: Four Closely Located Point Microphones Method

Reverberation times and sound pressure levels are very important parameters to evaluate sound fields [9–10]. However, we often get quite different acoustic impressions in sound fields that have about the same reverberation time and sound pressure level. This difference might come from spatial information differences, especially in early reflection periods. Hence, it is important to get spatial information for sound fields. A lot of effort has gone into solving this problem; for example, it is possible to use a microphone that has very sharp directivity. Here, we will discuss a way to get spatial information from impulse responses measured at four closely located points, the origin and three points at the same distance (3–5 cm) from the origin on cartesian coordinate axes. From the small differences among these four impulse responses, we try to estimate the positions and powers of direct and reflected sound sources by using a correlation technique or an intensity technique. This method could also be called 3-D triangular surveying [11–13].

3.1 Microphone Arrangement

The four-point method uses four impulse responses to get the spatial information for a sound field. Usually, we use a four-point microphone system as shown in Figure 2. One microphone is located at the origin of the coordinates and three on cartesian coordinate axes at the same distance from the origin. Theoretically, the four microphones should not be placed on the same plane and the distances of microphones should be sufficiently shorter than the sound wavelength. Although shorter separations are better in theory, in practice, we must consider the

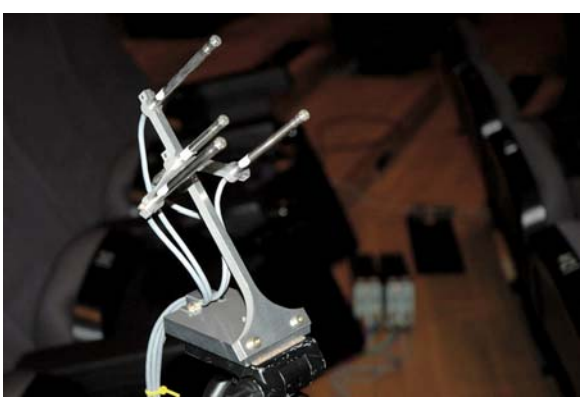


FIGURE 2
Four-point microphone.

accuracies of the mechanical set up and the numerical computation when deciding the distance. We used 1/4-inch diameter diaphragm electrostatic microphones with the distance to the origin of 50 mm for ordinary measurements in concert halls or opera houses.

3.2 Calculation of Coordinates of Virtual Image Sources

We can calculate the coordinates of the virtual image sources from the measured impulse responses as follows. The local peaks in the impulse responses represent the traveling distance of the reflected wave. The sound source corresponding to each local peak should be on a spherical surface of radius r_{in} whose center is at the observation point as shown in Figure 3, where $r_{in}(i = o,x,y,z)$ expresses the traveling distance from each microphone. The four local peaks make four spheres. The intersection of these spheres is the point of the sound source corresponding to the direct sound or the reflected sound. Of course, sound fields are very complicated, and we can estimate only the source points corresponding to the direct sound and a few reflections. As will be shown later, the estimated source points corresponding to the reflections are considered to be of degenerated equivalent reflections. We thus call them virtual image sources. The distances are calculated as

$$r_{on} = c \cdot t_{on}, \tag{19}$$

$$r_{xn} = c \cdot t_{xn}, \tag{20}$$

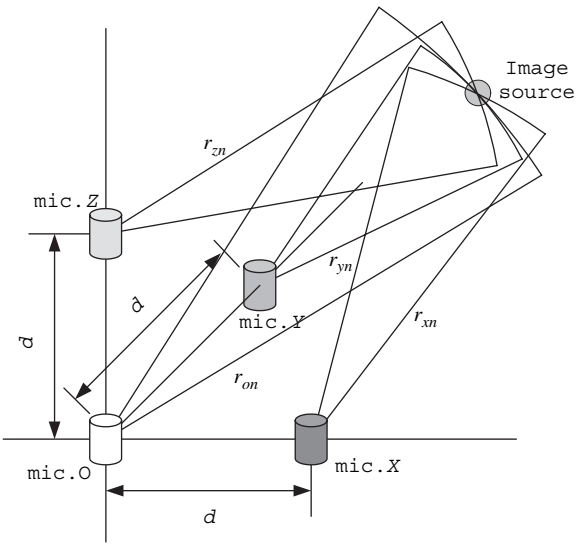


FIGURE 3 Reflection calculation.

$$r_{yn} = c \cdot t_{yn}, \tag{21}$$

$$r_{zn} = c \cdot t_{zn}, \tag{22}$$

where c is the velocity of sound and t_{on}, t_{xn}, t_{yn}, and t_{zn} correspond to the distances between a certain source and each observation point. The coordinates of a virtual image source (X_n, Y_n, Z_n) from the spherical equations are

$$X_n = (d^2 + r_{on}^2 - r_{xn}^2)/2d, \tag{23}$$

$$Y_n = (d^2 + r_{on}^2 - r_{yn}^2)/2d, \tag{24}$$

$$Z_n = (d^2 + r_{on}^2 - r_{zn}^2)/2d, \tag{25}$$

where d is the distance from the origin to the microphones.

3.3 Correlation Technique

Here, we will explain the process to calculate the coordinates of the virtual image sources by correlation:

1. Four impulse responses are interpolated 256 times to keep the resolution high enough in the following process.
2. The corresponding reflections are selected by use of the cross-correlation function.
3. The coordinates of the virtual image source are calculated by Equations (23), (24), and (25).

3.4 Short-Time Intensity Technique

Four microphones are set up at the origin and along each of the cartesian coordinate axes. Thus, we can calculate the intensity vector for three orthogonal directions. From the short duration intensity, we can determine the direction and distance of the virtual image sources in polar coordinates as well as their power. Strictly speaking, these are not image sources, but the equivalent centers of gravity or reflections of moments. For the sake of convenience, we also call these virtual image sources. By using the intensity technique, we can easily calculate virtual image sources within limited frequency bands.

3.5 Distribution of Virtual Image Sources

Figures 4 and 5 show impulse responses and distributions of virtual image sources of the Musikvereinssaal and Munchen Philharmonie Hall [14–15]. The distribution of the virtual image sources are projected on three planes determined with the correlation technique. The center of the circle represents the estimated points of the virtual image source, and the area of each circle represents the power of the corresponding source. The crossing point of the two orthogonal lines is the observation point, and the distance between the sound source and the observation point is 12 m. The outline of the concert hall is also shown.

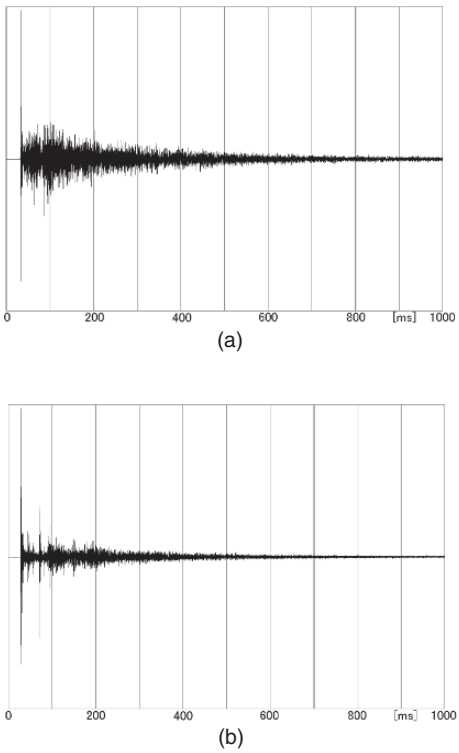


FIGURE 4 Impulse response, (**a**) Musikvereinssaal and (**b**) Munchen Philharmonie Hall.

Time-divided distributions of virtual image sources can also be displayed in color.

3.6 Directivity Patterns (Hedgehog Patterns)

We can calculate directivity patterns from the 3-D distributions of virtual image sources through the supposed microphone with vertical open angle degrees. Figure 6 shows directivity patterns of Musikvereinssaal and the Munchen Philharmonie Hall. This shows the total power of the virtual image sources with a horizontal open angle of 1°, vertical open angle of ±45°, and 1 division represents 10 dB.

4 • Sound Field Measurement Using a Laser

Visualizations can help us to understand the behavior of the sound field. Many researchers have recently studied numerical computations and sound field visualizations because it has become easier to use high-performance computers [16]. Sound fields visualized as real-world experiences have also been researched. It is difficult to visualize real-world sound fields because we have to measure the sound pressure or intensity

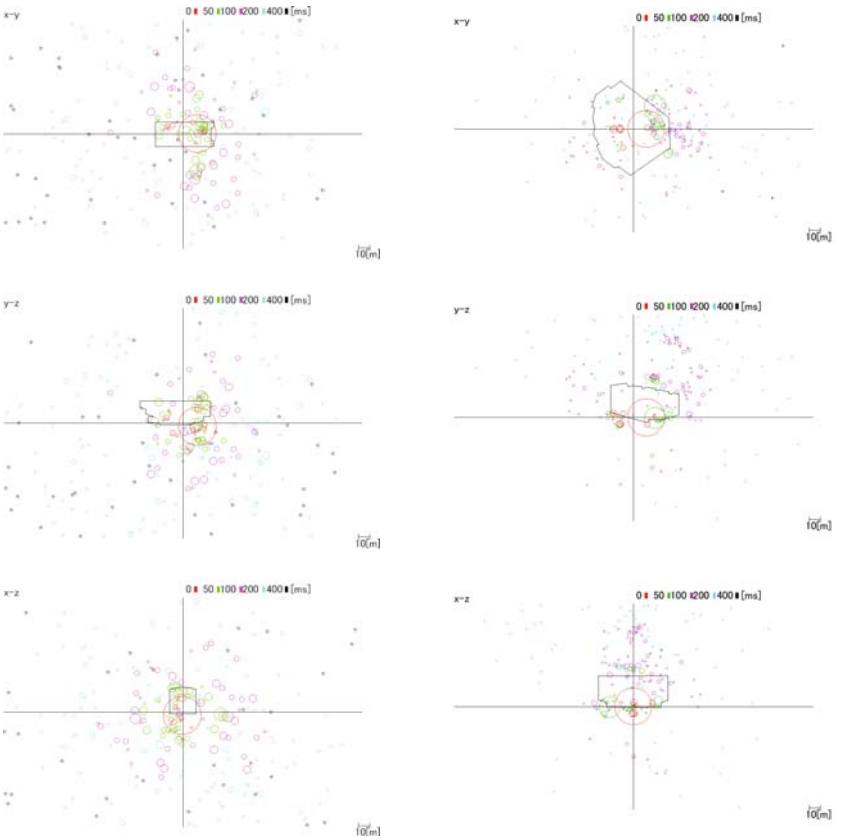


FIGURE 5 Distribution of virtual image sources, (**a**) Musikvereinssaal and (**b**) Munchen Philharmonie Hall.

correctly at so many points using microphones. On the other hand, the four-point microphone method can acquire the spatial information from four impulse responses measured using four closely located microphones [13]. We have measured many sound fields by using this method.

We usually use a number of microphones to measure sound fields. Another way is to measure sound pressure level of a broad sound field by using a LDV to measure the average sound pressure level over the laser path. The measured value includes all information on the path of laser. If we have LDV-measured data on an area from different directions, we can estimate the sound field in the area without having to measure at many points using microphones. This kind of signal processing is known as CT based on reconstruction from projections. X-ray CTs are used in

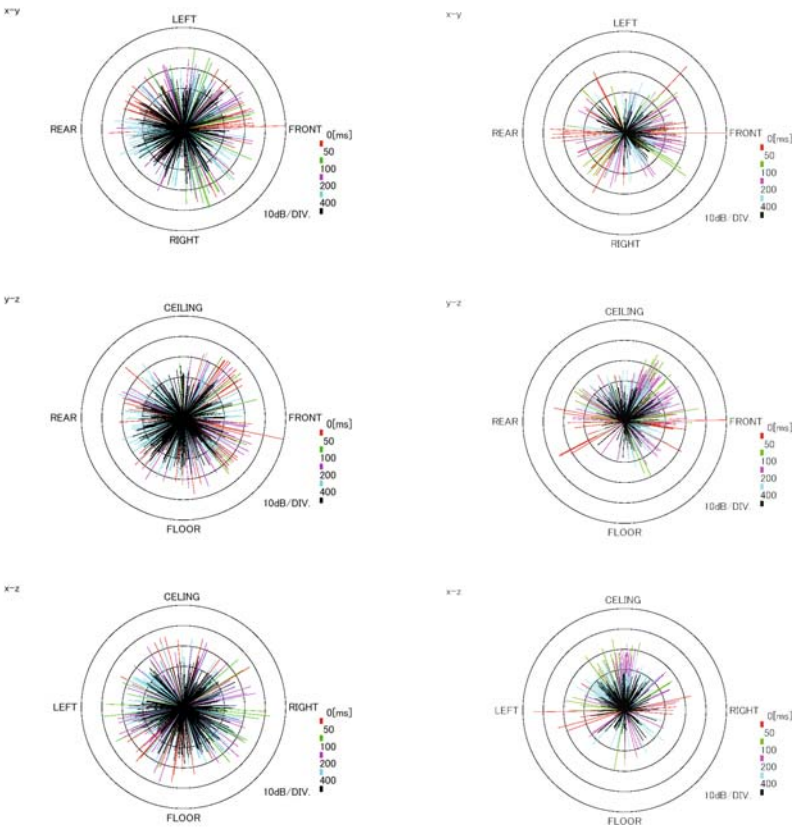


FIGURE 6 Directivity patterns, (**a**) Musikvereinssaal and (**b**) Munchen Philharmonie Hall.

medicine to observe inside the body without contact or damage. Similarly, we can observe a sound field by using laser CT.

Below, we describe sound pressure measurements using LDV and also suggest a way of sound field reconstruction from laser projections. With this method of reconstruction, we can visualize 2-D or 3-D sound fields in the real world.

4.1 Sound Pressure Measurement Using LDV

It is possible to measure changes in the sound pressure from changes in the refractive index of air [17]. From the state equation for the adiabatic change of the gas, we get the relationship between the volume change and the refractive index change:

$$\frac{\Delta V}{V} = -\frac{\Delta n}{n-1}, \tag{26}$$

where V is the volume and n is the refractive index. Furthermore, we have

$$\frac{P}{P_0} = -\gamma \frac{\Delta V}{V} \text{ and} \tag{27}$$

$$P_0 \gamma = c^2 \rho, \tag{28}$$

where P_0 is the atmosphere pressure, γ is the specific heat ratio, ρ is the density, and c is the sound velocity. From Equations (26), (27), and (28), we get

$$\Delta n = \frac{n-1}{c^2 \rho} P. \tag{29}$$

If we pass the laser from the LDV to the sound field against the reflection wall and measure the reflected laser light, we can calculate the refractive index change that produces the sound pressure changes according to distance. Hence, the relationship between the refractive index change and the distance change is

$$\Delta n \cdot l = n \cdot \Delta l. \tag{30}$$

Using LDV, the refractive index change is

$$\Delta n = \frac{n}{2\pi f l} v_{\mathrm{LDV}}, \tag{31}$$

where l is the distance of the laser path, f is the frequency of sound, and v_{LDV} is the measured velocity. From Equations (29) and (31), we get

$$P = \frac{n}{n-1} \frac{c^2 \rho}{2\pi f l} v_{\mathrm{LDV}}. \tag{32}$$

4.2 Visualization of Sound Fields by Using SVM

We visualized the sound field by using the sound pressure measurement method, which is described in Section 4. We measured the sound pressure at many points using a SVM and visualized the sound pressure distribution. Figure 7 shows the experimental conditions. We put a loudspeaker between the reflection wall and the laser head. In these experiments, we used an ordinary loudspeaker (NS10M from Yamaha) and a flat panel loudspeaker (W-3232 from Waseda EE) [18]. The measurement signal was a 4-kHz sinusoidal tone. We used a heavy concrete wall of the reverberant chamber as the reflection wall. Measurements from another LDV indicated that this wall did not move.

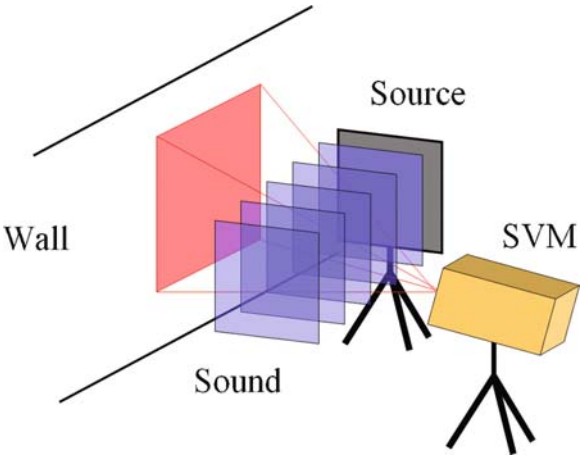


FIGURE 7 Measurement conditions. The loudspeaker is located between SVM and the reflection wall. (From Ref. [22]).

Figures 8 and 9 show the sound fields made by the ordinary and flat panel loudspeakers.

4.3 Reconstruction from Laser Projections

Reconstruction from projections was first suggested by Radon [19], and this sort of signal processing is known as CT based on reconstruction from projections [20–21]. We performed sound field reconstruction experiments based on the laser



FIGURE 8 Sound field made by ordinary loudspeaker. A spherical sound wave is constructed.

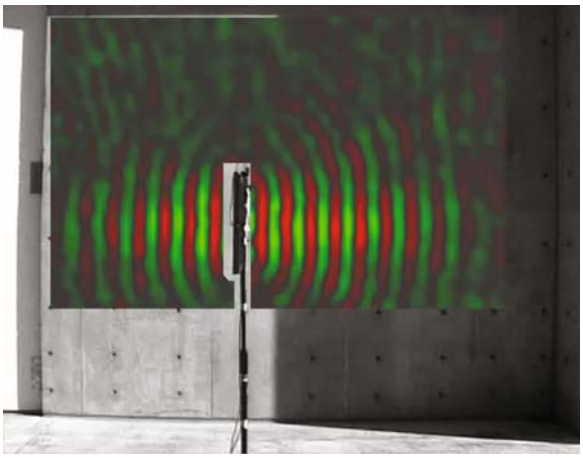


FIGURE 9 Sound field made by flat panel loudspeaker. A planer sound wave is constructed.

CT technique [22–23]. We also used an ordinary loudspeaker (Yamaha NS10M) and a flat panel loudspeaker (Waseda EE W-3232). The loudspeakers were put face up on a table in the echoic chamber as shown in Figure 10. Because the symmetric condition was satisfied for these speakers, we rotated them instead of the LDV head. The laser scanned a plane 10 cm above the loudspeaker. There were 100 scanning points. The interval between measurement points was about 10 mm. We took 36 laser projections and used them to make the reconstruction. The measurement signal was 4 kHz sinusoidal.

Figures 11 and 12 show the 3-D sound field reconstructions. Because the reflection wall was irradiated with laser light having fan beam shape, we reconstructed from fan beam projections.

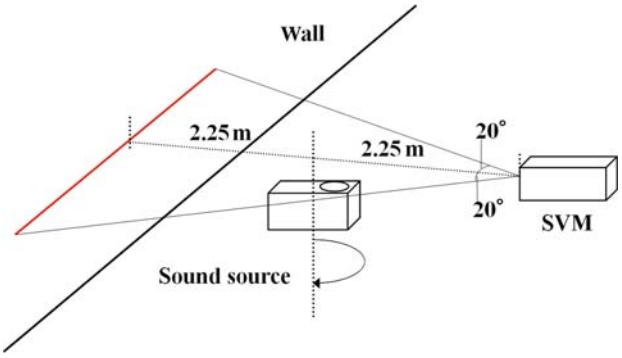


FIGURE 10 Experimental conditions.

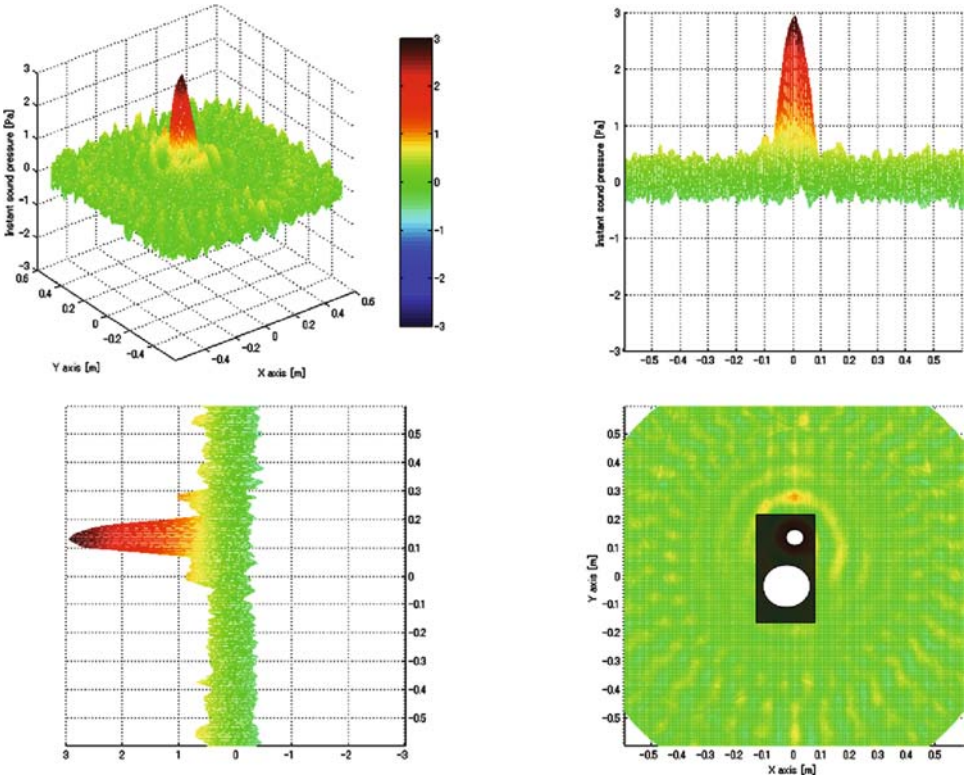


FIGURE 11 Reconstruction from 36 projections. A spherical sound wave is made by the ordinary loudspeaker.

4.4 Sound Field Reconstruction in the Room

We reconstructed the sound field in the room from the laser projections. We took this measurement in the echoic chamber with sound-absorbing materials. We used the non-directional loudspeaker as the point sound source. We also took 24 laser projections around the loudspeaker. Figure 13 shows the sound field reconstruction. We found the direct sound near the loudspeaker and the reflected sounds on the outside.

5 • Summary

We can estimate more DOAs than microphones by using a matching pursuit algorithm under reverberant conditions. Source separation has recently been studied by many researchers to separate sources under real-world conditions [24–35]. Thus, one of the research topics is to separate sources for convoluted mixtures consisting of long impulse responses in acoustics.

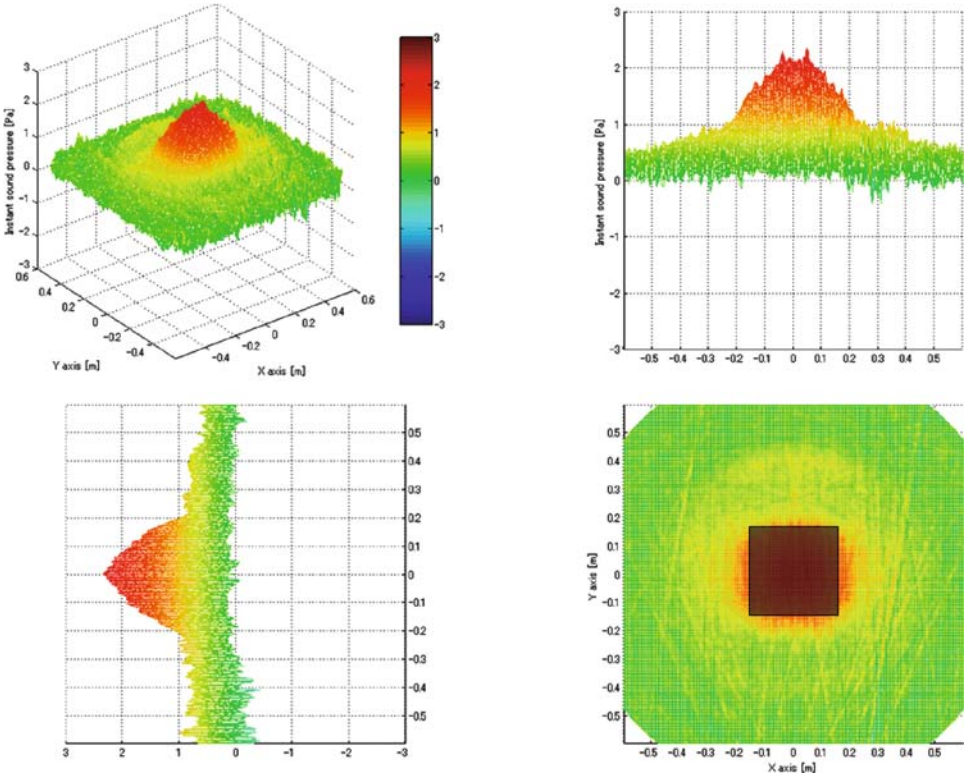


FIGURE 12 Reconstruction from 36 projections. A flat planer sound wave is made by the flat panel loudspeaker.

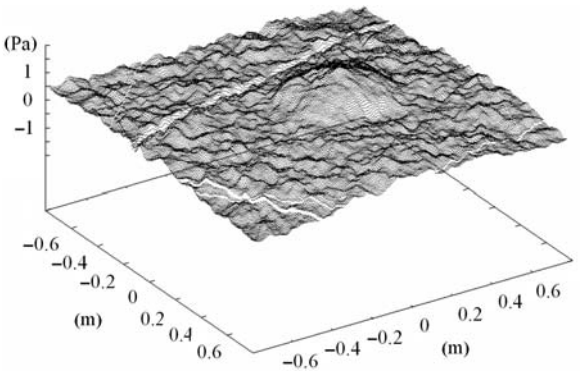


FIGURE 13 Reconstruction of the room sound field. (From [7]).

We expect that additional improvements can be achieved by considering the spatial information derived from this method.

We can get the spatial information in a sound field by using a four-point microphone method. We discussed the principle of this method, measurements in actual sound fields, and the development of virtual image sources. The method can be used to show spatial information of a sound field visually [the rewrite is a guess.]. For example, we could see that the Musikvereinssaal has much richer side reflections compared to the Munchen Philharmonie. The former also has rich distributions from the bottom; maybe the floor material of this hall is much harder. We have used the four-point microphone method to measure the sound fields of many concert halls and opera theaters all over the world. We introduced this technique to scale model measurements, as well, and compiled databases to estimate the sound fields and feed them back to acoustic designs.

As for sound field measurements using SVM, we could measure the sound pressures by using LDV and visualize the sound pressure distribution. We were also able to use CT to reconstruct a sound field visually from laser projections. Although we assumed that the laser beams were parallel, there is a cone-beam algorithm for CT [36], and we expect that we can use it to achieve additional improvements in real-world measurements.

References

1. J. Ohga, Y. Yamasaki and Y. Kaneda, *Acoustic System and Digital Processing* (Korona Publishing Co., Ltd., Tokyo, 1995) (in Japanese).
2. Y. Oikawa and Y. Yamasaki, "Direction of arrival estimates by two microphones using matching pursuit," *Proc. Autumn Meet. Acoust. Soc. Japan.*, pp. 501–502, 2003 (in Japanese).
3. D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Prentice-Hall Inc., New Jersey, 1993).
4. Y. Oikawa and Y. Yamasaki, "Direction of arrival estimation using matching pursuit under reverberant condition," *Acoust. Sci. & Tech.*, **26**(4), 365–367, 2005.
5. S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, **41**(12), 3397–3415, 1993.
6. K. Vos, R. Vafin, R. Heusdens and W. B. Kleijn, "High-quality consistent analysis-synthesis in sinusoidal coding," *Proc. 1999 Audio Eng. Soc. 17th Conf. "High Quality Audio Coding"*, pp. 244–250, 1999.
7. Y. Oikawa and Y. Yamasaki, "Direction of arrival estimates using matching pursuit," *Proc. ForumAcusticum2005*, pp. 2619–2624, 2005.
8. Y. Oikawa and Y. Yamasaki, "Direction of arrival estimation using matching pursuit and its application to source separation for convolved mixtures," *Acoust. Sci. & Tech.*, **26**(6), 486–494, 2005.

9. N. Aoshima, "Computer-generated pulse signal applied for sound measurement," *J. Acoust. Soc. Am.*, **69**, 1484–1488, 1981.
10. H. Yanagawa, Y. Yamasaki and T. Itow, "Effect of transient signal length on cross-correlation functions in a room," *J. Acoust. Soc. Am.*, **84**, 1728–1733, 1988.
11. Y. Yamasaki, S. Yamasuda, H. Saito, K. Yasukawa and T. Itow, "An application of digital techniques to obtain the spatial information of a room," *Joint Meet. ASA/ASJ*, BBB6, (1978).
12. K. Endoh, T. Horikoshi, Y. Yamasaki and T. Itow, "Grasp and estimation of spatial information method," *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, pp. 909–912, 1986.
13. Y. Yamasaki and T. Itow, "Measurement of spatial information in sound fields by closely located four point microphone method," *J. Acoust. Soc. Japan(E)*, **10**(2), 101–110, 1989.
14. Y. Yamasaki, H. Tachibana, M. Morimoto, Y. Hirasawa and Z. Maekawa, "Survey on the acoustics of concert halls in European countries (The 2nd Report) – Measured results by four-point microphone method," *J. Acoust. Soc. Japan.*, **43**, 277–285, 1987 (in Japanese).
15. H. Tachibana, Y. Yamasaki, M. Morimoto, Y. Hirasawa, Z. Maekawa and C. Posselt, "Acoustic survey of auditoriums in Europe and Japan," *J. Acoust. Soc. Japan.(E)*, **10**, 73–85, 1989.
16. T. Yokota, M. Iiduka, S. Sakamoto and H. Tachibana, "Differences of sound pressure level distribution due to the variations of room shape," *Proc. WESPAC VIII*, 2003.
17. K. Nakamura, "Measurements of high power ultrasound in the air through the modulation in the refractive index of air," *Tech. report IEICE*, US2001-9, pp. 15–20, 2001 (in Japanese).
18. Y. Yamasaki, H. Nakajima, M. Okazaki and Y. Kato, "Transducers considering energy saving -super conductive and electrostatic loudspeaker driven by 1-bit amplifier-," *Proc. Int. Cong. Acoust.*, **IV**, pp. 2427–2432, 2004.
19. J. Radon, "Uber die Bestimmug von Funktionen durch Ihre integralwerte Laengs Geweisser Mannigfaltigkeiten," *Berichte Saechsishe Acad. Wissenschaft. Math. Phys.*, Klass **69**, pp. 262–277, 1917.
20. A. C. Kak and Malcolm Slaney, "Principles of computerized tomographic imaging," *IEEE Press*, 1998.
21. L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, **NS-21**, pp. 21–43, 1974.
22. Y. Oikawa, M. Goto, Y. Ikeda, T. Takizawa and Y. Yamasaki, "Sound field measurements based on reconstruction from laser projections," *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, **IV**, pp. 661–664, 2005.
23. Y. Ikeda, M. Goto, N. Okamoto, T. Takizawa and Y. Yamasaki, "A measurement of reproducible sound field with laser computed tomography," *J. Acoust. Soc. Japan*, 62(7), 2006.
24. B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Processing Mag.*, **5**, 4–24 (1988).

25. A. J. Bell and T. J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," *Neural Computation*, **7**, 1129–1159 (1995).
26. J. F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, **86**(10), 2009–2025 (1998).
27. S. Amari and A. Cichocki, "Adaptive blind signal processing – Neural network applications," *Proc. IEEE*, **86**(10), 2026–2048 (1998).
28. T. W. Lee, *Independent Component Analysis: Theory and Applications* (Kluwer Academic Publishers, Boston, 1998).
29. P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, **22**, 21–34 (1998).
30. S. Ikeda and N. Murata, "A method of ICA in time-frequency domain," *Proc. Int. Workshop Independent Comp. Analysis Blind Sign. Separation (ICA'99)*, pp. 365–371 (1999).
31. S. Kurita, H. Saruwatari, S. Kajita, K. Takeda and F. Itakura, "Evaluation of blind signal separation method using directivity pattern under reverberant conditions," *Proc. IEEE Int. Conf. Acoust., Speech, Sign. Process.*, pp. 3140–3143 (2000).
32. H. Saruwatari, T. Kawamura and K. Shikano, "Blind source separation for speech based on fast-convergence algorithm with ICA and beamforming," *Proc. Eurospeech*, pp. 2603–2606 (2001).
33. L. C. Parra and C. V. Alvino, "Geometric source separation: Merging convolutive source separation with geometric beamforming," *IEEE Trans. Speech and Audio Process.*, **10**(6), 352–362 (2002).
34. A. Jourjine, S. Rickard and Ö. Yilmaz, "Blind separation of disjoint orthogonal signals: demixing n sources from 2 mixtures," *Proc. IEEE Int. Conf. Acoust., Speech, Sign. Process.*, pp. 2985–2988 (2000).
35. R. Gribonval, "Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture," *Proc. IEEE Int. Conf. Acoust., Speech, Sign. Process.*, pp. 3057–3060 (2002).
36. L. A. Feldkamp, L. C. Davis and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am.*, **1**(6), pp. 612–619, 1984.

PART XIV
INFRASONICS

Douglas Christie

Earth Physics, Research School of Earth Sciences, The Australian National University, Canberra, ACT, Australia

77 Infrasound Event Detection Using the Progressive Multi-Channel Correlation Algorithm

Yves Cansi and Alexis Le Pichon 1425

Introduction • Correlation: The Basic Component of PMCC • PMCC Configuration • Processing Examples

78 Ray Tracing in an Inhomogeneous Atmosphere with Winds

David J. Brown and Milton A. Garcés 1437

Historical Introduction • Atmospheric Variability • Geometrical Acoustics • Concluding Remarks

79 Infrasonic Data Processing at the Provisional Technical Secretariat of the Comprehensive Nuclear-Test-Ban Organization

Nicolas Brachet, Petr Firbas, Thomas Hoffmann, Paola Campus, Stefka Stefanova 1461

Overview of IDC Processing • Infrasonic Station Processing at the IDC • Automatic Network Processing • Interactive Processing • Post-location Processing

80 The Design and Performance of Infrasound Noise-Reducing Pipe Arrays

Benoit Alcoverro 1473

Introduction • Historical Background • Design Principles for a Robust Infrasonic Noise-Reducing System • Noise-Reducing Pipe Array Design and Performance • Conclusion

81 Processing Infrasonic Array Data

John V. Olson and Curt A.L. Szuberla 1487

Array Processing • Bearing Estimation • Signal Detection and Enhancement

82 Infrasound Propagation

Rodney W. Whitaker and David E. Norris 1497

Introduction • Relevant Physics and Issues • Propagation Models • Summary and Future Direction

77

Infrasound Event Detection Using the Progressive Multi-Channel Correlation Algorithm

Yves Cansi and Alexis Le Pichon

CEA/DASE/LDG, BP12, 91680 Bruyères-le-Châtel, France

1 • Introduction

Infrasonic waves can be represented by a set of plane waves using the well-known relation $f(\vec{r},t) = e^{i(\omega t - \vec{k}\vec{r})}$, where $\|\vec{k}\| = 2\pi f/c$ is the wave vector, f is frequency, c is phase velocity and $\omega = 2\pi f = 2\pi/T$ is the angular frequency. The frequency content of a recorded wave can easily be determined using data from a single station. In contrast, data from an array of at least three stations are needed to specify the propagation parameter $\vec{k}$. When the aperture of the array of sensors is roughly equal to the wavelength of the wave, the array is referred to as a mini-array. At the opposite extreme, when the aperture is far larger than the wavelength, the array is referred to as a network of sensors.

In the case of a network, the morphology of the signal is often very different from one sensor to another and the measure of the propagation parameters is usually derived from the set of arrival times by an inversion process, as described by Husebye (1969). In the case of a mini-array, we can take advantage of the high degree of signal coherence to compute arrival time differences using classical techniques of signal processing theory. Husebye's method may then be used to evaluate the propagation parameters from the derived set of arrival time differences.

The classical method for estimating these wave parameters in the case of a mini-array is based on a systematic search in a specific wave vector domain. For example, the area defined in the wavenumber plane by the relation $\|\vec{k}\| < 2\pi f/V_{\min}$ corresponds to all the waves from any azimuth with a frequency f and a velocity $V > V_{\min}$. The time delay of each sensor is calculated for each discrete wave vector in this regularly discretized domain and the delayed signals are summed. If the data consist primarily of random background noise, the energy variation of the sum will be small over the entire wave vector field. In contrast, if the data include a signal with specified wave vector $\vec{k}_0$ and if the signal-to-noise ratio is large, then the energy corresponding to $\vec{k}_0$ will usually be much larger than the energy for all the other vectors. A number of methods have been proposed by different authors to find the wave vector that produces the maximum energy (Capon, 1969). This is not a trivial problem because data are discrete in the spatial domain (since only a few sensors are used). This leads to the possibility that for each frequency, false results may be obtained due to spatial aliasing of signals over one or more periods. Capon suggested that these effects can be studied by computing the array beam forming function : $B(\vec{k}) = \left|\sum e^{i\vec{k}\cdot\vec{r}_j}\right|$. This represents the array response to a downward propagating vertical wave, i.e. to a wave that arrives at all sensors at the same time. The measured phase velocity of the wave is then infinite, which leads to a maximum amplitude at $\vec{k} = 0$. The main assumption in the classical methods is the search for a signal that is modelled by a plane wave recorded on all the sensors of the array. In practice, due, for example, to local heterogeneities, which can generate interferences, this assumption seems to be too restrictive. A more flexible analysis method with fewer model-related constraints is needed. The method proposed here to detect a stable signal in two or more data records is based on conventional signal processing techniques with a partial relaxation of the requirements of a rigid plane wave model. We have named this method the progressive multi-channel correlation (PMCC) algorithm.

2 • Correlation: The Basic Component of PMCC

A signal in the time domain, $s(t)$, can also be represented by its Fourier transform $S(f) = A(f)e^{i\varphi(f)}$, where $A(f)$ represents the spectral amplitude and $\varphi(f)$ is the phase. The background noise is characterized by a rapid variation of both $A(f)$ and $\varphi(f)$ from one sensor to another, even when the sensors are separated by less than one wavelength. We assume, in the case of a signal, that the following relations are valid between any two sensors in the array:

$$A_2(f) = A_1(f)$$
$$\varphi_2(f) = \varphi_1(f) - \Theta(\vec{r}_2 - \vec{r}_1).$$

These relations indicate that the morphology of the signals is identical at all sensors in the array; the only difference is a time delay that depends on the relative positions of the sensors (i.e. $\Theta(\vec{r}_2 - \vec{r}_1) = \vec{k}(\vec{r}_2 - \vec{r}_1)$ in the case of a plane wave).

Based on these two assumptions, the correlation function $Cor_{ij}(\tau)$ between $s_i(t)$ and $s_j(t)$ can be used to detect whether a signal is present in the records. $Cor_{ij}(\tau)$ is computed in the time domain and is normalized to values ranging from -1 to 1. Taking into account all frequencies, the correlation function evaluated over a window of length W provides a measure of the similarity of the signals when shifted relative to each other by a time interval τ. The maximum of the correlation function gives the time delay between the signals.

Correlation is the basis of the PMCC algorithm. This algorithm can be used to detect signals with small signal-to-noise ratio and to accurately measure the detected wave parameters (Cansi, 1995; Cansi and Klinger, 1997).

The first step in the implementation of the PMCC method is the use of the correlation function to measure the time delay Δt_{ij} between all pairs of signals, $s_i(t)$ and $s_j(t)$, in each triangular sub-array. In the case of a wave propagating without distortion (*a fortiori* in the case of a plane wave), this delay is the same for all frequencies in the contributing signals :

$$\Delta t_{ij} = \frac{1}{2\pi f}(\varphi_j(f) - \varphi_i(f)).$$

For each group of three sensors, the sum of the time delays for a plane wave signal obeys a closure relation:

$$\Delta t_{ij} + \Delta t_{jk} + \Delta t_{ki} = 0.$$

In contrast, in the case of background noise, the phase is very unstable from one frequency to another. Thus, the delays measured in this case are the result of random phase combinations. These delays are independent of the amplitude of each elementary wave component and are therefore random. Hence, the closure relation given above is no longer valid.

The use of correlation functions to measure time delays provides a very powerful, self-contained detector. This detector uses only the intrinsic information in the processed signals. This method enables a decision to be made on whether there is a signal in a set of simultaneous records, independent of any information in previous records. Furthermore, the method can be used to detect signals in short-duration records (a few signal periods).

The second step in the implementation of the PMCC method is concerned with a progressive procedure that can be used to avoid interference due to large random signals on certain sensors and to minimize potential errors due to spatial aliasing when the correlation is performed on sensors that are very far apart. This part of the computation is initialized using a sub-network R_n composed of n closely spaced sensors. The consistency of the set of delays obtained using all the sensors of R_n is defined as the mean quadratic residual of the closure relations:

$$\left.\begin{aligned} r_{ijk} &= \Delta t_{ij} + \Delta t_{jk} + \Delta t_{ki} \\ c_n &= \sqrt{\frac{6}{n(n-1)(n-2)} \sum_{i>j>k} r_{ijk}^2} \end{aligned}\right\} \; i,j,k \in R_n.$$

If this consistency is below a certain threshold, $c_{\text{Threshold}}$, a detection is declared on R_n. Attempts can now be made to add new stations in this detection. In order to avoid potential problems due to spatial aliasing when signals from distant sensors are included, the propagation information obtained on R_n is used to "direct" the search of other sensors. The delays measured on R_n can be used to estimate the propagation parameters of the detected wave, assuming it to be planar. Using these parameters, the value of the expected delay for a pair of sensors that includes one sensor from R_n and a new one from outside R_n can be estimated. The new measured time delay corresponds to the relative maximum that is the closest to the computed value. As long as the detection criterion remains valid, the use of sensors from increasingly greater distances gives more and more precise propagation parameters since the aperture of the network increases with each new sensor.

Processing using the PMCC algorithm is performed consecutively in several frequency bands and in adjacent time windows

covering the whole signal. Each elementary detection is therefore defined by four parameters:

– the number of sensors in the final sub-network;
– the associated consistency;
– the detected wave velocity;
– azimuth.

A given wave may be represented by several elementary detections (corresponding, for example, to different frequency bands or to adjacent windows for which propagation parameters – namely velocity and azimuth – are similar). Conversely, several waves with different parameters may coexist in the same time window but in different frequency bands. Each wave must be identified separately. To do this, a nearest-neighbour search method, which forms aggregates of elementary detections in (t, f, V, θ) space, is used. In this four-dimensional space, different distances can be defined to connect all sufficiently close points. The more usual definition is a weighted distance defined by $d(P_1, P_2) = \sqrt{\frac{(t_2-t_1)^2}{\sigma_t^2} + \frac{(f_2-f_1)^2}{\sigma_f^2} + \frac{(V_2-V_1)^2}{\sigma_V^2 V_2 V_1} + \frac{(\theta_2-\theta_1)^2}{\sigma_\theta^2}}$, where P_1 and P_2 are two elementary detections.

The σ s are used as weighting factors to allow the comparison of quantities with different meanings. They represent the accepted variation of the corresponding parameters for a given wave. Nevertheless, because slowness – and not velocity – is the direct result of the computation, the weighting formulation applied to the velocity in the weighted distance equation is given in terms of a relative coefficient, leading to a dimensionless σ_v in contrast to other σ s.

3 • PMCC Configuration

Depending on the type of signals to be processed (e.g. waveforms and frequency content) and the background noise level, the PMCC parameters need to be set in order to ensure optimum detection by providing a compromise between a low-detection threshold and low number of false detections. Threshold values that are too large will result in a low level of detection with a relatively high number of false detections. At the opposite extreme, threshold values that are too small will not be appropriate for the detection of coherent waves with poor signal-to-noise ratio even though the number of false detections is reduced.

The three main parameters that need to be specified are the following:

– *WindowLength* is the length of the time window used to compute the correlation. It cannot be lower than the maximum propagation time of the wave between two sensors. This value is generally increased to ensure a smoothing of the time base, which significantly reduces the number of false detections and improves statistics on the measurements of the wave parameters. A reduction in the number of PMCC detection families (aggregates of pixels that in the time/frequency/speed/azimuth domain) is then expected, providing a synthesis of information. In the case of long-duration phenomena such as microbaroms or Mountain-Associated Waves (MAW), *WindowLength* can contain several periods (up to 10). In the case of transient signals, *WindowLength* can also be greater than the maximum propagation time between sensors.
– *TimeStep* is the time shift between two consecutive windows. It cannot exceed *WindowLength.* It defines the time width of the elementary detection in the time/frequency domain, which we refer to as a detection pixel. A small value for *TimeStep* provides greater resolution in the time variation of the wave parameters without interfering with the detection performance.
– *ThresholdConsistency* (s) is the maximum consistency threshold for detection. This parameter is the first criterion that is used to reject a detection pixel, and the choice of its value has a major impact on the detection performance. *ThresholdConsistency* has to be higher than the sampling period, otherwise the closure relation will never be satisfied and no detection will be obtained. In unfavourable background noise conditions, which include interferences in the inter-correlation functions, the *ThresholdConsistency* must be raised. However, a value that is set too high will noticeably increase the number of false detections due to the incoherent nature of the background noise. The optimal tuning of this value is achieved by taking into account the frequency band that is being studied and the mean noise level prevailing at the station.

4 • Processing Examples

4.1 Infrasound from Volcanoes

The first example is a long sequence of signals associated with volcanic activity at Lopevi and Yasur volcanoes in Vanuatu (Le Pichon et al., 2005). Infrasonic waves from both of these volcanoes were recorded simultaneously at the I22FR infrasound station installed in New-Caledonia, which is part of the global infrasonic network of the international monitoring system (IMS) (Figure 1). Figure 2 displays the PMCC results. Shading is used to distinguish azimuth and trace velocity in a frequency

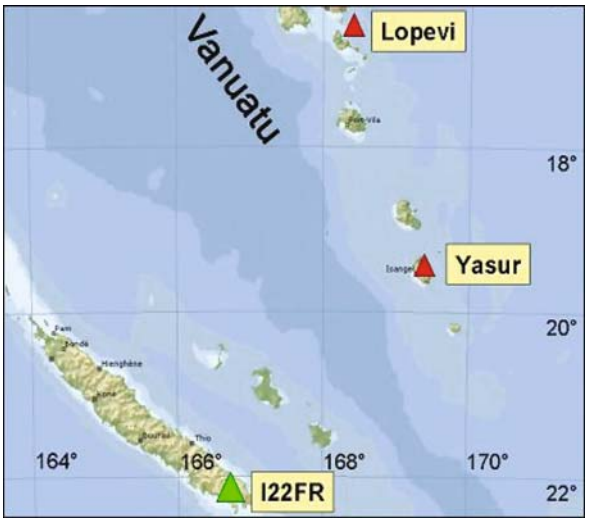


FIGURE 1 Map of the Vanuatu region showing two active volcanoes detected by infrasound station I22FR located in New-Caledonia.

versus time representation. The same results are plotted in a polar representation where the radius is the velocity and the angle is the azimuth. The three groups of detections enclosed by ellipses correspond to the three types of event previously described (the two volcanoes in the North and North-East directions and the microbaroms in the East). This example shows that infrasound signals with a great variety of characteristics and low signal-to-noise ratios can be detected using the PMCC algorithm.

4.2 Infrasound from Earthquakes

The second example was recorded at I34MN installed in Mongolia, also part of the IMS network (Figure 3). The signal displayed in Figure 4 was generated by a large shallow seismic event (Mw = 7.8) that occurred near the Qinghai–Xinjiang border with China on 14 November 2001, approximately 1800 km from the infrasonic array. In the first part of the signal, coherent wave trains with values of trace velocity greater than 3 km/s are associated with the passage of seismic waves across the infrasonic array. These seismic wave-associated signals are referred to as group A. Group B corresponds to infrasonic waves generated over a large effective source area surrounding the epicentre. Clear variations of azimuth are noted between 180° and 220° while the trace velocity ranges from 0.34 to 0.38 km/s. The main period is around 10 s with a peak-to-peak amplitude of ~2 Pa. A feature of the PMCC algorithm is the ability to make high-resolution continuous detection and propagation parameter measurements. This allows a precise description of the global wave train, which

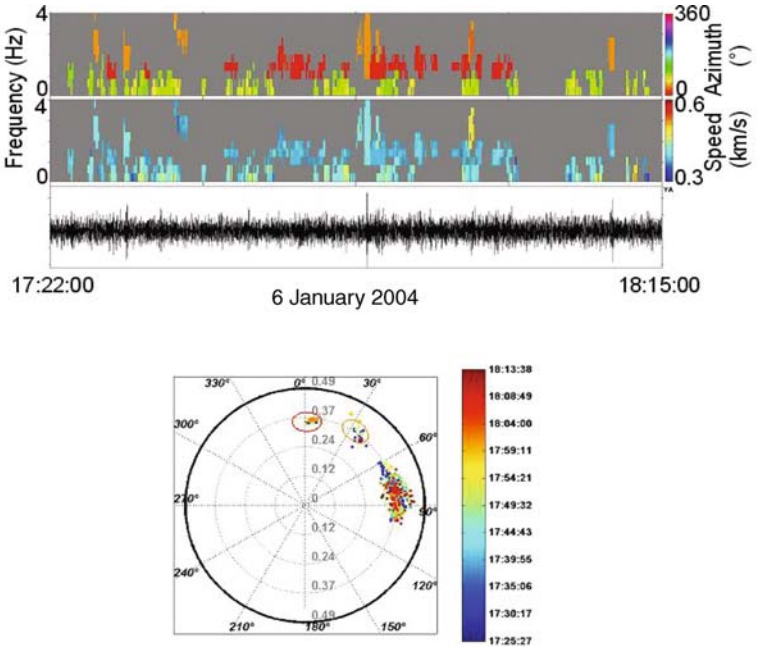


FIGURE 2 Results of the PMCC algorithm applied to typical recordings from I22FR. The data are filtered from 0.5 to 4 Hz in 10 equally spaced frequency bands and displayed as a function of time in the upper plots, where azimuth and trace velocities are indicated by shading. Azimuths are given with respect to North. In the lower plot these same data points are arranged in a polar plot according to azimuth (polar angle) and trace velocity (polar radius) with the arrival time indicated by shading. Pixels corresponding to the Lopevi and Yasur volcanoes form clusters that are circled in the lower (polar) plot. A third cluster is associated with microbaroms

shows complex variations of both azimuth and velocity. The observed variations in azimuth and trace velocity and the long duration of the signal are explained by (1) a local radiation area along the fault rupture and (2) the increase of the effective source region due to the motion of high mountains excited by seismic surface waves propagating away from the fault rupture (LePichon et al., 2003). Figure 5 (top) reconstructs the distant source regions using a celerity of 0.29 km/s. The spatial extent of the radiating zone is about 9° in latitude and 10° in longitude. The source distributions fall into line with the Qilian range, that borders the eastern part of the Qaidam basin and joins the Kunlun range. To the south of Qaidam basin, more scattered source distributions follow the Bayan Har mountains.

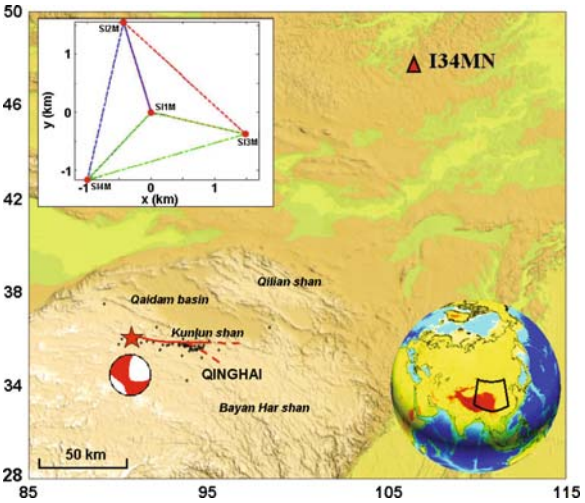


FIGURE 3 Geographical map (GMT) with the epicentre (star) of the Qinghai–Xinjiang earthquake on 14 November 2001 and its focal mechanism (Havard CMT solution) as delimited by the black outlines on the globe. The distribution of aftershocks (black dots, USGS) is consistent with the Kunlun Fault plane (grey line). Infrasound station I34MN (47.8°N–106.4°E, altitude: 1500 m) in Mongolia is composed of four microbarometers (SI1M–SI4M) 1.6 km apart. This station is part of the global infrasonic network of the IMS.

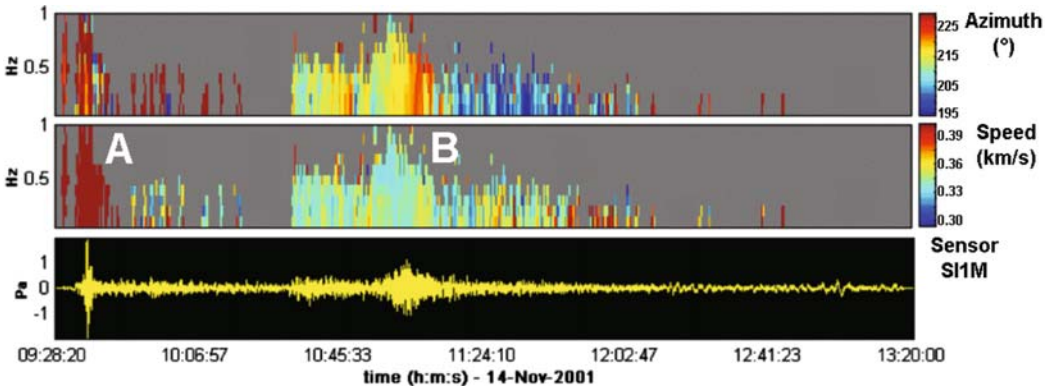


FIGURE 4 Results of PMCC calculation. The results are presented from 0.05 to 1 Hz. The bottom trace shows the atmospheric pressure fluctuations recorded at the central element filtered between 0.05 and 4 Hz. In the first part of the signal, coherent wave trains referred to as group A are related to local seismically coupled air waves. Due to the coupling at the earth–air interface, the horizontal trace velocity of the ground-coupled air waves and the associated seismic waves are identical (greater than 3 km/s). Group B corresponds to the propagation of infrasound generated in the area around the epicentre.

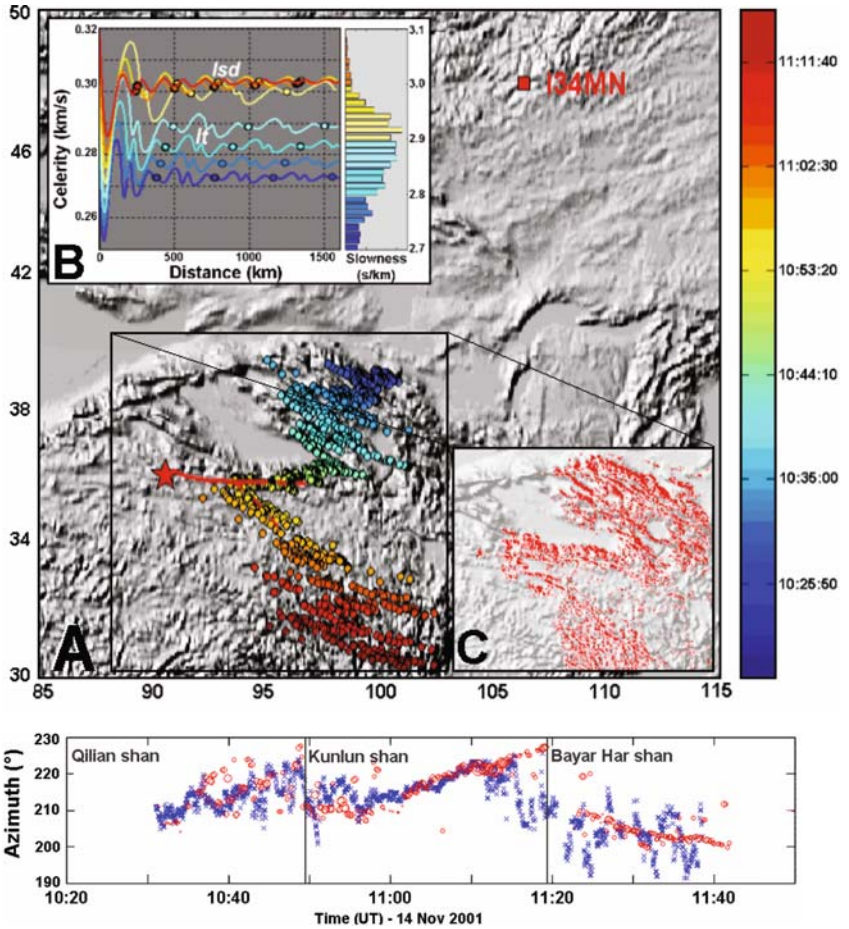


FIGURE 5 Propagation modelling and location of distant source regions of infrasonic waves. ***Top***: (**A**) The dots indicate the sources location according to the detected arrival times (UTC) of the infrasonic waves. Taking into account uncertainties due to the measurements and the propagation modelling, a maximum location error of 20 km is estimated for each dot. (**B**) Predicted celerity models versus slowness and propagation range for a source located at the main shock epicentre. The definition range of the celerity is given by the maximum of the slowness distribution derived from the measured trace velocities (Figure 2). The circles indicate the locations of the ground reception of both ducted stratospheric (*Isd* phases) and thermospheric paths (*It* phases). (**C**) Normalized surface pressure distribution along the Kunlun fault. ***Bottom***: Reconstruction of the azimuth variation at the station. Comparison between the measured (crosses) and the simulated (circles) azimuths. The radius of each circle is proportional to the simulated pressure amplitude radiated by each source element.

Figure 5 (bottom) compares the simulated azimuthal variation at I34MN to the measurements. The time variation of the azimuth is based on the evaluation of the relative contribution to the signal at I34MN from each element of the topography excited by seismic surface waves. Good agreement is found between the observed and the modelled azimuths, especially in the case of infrasonic waves generated by Kunlun Shan belt. The first part of the infrasonic wave signal is generated by seismically induced ground motion within the Qilian Shan range. The apparent scattering of the detection in this case is due to the fact that this mountain range is geographically wide with a succession of ridges which generate simultaneous arrivals with different azimuths. In contrast to the Qilian Shan range, the Kunlun mountains are located along the edge of the Tibetan plateau. Signals generated by this quasi-linear mountain range result in a smooth variation in detected azimuth, and this leads to good agreement between the observed and the simulated arrivals.

These results demonstrate the capabilities of PMCC in the description of complex signals. In particular, since the PMCC method can be used to make detailed continuous observations of individual propagation parameters, the use of the PMCC method in this case provides a detailed description of secondary source phenomena, both in space and time.

References

Cansi, Y., An automatic seismic event processing for detection and location: the PMCC method, *Geophys. Res. Lett.*, 22, 1021–1024, 1995.

Cansi, Y. and Y. Klinger, An automated data processing method for mini-arrays, *CSEM/EMSC European-Mediterranean Seismological Centre,* NewsLetter 11, 1021–1024, 1997.

Capon, J., High resolution frequency wavenumber spectrum analysis, *Proc. IEEE*, 57, 1408–1418, 1969.

Husebye, E. S., Direct measurement of dT/dΔ, *BSSA*, 1969. 59, 2, 717–727.

Le Pichon, A., E. Blanc, D. Drob, S. Lambotte, J. X. Dessa, M. Lardy, P. Bani and S. Vergniolle, Infrasound monitoring of volcanoes to probe high-altitude winds, *J. Geophys. Res.*, 110, D13106, doi:10.1029/2004 JD005587 (2005).

Le Pichon, A., J. Guilbert, M. Vallée, J. X. Dessa and M. Ulziibat, Infrasonic imaging of the Kunlun Mountains during the great 2001 China earthquake, *Geophys. Res. Lett.*, 1029/2003GL017581, 2003.

78

Ray Tracing in an Inhomogeneous Atmosphere with Winds

David J. Brown[1] *and Milton A. Garcés*[2]

[1] *Geoscience Australia, Canberra, Australia, (Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria)*
e-mail: David.Brown@CTBTO.org
[2]*Infrasound Laboratory, University of Hawaii, Manoa, HI, USA,*
e-mail: milton@isla.hawaii.edu

1 • Historical Introduction

A wavefront, or surface of constant phase, is the spatial surface containing sound emitted from a source at the same time (Groves, 1955). Rays are curves normal to these wavefronts.

Although widely used in optics, ray theory was not applied to atmospheric acoustics until the nineteenth century. Delaroche (1816) and Arago (1822), for example, utilize an ad hoc ray-based analysis in their experimental confirmation of enhanced audibility in down-wind sound propagation, as does Tyndall (1867) in his studies of the erratic reception of fog horn signals. The excellent reviews by Wescott and Kushner (1965), and Brown and Hall (1978) discuss much of the earlier work on ray theory as applied to atmospheric acoustics, in particular the important contributions of Stokes (1857), Reynolds (1874, 1876), Rayleigh (1894) and Barton (1901).

The first quantitatively accurate presentation of the acoustic ray equations in a two-dimensional moving medium was presented by Kornhauser (1953), where an error in Rayleigh's earlier work is corrected. Thompson (1972) provides a thorough treatment of the acoustic ray equations in three dimensions and later uses a ray-theoretic treatment to estimate acoustic intensities in a moving medium (Thompson, 1974a, b).

Georges (1972) provides a very comprehensive ray-based analysis of acoustic-gravity wave propagation in an atmosphere, which is applicable both above and below the Lamb wave asymptote. Georges's solution subsequently formed the basis of the very comprehensive *HARPA* computer program that models acoustic-gravity wave propagation in an atmosphere with realistic ambient temperature and wind fields over irregular terrain (Jones et al., 1986). Georges formulated his model equations in Earth-centered spherical coordinates, and hence, these equations apply to acoustic-gravity wave propagation over larger distances. Yan and Yen (1995) subsequently provide the three-dimensional ray-tracing equations in ellipsoidal coordinates, thus accommodating earth flattening, an effect that may be significant for long-range acoustic propagation in the ocean.

Garcés et al., (1998) adopt a second ray-tracing approach to the problem of atmospheric sound propagation. These authors reformulate the tau-*p* model developed originally by Buland and Chapman (1983) for the determination of seismic travel-times and apply it to the problem of atmospheric sound propagation in the presence of winds.

Infrasonic ray tracing has found application in many areas of atmospheric physics. For example, a three-dimensional ray-tracing algorithm (Jones et al., 2004) has been used in an attempt to better define the zones of silence surrounding tornadic supercells and landspouts for the purpose of improving tornado warnings. An infrasonic ray-tracing analysis was used by Brown et al. (2003) to determine that both a cylindrical line source and fragmentation effects are responsible for the acoustic signal generated by the Morávka meteorite and, in addition, to confirm that an unusual horizontally propagating stratospheric mode, which leaked to the receiver, was present. Using the tau-*p* model with enhanced climatological models, Garcés et al. (2002a) show that solar-tidal and geomagnetic effects can significantly influence the properties of infrasound causing azimuthal deviations from the geodesic of up to 10°. Recently, Le Pichon et al. (2005) use ray-tracing techniques as part of an inversion procedure to refine upper atmospheric wind models using acoustic signals from volcanoes as a source.

2 • Atmospheric Variability

The atmospheric sound speed and wind field define the sound propagation path. An early significant development in the area of mid- to upper-level atmospheric physics was the deduction of the generic atmospheric temperature variation with height. Gutenberg (1939) discussed observations, supported strongly by ray analysis, that helped reveal the nature of this temperature variation. Wind fields vary spatially, seasonally and diurnally (Donn and Rind, 1971; Georges and Beasley, 1977), and may dramatically affect the spatial distribution of the sound field.

A new generation of atmospheric models integrates high-resolution meteorological models in the troposphere and stratosphere with climatological models of the mesosphere and lower thermosphere (Drob et al., 2003). These global models generally include diurnal tides, jet streams, and planetary waves, but do not incorporate turbulence and other mesoscale features. Acoustic scattering from turbulent structures in the lower troposphere have been studied in detail (e.g., Brown and Hall, 1978; Ostashev, 1997), although most of these results have been restricted to the boundary layer and acoustic frequencies greater than 10 Hz. Controlled studies on the scattering and refraction of infrasound from large-scale turbulent inhomogeneities and surface irregularities have been limited, in large part due to the absence of a calibrated, stable infrasound source that can penetrate beyond the troposphere. Thus, very little is known about the effects of scattering structures on infrasound propagation, although preliminary studies on the effects of gravity waves have shown that substantial variability may be introduced (Gibson and Norris, 2002). Recent observations of ground truth events and microbarom signals suggest that atmospheric turbulence and surface irregularities may allow trapped infrasonic energy in elevated acoustic waveguides to leak back to the ground through scattering and diffraction (Garcés, et al. 2004).

3 • Geometrical Acoustics

Geometrical acoustics treats sound ray-paths as the extremal curves in Fermats principle of stationary time. Numerous methods have been developed to deduce the ray paths. In practical applications two methods are typically used. The first method defines a pair of Hamiltonian equations that describe the ray-path. Here, the wave number vector is the analogue of the momentum vector, and the angular frequency the analogue of the Hamiltonian function in classical mechanics (Landau and

Lifshitz, 1951; Whitham, 1960, 1961). Although the Hamiltonian treatment provides a full three-dimensional modelling capability, it can be time consuming for numerical applications. This is addressed in the second approach where the tau-p method of Buland and Chapman (1983) is reformulated for sound waves propagating in a stratified atmosphere under the influence of height-dependent winds (Garcés et al., 1998). This approach assumes the vertical component and horizontal gradients of the ambient wind field to be negligible, but provides performance and stability for numerical algorithms in an operational environment.

In geometrical acoustics, various simplifying assumptions are made in order to achieve a tractable problem that is also computationally efficient. As a consequence, the following caveats need to be kept in mind when considering each of the following ray-theoretic treatments:

i. The wave motion is assumed to take place on spatial and temporal scales much shorter than those of the supporting medium.
ii. Areas of convergence (caustics) or rarefaction are regions that may challenge the assumptions of the model, and thus, model predictions may be less accurate in these regions.
iii. Linear wave theory is assumed throughout. In realistic scenarios, where wave amplitude is becoming increasingly nonlinear, the model results may not be valid. This may, for example, be the case for acoustic wave propagation in the upper atmosphere where the exponential decrease in density with height leads to an exponential increase in wave amplitude.

Other treatments such as Full Wave Modelling may need to be employed to handle these situations.

3.1 Asymptotic Ray Theory

In the case of plane harmonic acoustic waves propagating with phase speed c, each dependent variable (pressure, density, etc.) is assumed to satisfy the wave equation

$$\nabla^2 \phi(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 \phi(\mathbf{x}, t)}{\partial t^2} = 0, \tag{1}$$

which is written in the frequency domain as

$$\nabla^2 \phi(\mathbf{x}, \omega) - \frac{(i\omega)^2}{c^2} \phi(\mathbf{x}, \omega) = 0 \tag{2}$$

and is assumed to have solutions of the form

$$\phi(\mathbf{x}, \omega) = \sum_{m=0}^{\infty} \frac{A_m(\mathbf{x})}{(i\omega)^m} \exp[i\omega T(\mathbf{x})] \tag{3}$$

(see, e.g., Jensen et al., 2000). Here, $\mathbf{x}$ is a Cartesian position vector, t is time, ω angular frequency, $A_m(\mathbf{x})$ are amplitude functions, and $T(\mathbf{x})$ is a time function known as the Eikonal that provides the surfaces of constant phase, i.e., the wavefronts, when $T =$ constant. This formulation is made with the intention that the series be asymptotic for large ω but not necessarily convergent as $m \to \infty$ (Burridge, 2002). This is the essence of the high-frequency or WKBJ approximation that is used in many physical systems (see, e.g., Tolstoy and Clay, 1987; Lindsay, 1960). This approximation assumes the wave motion takes place on spatial and temporal scales that are substantially shorter than those that characterize the background and thus certain higher order terms can be ignored.

Substitution of this expression for ϕ into the wave equation and considering only first order terms yields the Eikonal Equation

$$|\nabla T|^2 = \frac{1}{c^2} = s^2, \tag{4}$$

where s is slowness magnitude. The inclusion of the second-order terms in the analysis yields the Transport Equation

$$2\nabla A_0 \cdot \nabla T + A_0 \nabla^2 T = 0. \tag{5}$$

Solution of these two equations provides an estimation of the wavefront location and wave amplitude respectively.

3.2 Hamiltonian Treatment

The Eikonal Equation (4) is in general a nonlinear partial differential equation that provides the travel time of a propagating wave. An alternative approach to directly solving the Eikonal Equation is to use the method of characteristics to transform the problem into a set of coupled ordinary differential equations. This approach may be more tractable. These equations are the Hamiltonian Equations of geometrical optics as presented in Landau and Lifshitz (1951). Whitham (1960, 1961) discusses the application to geometrical acoustics. Whitham's derivation is briefly given here.

If $\mathbf{k}$ is the wave vector for our wave motion and $\mathbf{r}$ the position vector for a point on a surface of constant phase, then the phase function ψ can be written as

$$\psi = \omega(T(\mathbf{r}) - t) + \psi_0 = \mathbf{r} \cdot \mathbf{k} - \omega t + \psi_0, \tag{6}$$

where ψ_0 is an arbitrary phase-value corresponding to zero-time.

Assuming a general functional dispersion relationship of the form $\omega = H(\mathbf{k}, \mathbf{x})$ implies that

$$\frac{\partial \mathbf{k}}{\partial t} + \frac{\partial H}{\partial \mathbf{k}} \cdot (\nabla \mathbf{k})^{\mathrm{T}} = \frac{\partial \mathbf{k}}{\partial t} + \left(\frac{\partial H}{\partial \mathbf{k}} \cdot \nabla \right) \mathbf{k} = -\nabla H, \tag{7}$$

where the superscript T indicates matrix transposition, and we have made use of the identity $\partial\mathbf{k}/\partial t + \nabla\omega = \mathbf{0}$ together with the irrotationality of $\mathbf{k}$. Provided we identify the group velocity $\mathbf{C}$ as $\mathbf{C} = d\mathbf{r}/dt = \partial H/\partial\mathbf{k}$, the left-side of Equation (7) is seen to be the total derivative of the wave number vector, $d\mathbf{k}/dt$, and we have the set of Hamiltonian Equations

$$\frac{d\mathbf{k}}{dt} = -\nabla H \quad \text{and} \quad \frac{d\mathbf{r}}{dt} = -\frac{\partial H}{\partial \mathbf{k}}. \tag{8a–b}$$

We now have six ordinary differential equations that for a specific functional form $\omega = H(\mathbf{k}, \mathbf{x})$ can be integrated forward in time from some initial state using well-defined procedures.

It is also common in the literature to write the dispersion relationship in the slightly different but equivalent form $H(\mathbf{x}, t, \mathbf{k}, \omega) = 0$, with ω as an independent variable. In this case, the governing equations are

$$\frac{d\mathbf{k}}{d\tau} = -\nabla H \quad \text{and} \quad \frac{d\mathbf{r}}{d\tau} = -\frac{\partial H}{\partial \mathbf{k}}, \tag{9a–b}$$

where a new parameter τ is introduced such that

$$\frac{dt}{d\tau} = -\frac{\partial H}{\partial \omega} \quad \text{and} \quad \frac{d\omega}{d\tau} = -\frac{\partial H}{\partial t}. \tag{10a–b}$$

This alternative procedure is discussed more fully in Jones (1969).

It is important to note that Equation (8a–b) determines the extremals to the integral $T = \int_P^Q (1/c)ds$, where ds is the distance differential along a ray path that extends between two fixed points P and Q, and T is total travel time from P to Q. This is a statement of Fermat's Principle of stationarity and can be demonstrated by considering the Lagrangean $L(\mathbf{x}, \dot{\mathbf{x}}, t) = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}/c$ in the variational problem $\delta \int_P^Q L dt = 0$. The required curves are solutions to the vector Euler–Lagrange Equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{x}}}\right) - \nabla L = \mathbf{0}. \tag{11}$$

For configurations where the phase-velocity has no explicit dependence on a coordinate $\Gamma = x, y$, or , Equation (11) implies that for pure acoustic propagation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\Gamma}}\right) = 0, \text{ i.e.,} \frac{\dot{\Gamma}}{c\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \frac{k_\Gamma}{ck} = \text{constant}, \tag{12}$$

which is a result of some importance when considering wave propagation in media that can be considered to be homogeneous in at least one coordinate. The quantity $p = k_\Gamma/(kc)$ is known as the ray parameter, or slowness component in the Γ-direction and remains unchanged along a ray. In configurations where an ambient wind field $\mathbf{u}(\mathbf{x})$ exists, the ray parameter becomes $p = k_\Gamma/(ck + \mathbf{k}\cdot\mathbf{u})$ (Garcés et al., 1998). For acoustic propagation in a stratified atmosphere, the ray parameter can be used to infer information about the wind-speed at altitude, provided the vertical temperature profile is approximately known.

Although providing a legitimate alternative to handling the difficulties encountered when considering the nonlinear partial differential Eikonal Equation, the Hamiltonian Equation (8a–b) is computationally intensive compared with other methods. The coupled ordinary differential equation numerical integration procedure, however, is generally straightforward, and this method does allow for the provision of multi-valued travel time information.

3.2.1 Applications: Georges Model of Acoustic-Gravity Wave Propagation in an Atmosphere with Ambient Winds in Earth-Centered Spherical Coordinates

The most comprehensive ray model developed to date that describes acoustic-gravity wave propagation in a realistic atmosphere is that due to Georges (1972).

Georges takes the Hamiltonian for acoustic-gravity wave propagation, given by $H = \omega^4 - \omega^2 c^2 k^2 - \omega^2\omega_a^2 + c^2 k_h^2 \omega_b^2 = 0$ and determines the Hamiltonian Equations for acoustic-gravity wave propagation in Earth-centred spherical coordinates (r, θ, ϕ) to be

$$\frac{dr}{dt} = c^2 k_r F + u_r,$$

$$\frac{d\theta}{dt} = \frac{1}{r}\left\{c^2 k_\theta \left(\omega^2 - \omega_b^2\right) F/\omega + u_\theta\right\},$$

$$\frac{d\phi}{dt} = \frac{1}{r\sin\theta}\left\{c^2 k_\phi \left(\omega^2 - \omega_b^2\right) F/\omega + u_\phi\right\},$$

$$\frac{dk_r}{dt} = \frac{\omega}{c}\frac{\partial c}{\partial r}\left(\omega_a^2 - c^2k^2\right)F + \left(k_\theta^2 + k_\phi^2\right)\frac{gF}{2\omega}\frac{\partial^2 c^2}{\partial r^2} - \mathbf{k}\cdot\frac{\partial \mathbf{u}}{\partial r} + k_\theta\frac{d\theta}{dt} + k_\phi \sin\theta\frac{d\phi}{dt},$$

$$\frac{dk_\theta}{dt} = \frac{1}{r}\left\{\frac{\omega}{c}\frac{\partial c}{\partial \theta}\left(\omega_a^2 - c^2k^2\right)F + \left(k_\theta^2 + k_\phi^2\right)\frac{gF}{2\omega}\frac{\partial}{\partial\theta}\left(\frac{\partial c^2}{\partial r}\right) - \mathbf{k}\cdot\frac{\partial \mathbf{u}}{\partial \theta} - k_\theta\frac{dr}{dt} + rk_\phi\cos\theta\frac{d\phi}{dt}\right\}, \text{ and}$$

$$\frac{dk_\phi}{dt} = \frac{1}{r\sin\theta}\left\{\frac{\omega}{c}\frac{\partial c}{\partial \phi}\left(\omega_a^2 - c^2k^2\right)F + \left(k_\theta^2 + k_\phi^2\right)\frac{gF}{2\omega}\frac{\partial}{\partial\phi}\left(\frac{\partial c^2}{\partial r}\right)\right.$$
$$\left. - \mathbf{k}\cdot\frac{\partial \mathbf{u}}{\partial\phi} - k_\phi \sin\theta\frac{dr}{dt} - rk_\phi\cos\theta\frac{d\theta}{dt}\right\}. \qquad (13a\text{–}f)$$

In these equations, $\mathbf{k} = (k_r, k_\theta, k_\phi)$ is the wave vector in spherical coordinates, $k_h^2 = k_\theta^2 + k_\phi^2$ is the square of the horizontal wave vector magnitude, c is the local speed of sound, $\omega_a = \gamma g/2c$ is the isothermal form of the acoustic-cutoff frequency with γ the ratio of specific-heats and g the acceleration due to gravity (see, e.g., Gossard and Hooke,[1] 1975), $\omega_b = (\gamma - 1)g^2/c^2 + g\left(\partial c^2/\partial z\right)/c^2$ is the square of the Brunt-Väisälä frequency (see, e.g., Gossard and Hooke 1975), $\mathbf{u} = (u_r, u_\theta, u_\phi)$ is the ambient wind-field in spherical coordinates, and $F = \left(2\omega^2 - \omega_a^2 - c^2k^2\right)^{-1}$. It is also assumed that the wave frequency, Ω, in a fixed frame of reference is given by $\Omega = \omega + \mathbf{k}\cdot\mathbf{u}$.

The solutions to Georges' model equation highlight the fundamentally different morphologies of acoustic and internal gravity waves. In the former case, the dispersion relationship has limiting form $\omega^2 = c^2|\mathbf{k}|^2$ with the phase velocity parallel to the direction of energy flow, and the latter $\omega^2 = \frac{k_h^2\omega_b}{|\mathbf{k}|^2}$ with the group velocity normal to the phase velocity and having vertical component opposite in sign to that of the phase velocity, so that the phase velocity is perpendicular to the direction of energy flow.

These ideas are highlighted in Figure 1, which shows ray-path solutions to Equation (13a–f) for both acoustic waves and internal gravity waves together with the instantaneous wave number vector directions for a source located at 20 km altitude and 0° latitude, and 90° longitude in January. Model equations were solved with a Runge-Kutta four-step algorithm, and ambient wind and temperature conditions were provided by the HWM93 (Hedin et al., 1996) and MSIS (Hedin, 1991) climatological models, respectively.

3.3 Paraxial Ray Tracing

An important addition to Hamiltonian ray tracing theory is the first-order linear process known as the paraxial approximation (Červený and Pšenčik, 1983). Using first-order perturbation

[1]Note that, although Georges' acoustic-gravity ray equations are reproduced in Gossard and Hooke (1975), caution should be used as the equations presented in Gossard and Hooke (1975) contain typographic errors.

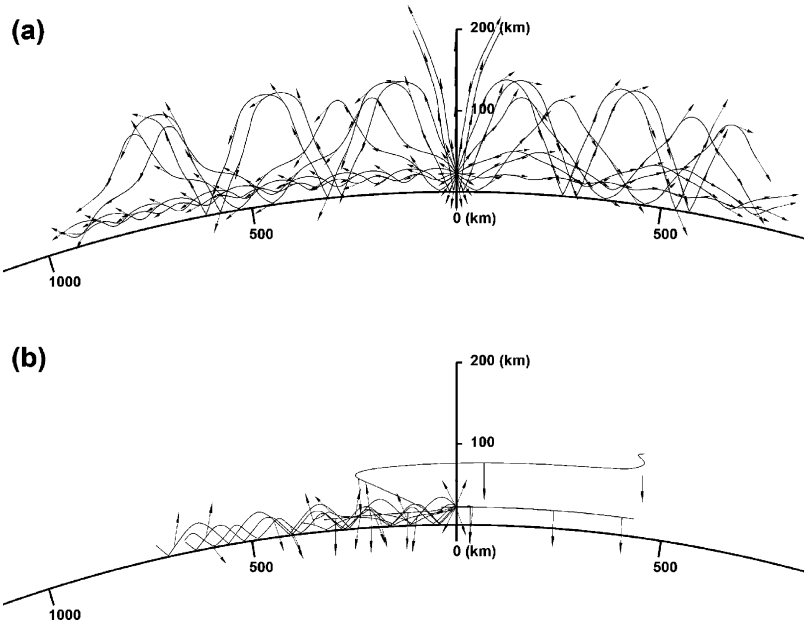


FIGURE 1 Ray-path solutions to Georges acoustic-gravity Hamiltonian ray-tracing equations (10 a–f). Ambient conditions are provided by the HWM93 (Hedin et al., 1996) wind and MSIS (Hedin, 1991) temperature models for a source located at 20 km altitude at 0° latitude, 90° longitude in January. The arrows indicate the direction of the wavenumber vector. (**a**) acoustic waves, period = 5 s, simulation time = 3000 s. (**b**) gravity waves, period 600 s, simulation time = 100000 s.

methods, the behaviour near a reference ray can be inferred. These neighbouring paths or paraxial rays can be used to detect caustics and provide an alternate method of determining amplitude along a ray. The material on Paraxial Ray Tracing presented here is more fully discussed in Virieux and Farra (1991).

The vector $\mathbf{q} = \nabla T$, where T is the eikonal taken from Equation (4), is a slowness vector orthogonal to the wavefront. By defining the Hamiltonian function $H(\mathbf{x}, \mathbf{q}) = \frac{1}{2}\left(\mathbf{q}^2 - s^2(\mathbf{x})\right)$, where s is the acoustic slowness, requires that $H = 0$ along a ray and permits the reformulation of the Eikonal Equation (4) as a pair of Hamiltonian Equations. The method of characteristics yields

$$\dot{\mathbf{x}} = \nabla_{\mathbf{q}} H \quad \text{and} \quad \dot{\mathbf{q}} = -\nabla_{\mathbf{x}} H, \tag{14}$$

where a sampling parameter ζ is defined by requiring $dT = \mathbf{q} \cdot d\mathbf{x} = s^2(\mathbf{x}) d\zeta$ and dots indicate differentiation with respect to ζ.

In addition, $\nabla_{\mathbf{q}}$ and $\nabla_{\mathbf{x}}$ denote the gradients with respect to $\mathbf{q}$ and $\mathbf{x}$, respectively.

By defining a reference ray as the vector $\mathbf{y}_0(\zeta) = (\mathbf{x}_0(\zeta), \mathbf{q}_0(\zeta))$, a paraxial ray is defined by the perturbation

$$\mathbf{y}(\zeta) = \mathbf{y}_0(\zeta) + \delta\mathbf{y}(\zeta) = (\mathbf{x}_0(\zeta) + \delta\mathbf{x}(\zeta), \mathbf{q}_0(\zeta) + \delta\mathbf{q}(\zeta)).$$

The perturbation $\delta\mathbf{y}$ satisfies the paraxial ray-tracing equations

$$\delta\dot{\mathbf{y}} = \mathbf{A}\delta\mathbf{y} \text{ where } \mathbf{A} = \begin{bmatrix} \nabla_{\mathbf{x}}\nabla_{\mathbf{q}}H & \nabla_{\mathbf{q}}\nabla_{\mathbf{q}}H \\ -\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}H & -\nabla_{\mathbf{q}}\nabla_{\mathbf{x}}H \end{bmatrix}, \tag{15}$$

which can be solved using standard techniques.

The new trajectory is also a ray provided that $\delta H = \nabla_{\mathbf{q}}H \cdot \delta\mathbf{q} + \nabla_{\mathbf{x}}H \cdot \delta\mathbf{x} = 0$.

Dessa et al. (2005) demonstrate that the determination of the amplitude evolution along a ray is now possible because the paraxial rays provide a measure of the ray-tube cross-sectional area.

3.4 Tau-*P* Treatment

The discussion on the tau-*p* model presented here is essentially that given in Garcés et al. (1998), with one addition that permits the computation of smooth curves for the azimuth deviation. A caveat in its use is that the model makes the assumption that both the vertical component and the horizontal gradient of the atmospheric wind velocity are negligible and are therefore set to zero. The primary advantage of this approach is stability and speed of computation, which is important when performing multiple forward modelling permutations during a source location search.

Straightforward manipulation of the Hamiltonian Equations (8a–b) show that the position of the wavefront $\mathbf{r}$ and its normal $\hat{\mathbf{n}}$ as a function of time are given by

$$\frac{d\mathbf{r}}{dt} = c\hat{\mathbf{n}} + \mathbf{u} = \mathbf{C}, \quad \text{and} \tag{16}$$

$$\frac{d\hat{\mathbf{n}}}{dt} = (\hat{\mathbf{n}} \cdot \nabla\mathbf{c})\,\hat{\mathbf{n}} - \nabla\mathbf{c}, \tag{17}$$

where $\mathbf{c} = (c + \hat{\mathbf{n}} \cdot \mathbf{u})\,\hat{\mathbf{n}}$ is the phase velocity and $\mathbf{C} = c\hat{\mathbf{n}} + \mathbf{u}$ the group velocity. Without loss of generality, we can align the ray normal along the x–z plane in the Cartesian system (x–z) and consider only the ray propagating along that direction. This is equivalent to considering propagation only along a given azimuth.

Assuming the form of the wind profile to be $\mathbf{u} = (u(z), v(z), 0)$ and the ray parameter to be

$$p = \frac{\sin\theta}{c}\left(1 + \frac{u\sin\theta}{c}\right)^{-1}, \tag{18}$$

where θ is the launch angle, Garcés et al. (1998) uses Equations (16) and (17) to show that the travel time, T, the range along the ray direction, X, and transverse offset, Y, of a propagating wave over a ray can be written as

$$T = \int_C s^2 \left(s^2 - \frac{p^2}{(1-pu)^2}\right)^{-\frac{1}{2}} dz, \tag{19}$$

$$X = \int_C \left(\frac{p}{(1-pu)} + s^2 u\right)\left(s^2 - \frac{p^2}{(1-pu)^2}\right)^{-\frac{1}{2}} dz, \text{ and} \tag{20}$$

$$Y = \int_C (s^2 v)\left(s^2 - \frac{p^2}{(1-pu)^2}\right)^{-\frac{1}{2}} dz, \tag{21}$$

where $s = c^{-1}$, and where the integral is taken over the vertical distance travelled by the ray path. Similarly, a quantity, W, can be defined as

$$W = \int_C v\left[\frac{p}{(1-pu)} + us^2\right]\left[s^2 - \frac{p^2}{(1-pu)^2}\right]^{-\frac{1}{2}} dz, \tag{22}$$

which is related to X but is used primarily to facilitate the interpolation of the transverse offset and the azimuth deviation.

It is easy to verify that the above equations reduce to their standard forms when the wind velocity vanishes.

As in the windless case, we can use the Legendre transformation (Buland and Chapman, 1983) to define the intercept time, τ, as

$$\tau = T - pX, \tag{23}$$

then

$$\tau = \int_C (1-pu)\left[s^2 - \frac{p^2}{(1-pu)^2}\right]^{\frac{1}{2}} dz. \tag{24}$$

The tau function, τ, emerges from the variational method of classical mechanics in the specific case where we seek to minimize the total travel time T along the wave propagation

path. By applying the Legendre transform (Goldstein, 1980), we obtain τ as our Hamiltonian, with the ray parameter, p, acting as a generalized momentum. The tau function is a monotonic function of the ray parameter, and it does not exhibit the square root singularities that X, Y and T show in Equations (19–21). Thus, for a specific travel time branch, we can construct an interpolation function for tau up to its turning point, and transform tau into a continuous function of ray parameter (Buland and Chapman, 1983). Then, the range for a specific ray parameter (or launch angle), p_k, can be obtained from

$$\frac{d\tau}{dp}\Big|_{p_k} = -X_k \tag{25}$$

and the travel time for a specific arrival at a given range from

$$T_k = \tau(p_k) + p_k X_k. \tag{26}$$

Similarly, we can define a function

$$\sigma = Y - pW = \int_C v(1-pu)\left[s^2 - \frac{p^2}{(1-pu)^2}\right]^{\frac{1}{2}} dz, \tag{27}$$

where

$$\left.\frac{d\sigma}{dp}\right|_{p_k} = -W_k \tag{28}$$

and

$$Y_k = \sigma_k + p_k W_k. \tag{29}$$

Both the τ and the σ functions are piecewise continuous functions whose discontinuities are defined by the slowness values at which the radical vanishes.

3.4.1 Applications

To apply Equations (18–29), we assume constant velocity layers and discretize the ray parameter, p. The discrete versions of Equations (24) and (27) are

$$\tau_{ij} = (1-p_j u_i)\left[s_i^2 - \frac{p_j^2}{(1-p_j u_i)^2}\right]^{\frac{1}{2}} \Delta z_i, \quad \text{and} \tag{30a–b}$$

$$\sigma_{ij} = v_i(1-p_j u_i)\left[s_i^2 - \frac{p_j^2}{(1-p_j u_i)^2}\right]^{\frac{1}{2}} \Delta z_i$$

where a subscript i denotes the index of the layer at elevation z_i, $\Delta z_i = z_{i+1} - z_i$, and j is the index of the discretized ray parameter, $p = \{p_j\}$.

Assuming a source at an elevation z_0, for each ray parameter we can identify a launch angle, θ_0, measured from the vertical, and given by Equation (18):

$$\theta_o = \sin^{-1}\left[pc_0\,(1 - pu_0)^{-1}\right], \tag{31}$$

where c_0 is the sound speed and u_0 the wind velocity along the ray path at the source. The discretized equations substitute an integral over a ray path by the sum of the layers over the given ray path. From the τ and σ functions, the travel time, range, and transverse offset can be readily computed as piecewise continuous, smooth functions of ray parameter. We can also estimate the azimuthal deviation for each ray from

$$\tan \xi_j = \frac{\sum\limits_i Y_{ij}}{\sum\limits_i X_{ij}}. \tag{32}$$

For each ray at each layer, real solutions can only be obtained if

$$p_j \leq \frac{1}{c_i + u_i}, \tag{33}$$

with the equality defining the turning point of a ray. Thus, we can define the turning slowness, p_i, at each layer by $p_i^{-1} = c_i + u_i$. The inverse of the turning slowness is often referred to as the effective sound speed of the atmosphere along the direction of propagation. Figure 2 shows a sound speed and wind speed profile evaluated for an infrasound station located at 47.8°N, 106.4°E at 1800 UT on 10 December 2003, and Figure 3 shows the turning slowness for rays arriving at the station with an azimuth

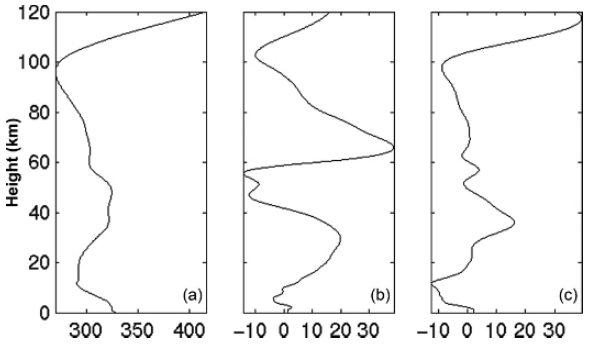


FIGURE 2 Sound speed, zonal (positive to the east), and meridional (positive to the north) wind speeds, in meter per second, as a function of atmospheric height (km) recorded at 47.8°N, 106.4°E on 10 December 2003 at 1800 UT.

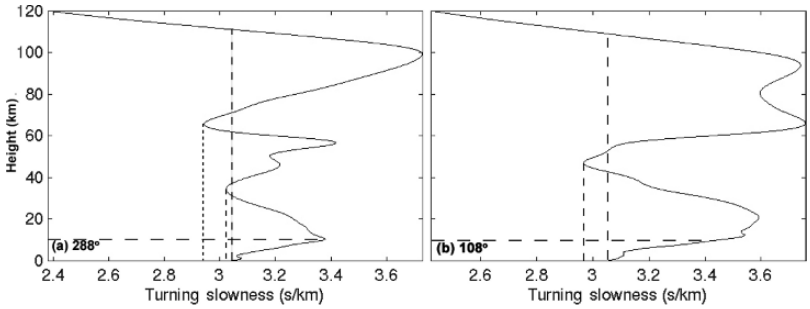


FIGURE 3 Turning slowness as a function of height for arrival azimuth 288° and 108° and the atmospheric conditions shown in Figure 2.

of 288° and 108°. The vertical dashed line furthest to the right denotes the limit set by the effective sound speed at the ground. In principle, ray parameters with values greater than the slowness at the surface cannot reach the ground and would not be observed. Thus, we can define all possible acoustic waveguide boundaries from the turning slowness without launching any rays.

The tau-*p* algorithm would yield the results shown in Figure 4 when tracing multiple rays through the atmospheric conditions shown in Figure 2, and arriving at the receiver from an azimuth of 288°. The celerity, in meter per second, is defined as the ratio of the range over the travel time and may be viewed as an effective propagation speed. Thus far, we have encountered the acoustic phase velocity, group velocity, effective sound speed, trace velocity, and the celerity. The first three quantities depend on the sound speed and wind speed at a specific region of the atmosphere. The trace velocity, which is the inverse of the slowness, is preserved for each ray and may be measured directly by observing the apparent horizontal speed of propagation of a signal across an array. The celerity, however, is an average propagation speed over the complete propagation path from source to receiver and may be also computed directly for an observed infrasonic arrival with a known source–receiver configuration. Although it is inviting to use the trace velocity for identifying a given arrival, measurement and calibration uncertainties make this procedure rather inaccurate for sound arriving near the horizontal plane from distant sources. Celerity estimates from various infrasonic events with known locations suggest that (1) our knowledge of the atmosphere is presently insufficient to reliably predict all infrasonic arrivals at a station under all conditions and (2) diffraction and/or scattering can feed acoustic energy to waveguides that are elevated above the ground, and

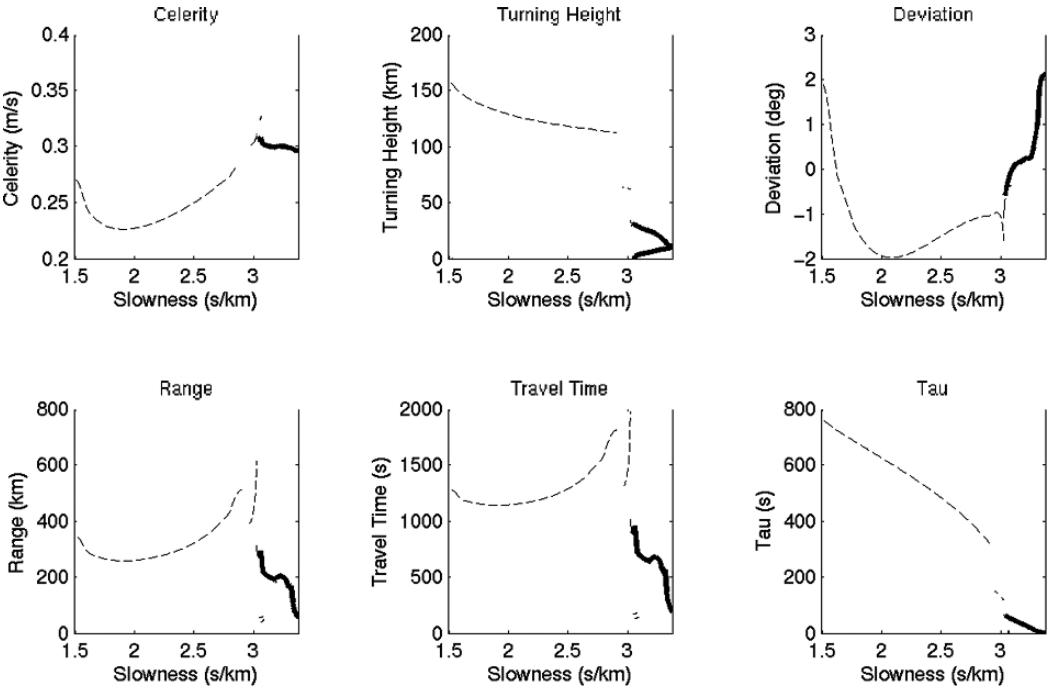


FIGURE 4 Celerity, turning height, azimuth deviation, range, travel-time, and tau as a function of the slowness of each ray for an arrival azimuth of 288° and the atmospheric conditions shown in Figure 2.

these elevated waveguides may also leak sound back to the ground. For the purposes of our discussion, let us assume that our receiver is located ~10 km above the ground so that we may consider all infrasonic arrivals at a given region.

The turning height of a ray defines the upper and the lower waveguide boundaries. The turning height in Figure 5 shows only the waveguides that either reach the ground or are close to the ground. This panel contains a subset of information from Figure 4, which shows that multiple elevated waveguides exist above 30 km altitude. We concentrate on waveguides with lower boundaries below 10 km, because energy from these waveguides may leak back to the ground. The azimuth deviation shows the angular deflection along the direction of propagation or the correction that has to be made at the station so that the observed arrival azimuth points back to the source location. A positive azimuthal deviation corresponds to a correction to the right of the observed backazimuth.

The interpolated range and travel times, all derived from the piecewise continuous tau function, can be used to determine

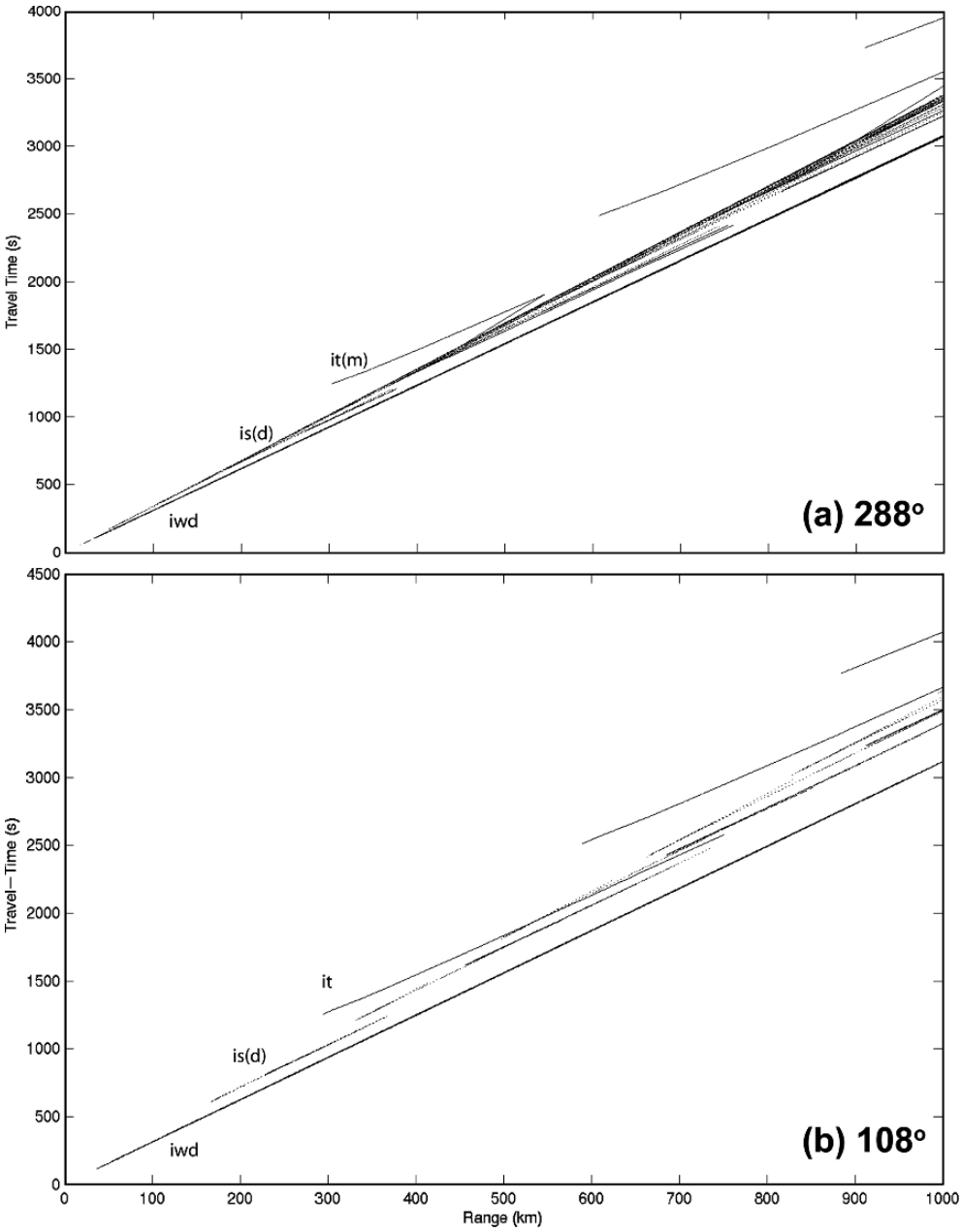


FIGURE 5 Travel–time curves for infrasound arriving with an azimuth of (**a**) 288° and (**b**) 108° assuming the atmospheric conditions shown in Figure 2.

travel–time curves where the slope of the curve is equal to the ray-parameter. Travel–time curves of this sort are often used in seismic source location applications (e.g. Schweitzer, 2001; Ballard, 2002). However, due to the highly variable nature of the effective sound-speed, infrasonic travel–time curves differ substantially from their seismic counterpart. This adds substantial dynamic complexity to the infrasonic source location problem.

The predicted arrivals at a ground-based station may be classified according to the atmospheric waveguide in which infrasonic energy is ducted. A continuously evolving arrival classification scheme (Brown, 2002; Garcés, 2002b), is given in Table 1.

With the notable exception of it, im, is, and iw, which may arrive straight from the source with very high apparent horizontal phase velocities, most arrivals observed at infrasound arrays from large distances have a trace velocity close to that of the local sound speed. Distant arrivals with high trace velocities and low celerity values are often difficult to observe because they are more heavily attenuated and reach the station in the coda of a complex signal consisting of multiple arrivals. The nomenclature in Table 1 is used in Figure 5, which shows travel-time curves for arrival azimuths 288 and 108° assuming the atmospheric profile displayed in Figure 2. Although discrete behaviour in the distribution of ground-based audibility is predicted to occur on shorter

TABLE 1 Arrival nomenclature for long-range infrasonic propagation based on waveguide boundaries

Phase ID	Waveguide boundaries (upper and lower)	Typical celerity of first arrival, m/s
iw	Troposphere and the ground. May not be stable.	310–340
iwd	Elevated waveguide in the troposphere, leaking to the ground.	310–340
is	Upper stratosphere and the ground.	290–330
isd	Elevated waveguide between stratosphere and the troposphere, leaking to the ground. May have higher frequency.	290–330
im	Mesosphere and the ground.	300–320
imd	Elevated waveguide between the mesosphere and the troposphere, leaking to the ground.	300–320
it	Lower thermosphere and the ground.	280–300
itd	Elevated waveguide between the lower thermosphere and the troposphere, leaking to the ground.	280–300
it, im, is, iw	Direct arrival from the source to the receiver – not a guided arrival. May have high apparent phase velocity	N/A

spatial and temporal distances from a source, resulting in a clear succession of distinct arrivals, or in other cases 'zones of silence', the possible arrival of signals from a myriad of potential propagation paths at distance can change the nature of the received signal substantially. Zones of silence disappear, with the resultant signal stretched in time, being the overlapping sum of contributions from a number of different arrivals.

The December 10 profile was chosen because it illustrates some of the complexities of infrasonic wave propagation and defies some of the expected results. For example, stratospheric arrivals are only expected along the downwind stratospheric wind direction. For the profiles in Figure 2, the zonal and meridional winds reverse direction between 40 and 50 km and conspire with a high sound-speed in that region to produce predicted stratospheric arrivals along opposing azimuths. Figure 6 shows five rays with prescribed slowness values of $p1 = 2.8300$ s/km, $p2 = 3.0221$ s/km, $p3 = 3.0225$ s/km, $p4 = 3.0592$ s/km, and $p5 = 3.0614$ s/km, received along the azimuths of 108 and 288°. Along the arrival azimuth of 288°, the predicted celerities (arrivals) are 277 m/s (it), 315 m/s (im), 311 m/s (is), 312 m/s (isd), and

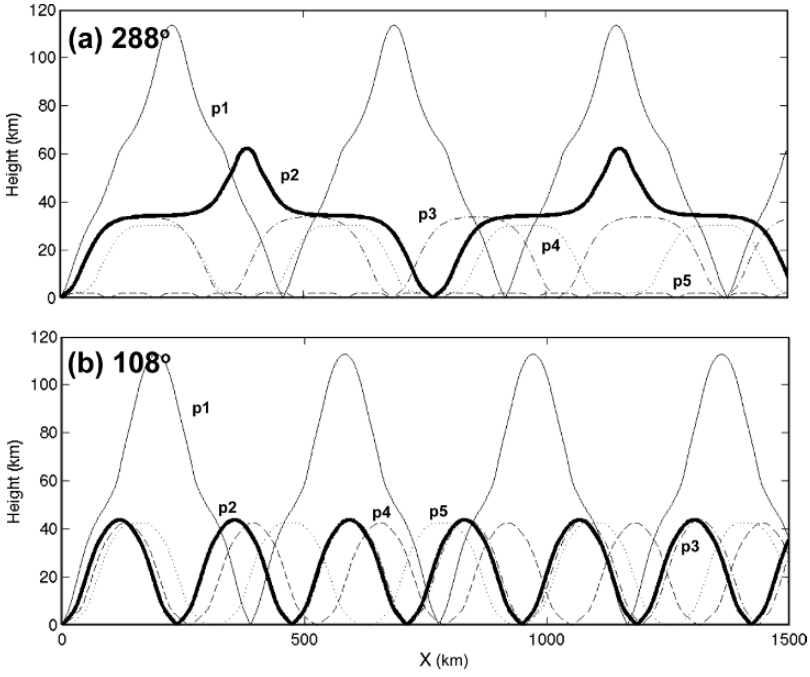


FIGURE 6 Propagation paths for five rays arriving with an azimuth of (**a**) 288° and (**b**) 108°.

326 m/s (iwd), respectively. However, along the arrival azimuth of 108°, for exactly the same rays, the predicted celerities (arrivals) are 258 m/s (it), 284 m/s (is), 284 m/s (is), 293 m/s (isd), 288 m/s (isd). Of particular interest is the predicted return from the mesosphere. Recent infrasonic observations suggest that winds in the mesosphere may be underestimated (Garcés, et al. 2004). Figure 6a also shows how nearly adjacent rays may take very different paths. For example, $p2$ and $p3$ may conceivably alternate paths over long ranges, leading to a hybrid propagation path with a celerity that would average over both rays. It would be relatively easy to slightly re-route a ray through surface roughness, topography, weather, and the boundary layer, which all rays have to pass through. For these reasons, tropospheric waveguides are bound to be unstable for long ranges. In addition, we know that the atmosphere is continuously changing and possesses inhomogeneities at a wide range of temporal and spatial scales (Brown and Hall, 1978). Propagating (as well as breaking) gravity waves could also induce substantial variability in the lower, middle, and upper atmosphere. Thus, during long-range propagation a given ray may propagate though more than one waveguide channel, further complicating the identification of an arrival.

The December 10 case study shows that

i. small changes in the ray slowness may yield completely different propagation paths and the same slowness value may produce different arrivals when encountering different propagation environments. Rays are particularly vulnerable when they are propagating close to the horizontal and may be deflected from the expected course by atmospheric inhomogeneity and surface variability
ii. stratospheric and tropospheric arrivals may arrive from opposing azimuths and, under certain conditions, mesospheric arrivals may be expected.

3.5 Other Ray Tracing Methods

Other ray-tracing methods have become popular in the geosciences and may find application in acoustics. Leidenfrost (1999) provide a comparison of some of these techniques.

Shooting and bending ray-tracing methods (Julian and Gubbins, 1977) are concerned with determining eigenrays, that is, those rays that connect given start-and end-point locations. Starting at a specified start-point location, the shooting method varies the launch angles of a ray in an iterative fashion until it passes through the specified end point location. The bending method chooses an initial curve that passes through both start and end point locations and deforms it over a series of iterations until it satisfies Fermat's principle of stationary time.

Wavefront construction (WC) techniques (Vinje et al., 1993; Lambaré et al., 1996) have also been employed to determine travel-time information. With this scheme, the location of the wavefront at time $T + \delta T$ is prognosed from the location of the wavefront at time T by numerically integrating the Hamiltonian equations for a large number of rays from time T to $\mathrm{T} + \delta T$. The new wavefront location is determined collectively by the new end point of each ray . In the situation where diverging rays cause the wavefront to be poorly sampled, new rays are introduced by interpolation. WC techniques are often favored because of their robustness and accuracy in complex smooth media together with their ability to provide multi-valued travel time information, but they can be computationally slow.

Finite difference (FD) methods have also been applied to the Eikonal Equation (Vidale, 1988). FD methods, however, are often challenged by regions where the evolving surface focuses or undergoes rarefaction. A significant development in the FD solution of the Eikonal Equation is the application of the "Entropy-Satisfying" upwind-differencing scheme (see, e.g., Sethian and Popovici, 1999) defined by

$$\begin{bmatrix} \max(D_a^{-x}T, -D_b^{+x}T, 0)^2 + \\ \max(D_c^{-y}T, -D_d^{+y}T, 0)^2 + \\ \max(D_e^{-z}T, -D_f^{+z}T, 0)^2 \end{bmatrix}_{ijk}^{\frac{1}{2}} = s_{ijk}, \tag{34}$$

where a,b,c,d,e,f are integers that define the order of the finite-difference operators, which may be mixed in Equation (34). For example $D_1^{-x}T_i = \frac{T_i - T_{i-1}}{\delta x}$ and $D_2^{-x}T_i = \frac{3T_i - 4T_{i-1} + T_{i-2}}{2\delta x}$. These Fast Marching Methods (FMM) (Sethian and Popovici, 1999; Rawlinson and Sambridge, 2004) are unconditionally stable for the first order scheme and will advance the wavefront in $O(N \log N)$ steps for a mesh containing N points in media of arbitrary complexity. FMM methods will, however, only compute the first arrival.

4 • Concluding Remarks

Ray theory is a high-frequency approximation that yields valuable insight into the propagation of sound in a temporally varying inhomogeneous, moving atmosphere. The geometrical acoustics solution is not above reproach, however. Barry (1963) notes that the justification for its use is that there is no simple alternative that yields comparable insight into the qualitative behaviour of acoustic energy. However, due to their relative simplicity, ray methods have a place in operational environments where

computational speed is essential and where environmental and measurement uncertainties limit the attainable accuracy (Jensen et al., 2000).

The few simple examples provided here indicate that the propagation of acoustic energy in the atmosphere can be accurately modelled, and hitherto complex observations resolved, if ray theory is sensibly applied. The ad hoc analysis in Tyndall's (Tyndall, 1867) interpretation of erratic fog-horn signals noted in the Historical Introduction has been refined and accurately quantified, but otherwise follows the same principles used by modern algorithms.

Acknowledgments M. Garcés thanks P. Caron and C. Hetzer for their analyses.

References

Arago, D.F.J. (1822), Resultats des experiences faites en 1822, par ordre du Bureau des Longitudes, por la determination de la vitesse du son dans l'atmosphere. *Ann. Chim. Phys.*, **20**, 210.

Ballard, S. (2002), Sesmic event location using Levenberg-Marquardt least squares inversion. *Sandia National Laboratories Report*, SAND2002-3083.

Barry, G. (1963), Ray tracings of acoustic waves in the upper atmosphere. *J. Atmos. Terr. Phys.*, **25**, 621–629.

Barton, E.H. (1901), On the refraction of sound by wind. *Phil. Mag.*, **1**, 159–165.

Brown, E.H. and Hall, F.F. (1978), Advances in atmospheric acoustics. *Rev. Geophys. Space Phys.*, **16**, 47–110.

Brown, D.J., Katz, C.N., LeBras, R., Flanagan, M.P., Wang, J., and Gault, A.K. (2002), Infrasonic signal detection and source location at the prototype international data centre. *Pure. Appl. Geophs.*, **159**, 1081–1125.

Buland, R. and Chapman, C. (1983), The computation of seismic travel times. *Bull. Seismol. Soc. Am.*, **73**, 1271–1302.

Burridge, R. (2002), *Ray Theory for the Elastic Wave Equation*. Institute for Mathematics and its Applications, University of Minnesota, Short Course.

Červený, V. and Pšenčik, I. (1983), Gaussian beams and paraxial ray approximation in three dimensional elastic inhomogeneous media. *J. Geophys.*, **53**, 1–15.

Delaroche, F. (1816), Sur l'influence que le vent exerce dans la propagation du son, sous le rapport se son intensite. *Ann. Chim.*, **1**, 176

Dessa, J.X., Virieux, J., and Lambotte, S. (2005), Infrasound modelling in a spherical heterogeneous atmosphere. *Geophys. Res. Lett.*, **32**, L12808.

Donn, W.L. and Rind, D. (1971), Natural infrasound as an atmospheric probe. *J. R. Astr. Soc.*, **26**, 111.

Drob, D., Picone, M., and Garcés, M. (2003), The global morphology of infrasound path partitioning. *J. Geophys. Res*., **108**, 4680.

Garcés, M., Drob, D., and Picone, M. (2002a), A theoretical study of the effect of geomagnetic fluctuations and solar tides on the propagation of infrasonic waves in the atmosphere. *Geophys. J. Int.l*, **148**, 77–87.

Garcés, M.A., Hansen, R.A., and Lindquist, K.G. (1998), Traveltimes for infrasonic waves propagating in a stratified atmosphere. *Geophys. J. Int.*, **135**, 255–263.

Garcés, M., Hetzer, C., Lindquist, K., and Drob, D. (2002b), Source location algorithm for infrasonic monitoring. 24th Annual DTRA/NNSA Seismic Research Review, Ponte Vedra, FL, 17–19 September 2002.

Garcés, M., Willis, M., Hetzer, C., Le Pichon, A., and Drob, D. (2004), On using ocean swells for continuous infrasonic measurements of winds and temperature in the lower, middle, and upper atmosphere. *Geophys. Res. Lett*., **31,** L19304, doi:10.1029/2004 GL020696.

Georges, T.M. (1972), 3D Ray tracing for acoustic-gravity waves. *Proceedings Conference on Effects of Acoustic Gravity Waves on Electromagnetic Wave Propagation*, AGARD Proc. **115,** 2–1 to 2–8.

Georges, T.M. and Beasley, W.H. (1977), Refraction of infrasound by upper atmospheric winds. *J. Acoust. Soc. Am.*, **61**, 28–34.

Gibson, R. and Norris, D (2002), Infrasound propagation modeling: near-real-time environmental characterization and variability studies. *Proceedings of the Infrasound Technology Workshop*, De Bilt, Netherlands, October 28–31, 2002.

Goldstein, H. (1980), *Classical Mechanics, Second Edition*. Addison-Wesley Publishing Company.

Gossard, E.E. and Hooke, W.H. (1975), *Waves in the Atmosphere: Atmospheric Infrasound and Gravity Waves – Their Generation and Propagation*. Elsevier Scientific Publishing Company.

Groves, G.V. (1955), Geometrical theory of sound propagation in the atmosphere. *J. Atmos. Terr. Phys*., **7**, 113–127.

Gutenberg, B. (1939), The velocity of sound waves and the temperature in the Stratosphere in Southern California. *Bull. Am. Met. Soc*., **20**, 192–201.

Hedin, A.E. (1991), Extension of the MSIS thermosphere model in the middle and lower atmosphere. *J. Geophys. Res*., **96**, 1159–1172.

Hedin, A.E., Fleming, E.L., Manson, A.H., Schmidlin, F.J., Avery, S.K., Clark, R.R., Franke, S.J., Fraser, G.J., Tsuda, T., Vial, F., and Vincent, R.A. (1996), Empirical wind model for the upper, middle and lower atmosphere. *J. Atmos. Terr. Phys*., **58**, 1421–1444

Jensen, F.B., Kuperman, W.A., Porter, M.B., and Schmidt, H. (2000), *Computational Ocean Acoustics*. Springer-Verlag, New York.

Jones, W.L. (1969), Ray tracing for internal gravity waves. *J. Geophys. Res*., **74**, 2028–2033.

Jones, R.M., Georges, T.M., and Riley, J.P. (1986), *HARPA: A versatile three-dimensional Hamiltonian ray tracing program for acoustic waves in the atmosphere above irregular terrain*. NOAA Special Report, 410pp.

Jones, R.M., Gu, E.S., and Bedard, A.J. (2004), *Infrasonic Atmospheric Propagation Studies Using a 3-D Ray Trace Model* NOAA, Environmental Technology Laboratory, Boulder, Colorado. Special Report, 9. pp.

Julian, B.R. and Gubbins, D. (1977), Three-dimension seismic ray tracing. *J. Geophys.*, **43**, 95–113.

Kornhauser, E.T. (1953), Ray theory for moving fluids, *J. Acoust. Soc. Am.*, **25**, 945–949.

Lambaré, G., Lucio, P.S., and Hanyga, A. (1996), Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of a ray field. *Geophys. J. Int.*, **125**, 584–598.

Landau, L.D. and Lifshitz, E.M. (1951), *The Classical Theory of Fields*. Addison-Wesley, Cambridge.

Leidenfrost, A., Ettrich, N., Gajewski, D., and Kosloff, D. (1999), Comparison of six different methods for calculating traveltimes. *Geophys. Prospect.*, **47**, 269–297.

Le Pichon, A. Blanc, E., and Drob, D. (2005), Probing high-altitude winds using infrasound. *J. Geophys. Res.*, **110**, D20104, doi:10.1029/2005JD006020.

Lindsay, R.B. (1960), *Mechanical Radiation*. McGraw-Hill, New York.

Ostashev, V. (1997), *Acoustics in Moving Inhomogeneous Moving Media*, E&FN Spon, imprinted by Thompson Science and Professional, London, UK.

Rawlinson, N. and Sambridge, S. (2004), Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophys. J. Int.*, **156**, 631–647.

Rayleigh (Strutt, J.W.) (1894), *The Theory of Sound*. Macmillan, New York.

Reynolds, O. (1874), On the refraction of sound by the atmosphere. *Proc. R. Soc. Lond.*, **22**, 531–548.

Reynolds, O. (1876), On the refraction of sound by the atmosphere. *Philos. Trans. R. Soc.*, **166** (Part 1), 315–324.

Schweitzer, J. (2001), HYPOSAT – An enhanced routine to locate seismic events. *Pure. Appl. Geophys.*, **158**, 277–289.

Sethian, J.A. and Popovici, A.M. (1999), 3-D traveltime computation using the fast marching method. *Geophysics*, **64**, 516–523.

Stokes, G.G. (1857), On the effect of wind on the intensity of sound, *Report, Brit. Assoc. Dublin*. (Reprinted in *Mathematical and Physical Papers of G. G. Stokes.*, **4,** pp. 110–111, Cambridge University Press, New York, 1904).

Thompson, R.J. (1972), Ray theory for an inhomogeneous moving medium. *J. Acoust. Soc. Am.*, **51**, 1675–1682.

Thompson, R.J. (1974a), Ray-acoustic intensity in a moving medium. 1. *J. Acoust. Soc. Am.*, **55**, 729–732.

Thompson, R.J. (1974b), Ray-acoustic intensity in a moving medium. 2. *J. Acoust. Soc. Am.*, **55**, 733–737.

Tolstoy, I. and Clay, C.S. (1987), *Underwater Acoustics: Theory and Experiment in Underwater Sound*. American Institute of Physics, New York

Tyndall, J., (1867), *On Sound*, D. Appleton, NT, NY.

Vidale, J.E. (1988), Finite-difference calculations of traveltimes. *Bull. Seism. Soc. Am.*, **78**, 2062–2076.

Vinje, V., Iversen, E., and Gjøystdal, H. (1993), Traveltime and amplitude estimation using wavefront construction. *Geophysics*, **58**, 1157–1166

Virieux, J. and Farra, V. (1991), Ray Tracing in 3-D complex isotropic media: An analysis of the problem. *Geophysics*, **56**, 2057–2069.

Wescott, J.W. and Kushner, S.S. (1965), *Propagation of Sound in Air: A Bibliography with Abstracts*. The University of Michigan, Ann Arbor.

Whitham, G.B. (1960), A note on group velocity. *J. Fluid. Mech.*, **9**, 347–352.

Whitham, G.B. (1961), Group velocity and energy propagation for three-dimensional waves. *Comm. Pure. Appl. Math.*, **14**, 675–691.

Yan, J. and Yen, K. (1995), A derivation of three-dimensional ray equations in ellipsoidal coordinates. *J. Acoust. Soc. Am.*, **97**(3), 1538–1544.

79

Infrasonic Data Processing at the Provisional Technical Secretariat of the Comprehensive Nuclear-Test-Ban Organization

Nicolas Brachet[1], *Petr Firbas*[2], *Thomas Hoffmann*[1], *Paola Campus*[1], *Stefka Stefanova*[1]

[1]*Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria*
[2]*Provisional Technical Secretariat, CTBTO, Vienna International Centre, Vienna, Austria (International Atomic Energy Agency, Vienna International Centre, Vienna, Austria)*

On 10 September 1996, the United Nations General Assembly in New York adopted the Comprehensive Nuclear-Test-Ban Treaty (CTBT) to prohibit nuclear explosions in any environment – underground, in the atmosphere or in water. A global verification regime has been set up to support the CTBT. The verification

regime relies on the International Monitoring System (IMS) consisting of a global network of 321 monitoring stations and 16 laboratories employing four technologies – seismic, infrasound, hydroacoustic and radionuclide. IMS data are collected, processed and archived at the International Data Centre (IDC) in Vienna, Austria. The IDC distributes both the IMS raw data and the IDC final products to the States Signatories to the CTBT.

In its final configuration, the IMS network will comprise 60 infrasound stations [1]. The location of the stations is defined in Annex 1 to the Protocol of the Treaty. Figure 1 shows the global distribution of the IMS infrasound network. These stations are arrays of microbarometers, sensors that are sensitive to acoustic pressure variations in the atmosphere. The array configurations include four to eight elements with apertures between 1 and 3 km (see Figure 2).

The IMS infrasound stations record low frequency acoustic waves propagating in the atmosphere. Natural sources – such as meteorological phenomena, explosive volcanic eruptions, meteorites, aurora – or man-made sources – such as atmospheric and surface explosions, industrial activity, rocket launches, aircraft sonic booms – are known to generate signals that may be detected at the IDC. Generally, the sensitivity of an infrasound station is limited by small-scale wind-induced turbulence in the vicinity of the infrasound sensors. Efficient noise reduction can be achieved by connecting the infrasound sensor to a multiport pipe filter. As shown in Figure 2, such a noise-reducing pipe filter averages over the micropressure fluctuations arriving at numerous inlet ports within a relatively large surface area.

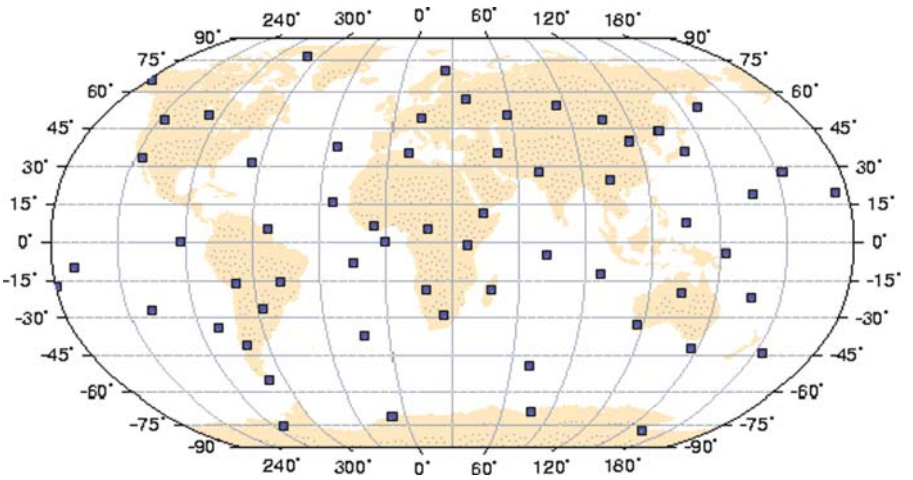


FIGURE 1 The 60 infrasound stations of the IMS infrasound network.

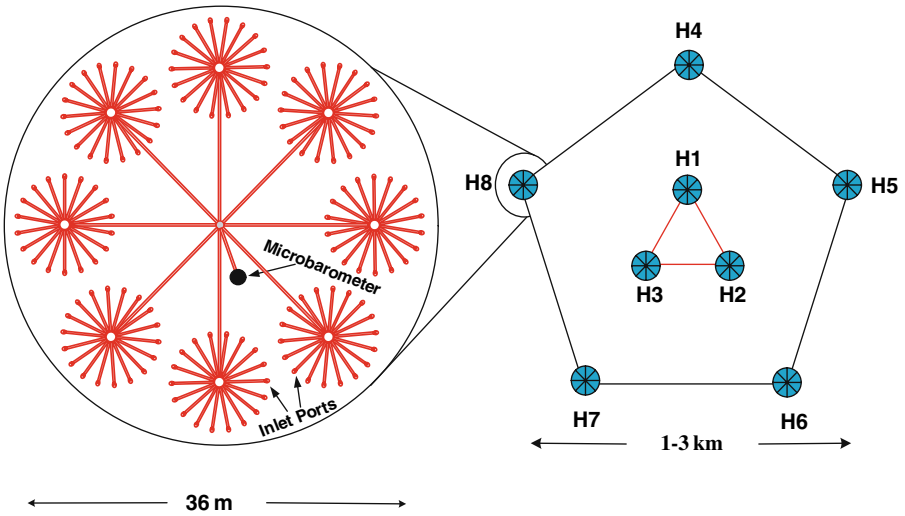


FIGURE 2 Typical eight-element infrasound array configuration (right) and details of an array element with a 36-m wind-noise-reducing port filter (left).

Incoherent fluctuations, such as wind generated eddies, average out while coherent infrasound signals sum up. In this way, noise-reducing pipe filters effectively increase the signal-to-noise ratio and thus improve the detectability of infrasound signals. Infrasound data contributes to the location of the sources of such events during IDC processing. The results of infrasound data processing are also combined with results obtained by the two other IMS waveform technologies (seismic and hydroacoustic).

1 • Overview of IDC Processing

The IDC receives the data from all IMS stations in near real time. The IMS data is stored and referenced in the IDC database, organised into 30 min time blocks and processed automatically. The first stage of automatic processing is "station processing", where the system attempts to detect signals and extracts their characteristics at the individual stations. Station processing includes an initial attempt to identify individual types of phases and another step where noise detections get separated. Station processing is followed by network processing, which combines all relevant detections from all available stations and attempts to build events from these associations. The result of automatic processing is a list of events called the Standard Event List (SEL), which facilitates the interactive analysis by the IDC analysts. The Reviewed Event Bulletin (REB) is issued after analyst review

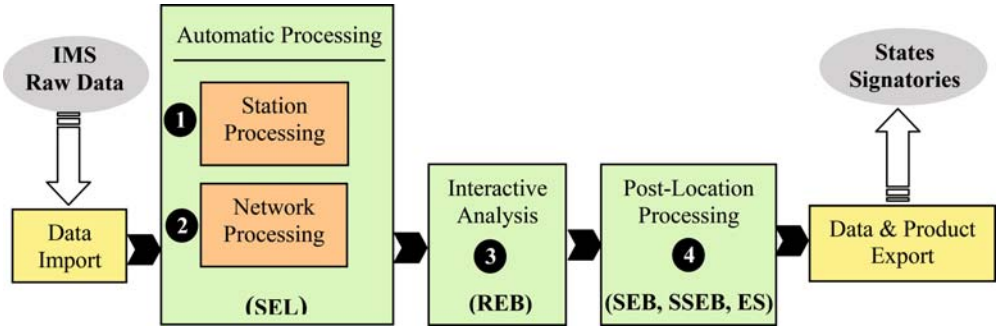


FIGURE 3 Functional model of the IDC automatic and interactive data processing.

of the results from automatic processing. Subsequent automatic processing generates the Standard Event Bulletin (SEB), the Standard Screened Event Bulletin (SSEB) and the Executive Summary, which constitute the final products provided to the States Signatories.

The flowchart in Figure 3 summarises the four major phases of the IDC processing: station processing, network processing, interactive analysis and post-location processing.

2 • Infrasonic Station Processing at the IDC

Infrasonic data received from the IMS are composed of data frames sampled at 20 Hz. For some stations, the atmospheric pressure data are accompanied by local meteorological data – e.g. local temperature, wind speed and direction – with typical sampling frequencies of 1 Hz.

As with the other waveform technologies, there is a provision to validate the integrity of each data frame before it is stored in the database. An automatic process constantly scans the database searching for available data from each array. When there are sufficient data to build a 30-min time interval for automatic processing, the system schedules the time block for station processing. This is usually done in quasi real-time.

The primary objective of station processing is to analyse the infrasonic data for a given time block in order to detect any signals whether or not they are visible above the background noise. The waveforms recorded at the stations are the signature of the wave trains propagating in the different layers of the atmosphere and the detected signals correspond to the different phases. The second function of station processing is to identify the type of phases detected at the station.

2.1 Pre-processing: Automatic Data Quality Control

The quality of the station processing results depends not only on the detection algorithm used, but also on the quality of the input data. A data quality control (QC) testing algorithm is applied as a pre-processing stage to ensure that subsequent detection processing runs under optimal conditions. QC searches for, and attempts to "repair" poor quality data samples in the waveforms or takes the decision to not use these samples. These include spikes (isolated large amplitude data values), dropouts (missing data), and extended sequences of repeated values.

2.2 Automatic Detection Processing

The IMS infrasound arrays are designed to include four or more sensors. This configuration with a large number of elements permits array-processing techniques to be applied to detection processing. The principle is that any signal crossing the array is recorded in a coherent way on the individual sensors, whereas noise is random and incoherent at different sensors. Coherence is therefore an efficient criterion for extracting signals and distinguishing them from the background noise.

The individual sensors record the signal at different times as it propagates across the array. Assuming a coherent signal, the cross-correlation between pairs of sensors allows the time delays between individual sensors to be determined, these are in turn dependent on the geometry of the array and on the wave front characteristics. The propagation parameters – azimuth and speed – of the signal are derived from the time delay analysis. IDC uses the Progressive Multi Channel Correlation (PMCC) technique [2] to detect coherent wave fronts that are consistently recorded at several sensors. The closure formula (1) on time delays is a simple and efficient technique to verify the consistency of the signal detections at individual sensors. The sum of time delays tends to zero for consistent signals.

$$\Delta t_{i,j} + \Delta t_{j,k} + \Delta t_{k,i} = 0 \quad \text{Closure formula for time delays } \Delta t \text{ between three sensors } i,j,k \tag{1}$$

The detection analysis is performed in the combined time-frequency domain. The frequency interval of interest between 0.1 and 4.0 Hz is divided into a series of relatively narrow bands. The data from each individual sensor is first filtered in all the filter bands. Then the detection processing is run using sliding, overlapping time windows. The cross-correlation and consistency values are computed for each time window and frequency band. Azimuth and speed values are obtained when the detection criteria are satisfied, i.e. when the closure formula is verified for a minimum number of sensors. This defines the so-called detection pixels with associated signal in the time-frequency domain.

A grouping algorithm is the next step in the PMCC detection process. Individual detection pixels that have similar signal attributes (i.e., are close enough in the time-frequency domain) constitute a "family". Only the largest and most stable families are preserved for subsequent processing. For each family, the azimuth and signal speed values are calculated as the average values of all pixels in the family. The PMCC time-frequency analysis provides rather complete information regarding the detected signal, and particularly the evolution of the propagation parameters and the frequency content with time.

Very small coherent signals with signal-to-noise ratio substantially lower than 1 are detected by the IDC automatic system. In a sense, this is very positive as it shows the very high detection capability of the IDC infrasound processing system. However, this also means that IDC may have to deal with a very large number of automatic detections, which can become a problem when the detections are to be associated with events during subsequent network processing. The number of bogus automatic events grows with the number of automatic detections and measures to limit the problem have to be taken.

The algorithm used initially at the IDC in its 1st years of existence was developed at the Prototype IDC in Arlington [3], USA, and was, in a way, similar to PMCC. It also extracted coherent signals from incoherent background noise. The detection process was controlled by two parameters: the Fisher statistics values (F-statistics) and the short-term average over long-term average ratio (STA/LTA). The coincidence detector required both F-statistics and the STA/LTA values to exceed some given thresholds to limit the detection process to the most coherent and energetic signals.

IDC decided against applying STA/LTA threshold criteria to reduce the number of coherent detections, as it introduces a high risk of missing signals of interest to CTBT. The primary task of the detector is to detect any coherent signal. A subsequent "cleaning" process analyses the signal characteristics in order to categorise and mask the uninteresting signals, and identify the phase names (see Part 2.3). Consideration is being given to computing F-statistics for each detected signal to provide information complementary to the PMCC detection attributes.

2.3 Automatic Signal Categorization and Phase Identification

The mission of phase identification processing is to examine the automatic detections and determine the signal type to be either a signal or noise and further to assign a phase name according to the extracted signal features (characteristics). The notion of noise or signal used at the IDC reflects the focus of monitoring at the PTS. As far as PTS is concerned, detection of nuclear explosions remains the main monitoring task of the IDC.

The IMS stations are distributed all over the world, in very different environments – e.g. mountainous areas, coastal, industrialized or rural environments, or extreme latitudes. The global network of IMS stations records a large number of signals from different sources of either man-made or natural origin every day. Preliminary studies at selected stations have allowed the IDC to build a knowledge base of types of signals detected by IMS infrasound stations. Despite the fact that local sources inevitably generate some very specific signals at the station, it is possible to make general observations that can be applied to the entire IMS network.

The IDC processing is focused on the frequency band from 0.1 to 4 Hz. This band is important for the detection of nuclear explosions, but there are many other types of event with transient or continuous characteristics recorded in this band. The duration of the signal wave trains, the frequency content and many other parameters are likely to be key factors in the identification of the type of signal detected. Transient signals with high frequencies, e.g. mining activity, lightning, ocean surf, are usually of local or regional origin and may not be of interest in view of the mission of the organization. Repetitive or long-duration signals – microbaroms, industrial activity – may also fall in this category.

The phase identification is currently done in a rather basic way at this level of the processing and it is applied to the signals that have not previously been declared as noise. Initial phase names are attributed to the detections based on the horizontal wave speed of signal propagation across the array. Infrasound phases I characterize acoustic waves with speed ranging from 280 to 450 m/s. Seismic phase names ISx and IPx are attributed to signals recorded at seismic speeds respectively greater than 2,900 m/s and 5,400 m/s, respectively (note, that microbarometers are known to behave as seismometers for signals from large earthquakes). All remaining phases are considered as noise N.

The IDC is studying the possibility of applying a more advanced nomenclature for infrasonic waves that would identify the arrivals according to the waveguide boundaries in the atmosphere. Three major phase types and relevant travel-time-tables are being considered, for infrasound waves propagating through the troposphere (Iw), the stratosphere (Is) and the thermosphere (It).

3 • Automatic Network Processing

The function of IDC network processing is to associate the signals detected and identified during station processing and to automatically "build" and locate events. Only those detections that have

not been identified as noise can contribute to the association process.

Network processing utilizes all phases available from seismic, hydroacoustic and infrasound (SHI) data processing. Three automatic event bulletins are produced in sequence. The first bulletin is the Standard Event List 1 (SEL1). It is produced approximately 2 h after real time and includes detections from the primary seismic and hydroacoustic IMS stations. Standard Event List 2 (SEL2), produced 6 h after real time, integrates the detections from auxiliary seismic IMS stations, and includes the infrasound detections (as the infrasonic signals propagate at much slower speed than seismic or hydroacoustic signals). The final Standard Event List 3 (SEL3), produced 12 h after real time, completes the IDC automatic processing and refines the location estimates. SEL3 is the final stage of network processing; it is the best automatic product and the baseline from which IDC analysts begin an interactive review.

3.1 Automatic Phase Association

The technique used to associate SHI phases identified during station processing is based on a global grid search. Grid cells completely cover the surface of the Earth with additional depth cells defined in the areas where deep earthquakes are known to occur. Each grid cell represents a hypothetical source location from which signal propagation times can be predicted at the IMS stations. Seed events are built when several detections are found to be compatible in time, slowness and azimuth within defined uncertainty limits. The location procedure is applied to all seed events, and conflicts in possible phase associations to more than one event are resolved to reduce the number of hypotheses and build the most probable events with the appropriate phases.

To ensure homogeneous detection capability over the surface of the Earth, the events built by the automatic processing are tested for compliance with the so-called weighted-count test. A phase weight is attributed to each type of phase and signal attribute. The sum of the weights must be greater than a certain threshold to pass the SEL test. In practice, this means that two infrasound phases or one infrasound phase combined with one primary seismic or hydroacoustic phase are sufficient to build an automatic event in the SEL2 and SEL3 bulletins.

3.2 Automatic Event Location

The algorithms used for locating events in both the IDC automatic and interactive processing systems are identical. Hypocentres are determined via an iterative non-linear least-squares inversion originally developed by Jordan and Sverdrup [4] and modified by Bratt and Bache [5] to include azimuth and slowness data.

At each iteration, residuals in time, slowness and azimuth are calculated between the measured and the predicted values. The event location determination is achieved when the residuals meet some convergence criteria.

3.3 Travel-Time Tables

The association and location algorithms use the same set of travel-time tables for infrasound data. These travel times have been computed with numerical acoustic ray tracing models, simulating the propagation of acoustic waves in the two principal atmospheric ducts: the stratosphere and the thermosphere. Each infrasound array utilizes a specific set of travel times, which are built with dependencies on azimuth and distance, currently for surface sources.

The propagation model takes into account seasonal variations of the atmospheric conditions, using the MSISE90 atmospheric model [6] combined with HWM-93 wind profiles [7]. The preferred approach is to utilize a large set of travel-time tables smoothly distributed throughout the year.

4 • Interactive Processing

Interactive analysis is the stage of IDC processing where results produced by automatic processing are reviewed and refined by analysts. This stage is concluded by the publication of the REB. For each automatic event built in the SEL3 bulletin, the analysts review the waveform signals and relevant phases detected at the IMS seismic, hydroacoustic and infrasound stations. Phases may need to be manually corrected – retimed, deleted, added, renamed – and the signal features may need to be recomputed and refined. The analysts may also modify the results from automatic phase association, by adding or disassociating phases to the events. The location of the events is performed using the procedures described above.

A weighted count test is applied to REB events in order to retain the largest events. The weights depend on the phase and station types and are slightly different from those used during SEL automatic processing. Events with total weight less than a given threshold are excluded from the REB.

The recent developments in detection processing, namely the introduction of the PMCC algorithm, have shown the importance and necessity of computing three-dimensional information – i.e. azimuth and speed versus time and frequency – during the analysis of infrasound data. The standard IDC tool for reviewing the SHI waveforms is currently not appropriate for a three-dimensional display and efforts are underway to update the IDC

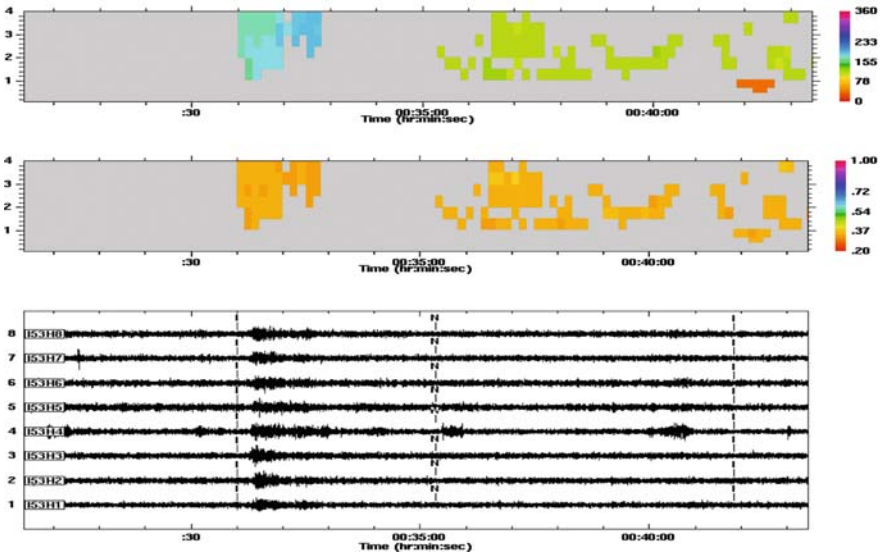


FIGURE 4 Interactive tool for infrasound data processing at the IDC.

interactive software by integrating a new infrasound module. The infrasound-specific tool is designed to display the waveforms along with a coloured spectrogram representing the derived azimuth and speed attributes computed by the detector in a time frequency space (Figure 4). The tool provides a comprehensive picture of the signal properties as they evolve with time.

As infrasound signals are often characterized by low signal-to-noise ratio, the interactive review tool emphasizes the derived azimuth and speed data more than the signal waveforms. In that sense, interactive processing of infrasound data is based on a significantly different approach compared to the seismic and hydroacoustic data, where the waveforms are the key time series for the analysts to identify useful phases.

5 • Post-location Processing

Post-location processing provides supplementary information about the characteristics of the events listed in the REB. The signal attributes are automatically re-computed providing an accurate estimate of various parameters including amplitude, period, magnitudes, spectral characteristics etc. The final event bulletin produced at this stage is called the Standard Event Bulletin (SEB).

The characteristics of the SEB events are automatically analysed and used to screen out events that have a natural or non-nuclear origin. Events that are found to pass the screening criteria are grouped in the SSEB.

An Executive Summary synthesises the results produced by IDC processing. The Executive Summary provides to the States Signatories high level information such as the total number of detected events, the categorization of these events based on the screening criteria, and the current status of the data acquisition for all stations in the IMS network.

Disclaimer

The views expressed herein are those of the authors and do not necessarily reflect the views of the CTBTO Preparatory Commission.

References

1. Christie, D. R., J. A. Vivas Veloso, P. Campus, M. Bell, T. Hoffmann, A. Langlois, P. Martysevich, E. Demirovic and J. Carvalho, *Detection of Atmospheric Nuclear Explosions: The Infrasound Component of the International Monitoring System*, Kerntechnik, 66, 96–101, 2001.
2. Cansi, Y., *An Automatic Seismic Event Processing for Detection and Location: The P.M.C.C. Method*, Geophys. Res. Lett., 22, 1021–1024, 1995.
3. Brown, D.J., C.N. Katz, R. Le Bras, M.P. Flanagan, J. Wang and A.K. Gault, *Infrasonic Signal Detection and Source Location at the Prototype International Data Centre*, Pure Appl. Geophy., 159, 5, 1081–1125, 2002.
4. Jordan, T.H. and K.A. Sverdrup, *Teleseismic Location Techniques and Their Application to Earthquake Clusters in the South-Central Pacific*, Bull. Seism. Soc. Am., 71, 1105–1130, 1981.
5. Bratt, S.R. and T.C. Bache, *Locating Events with a Sparse Network of Regional Arrays*, Bull. Seism. Soc. Am., 78, 780–798, 1988.
6. Hedin A. E., *Extension of the MSIS Thermosphere Model in the Middle and Lower Atmosphere*, J. Geoph. Res., 96, 1159–1172, 1991.
7. Hedin, A. E, E. L. Fleming, A. H. Manson, F. J. Schmidlin, S. K. Avery, R. R. Clark, S. J. Franke, G. J. Fraser, T. Tsuda, F. Vial and R. A. Vincent, *Empirical Wind Model for the Upper, Middle, and Lower Atmosphere*, J. Atmos. Terr. Phys., 58, 1421–1444, 1996.

80

The Design and Performance of Infrasound Noise-Reducing Pipe Arrays

Benoit Alcoverro

CEA/DASE, BP 12, 91680 Bruyères Le Châtel, France

1 • Introduction

The detection of infrasonic waves in the atmosphere is almost always limited by background noise associated with wind-generated turbulence. Traditionally, this problem has been resolved by using a pipe array on the input to the infrasonic sensor to spatially average the surface micropressure over the area surrounding the sensor. Since wind speed and turbulence increase with height in the lower boundary layer, the inlets of this pipe array must be located close to the surface. The pipe array is designed specifically to take advantage of the spatially incoherent structure of the background noise compared with the coherent infrasonic micropressure signal (McDonald and Herrin, 1975; McDonald et al., 1971). The turbulent motion of

air in the boundary layer is very complex. Atmospheric turbulence near the surface consists of relatively short-lived, rapidly evolving eddies that are advected by the mean wind. The effective scale of these eddies ranges from less than 1 cm to more than 10 m. The scale (S) of an eddy is given in terms of the frequency, f, of the pressure fluctuations and the wind speed, V, by the well-known relation, $f = V/S$. The ratio between the propagation velocity of the wind-generated noise and the propagation speed of the infrasonic signal across the array, v, is much less than unity. For a given frequency, this implies that several wavelengths of atmospheric noise exist in the same space occupied by only one wavelength of the acoustic signal. The summation of pressure – comprising both acoustic signal and atmospheric noise – from a spatial array of inlets results in a significant increase in the signal-to-noise ratio. In this chapter, we describe the acoustics of noise-reducing pipe arrays and the influence of the pipe array response on the performance of an infrasonic array element.

2 • Historical Background

Almost all noise-reducing pipe array systems in use at the present time are based on the original linear array design proposed by Daniels (Daniels, 1959). The basic Daniels array consists of a long (typically 300 m) linear succession of rigid pipes of increasing diameter, with high impedance capillary inlet ports located near each junction. This array, which is designed to eliminate acoustic reflections at each inlet port, provides a large increase in the signal-to-noise ratio for frequencies lower than 1 Hz. However, this linear pipe system exhibits a non-isotropic noise reduction and sensitivity (Noel and Whitaker, 1991). A circular pipe array system, such as that proposed by Burridge (Burridge, 1971) and Grover (Grover, 1971) has a better and more omnidirectional response.

A comparison of the performance of different systems shows that the long Daniels pipe system can be replaced by an equally effective circular system of smaller size (Noel and Whitaker, 1991; Grover, 1971). The performance of a circular system with a diameter of 100 m is comparable with the performance of a 300-m-long Daniels line array at frequencies below 0.5 Hz. The azimuthal dependence of the response function for a circular noise-reducing system is also fairly isotropic when the wavelength exceeds the diameter of the array and when the impedance of the inlet ports is sufficiently large to allow perfect acoustic summation over the entire circumference of the circle (Burridge, 1971; Haak and de Wilde, 1996). However, there are a number of problems with these

early circular pipe array designs. These large-diameter pipe arrays require a large area of land and they are difficult to install. They also require constant maintenance to ensure that the high-impedance capillary inlet ports are not obstructed. In addition, large-diameter circular pipe arrays are subject to unwanted acoustic resonances in the [0.02–4] Hz band. Finally, we note that pipe arrays with high impedance inlets are sensitive to the ambient temperature variations. These temperature variations may induce pressure fluctuations that are comparable to signal levels. Another type of noise reduction system based on the use of microporous hose has been described by Noel and Whitaker (1991). The porous pipe solution is inexpensive and provides a significant improvement in the signal-to-noise ratio in the [0.1–10] Hz frequency band (Haak and de Wilde, 1996). However, the performance of these systems deteriorates rather quickly with time in some environments, especially in tropical equatorial environments. The major drawback with the use of porous pipe noise reducing systems is the variation in the physical properties when the hose becomes saturated with rain or partially blocked by dust and dirt. Although this solution is attractive as well as efficient, it is not sufficiently robust to be used under operational conditions at many infrasonic monitoring stations. The difficulties outlined above with the use of high impedance inlet ports and porous hoses have led in recent years to the development of a new type of robust infrasonic noise-reducing pipe array based on the use of low-impedance inlet ports.

3 • Design Principles for a Robust Infrasonic Noise-Reducing System

In order to avoid the problems associated with the use of high impedance inlets and porous materials, most modern pipe arrays are designed to be perfect isotropic summing systems with low-impedance inlets arranged symmetrically as close as possible over a circular area. A key feature of this design is the fact that the acoustic impedances between each inlet port and the sensor are identical. The simplest way to achieve this design is to connect each inlet directly to a central (or primary) summing manifold (the sensor). However, in order to reduce the complexity of the pipe array, intermediate (or secondary) summing manifolds are inserted between groups of inlets and the sensor. In order to satisfy the requirement that the acoustic impedance between each inlet port and the sensor are identical for all ports and to keep the design as simple as possible, the inlet ports are usually distributed in a circle or semi-circle around the secondary

manifolds, which in turn are evenly distributed on a circle with the primary manifold at the center. This symmetrical design also ensures that the response does not depend on signal azimuth. A typical design is shown in Figure 1. This pipe array design was proposed by Alcoverro (1998) and enhanced for use in the global Comprehensive Nuclear Test Ban Treaty Organization (CTBTO) network by Christie et al. (2001). The design shown in Figure 1 consists of 32 low-impedance inlet ports equally spaced on an annular area with a 16-m overall diameter. Connections between the inlet ports and the intermediate manifolds and between the manifolds and the fore-volume of the microbarometer sensor are made using 15-mm I. D. pipes. Straight pipes ensure shorter acoustic paths, but pipes with large curvatures can be used when necessary.

The response of the noise-reducing system illustrated in Figure 1 depends primarily on the length and inside diameter of the connecting pipes, the volume of the manifolds, and to a lesser degree, on the number of inlets and the properties of the sensor itself. Noise reduction efficiency at a given frequency is mainly governed by the number of inlets, the diameter of the pipe array, D, and the wind speed, V, in accord with the relation $f = V/D$. The following discussion is focused on the optimization of the design parameters.

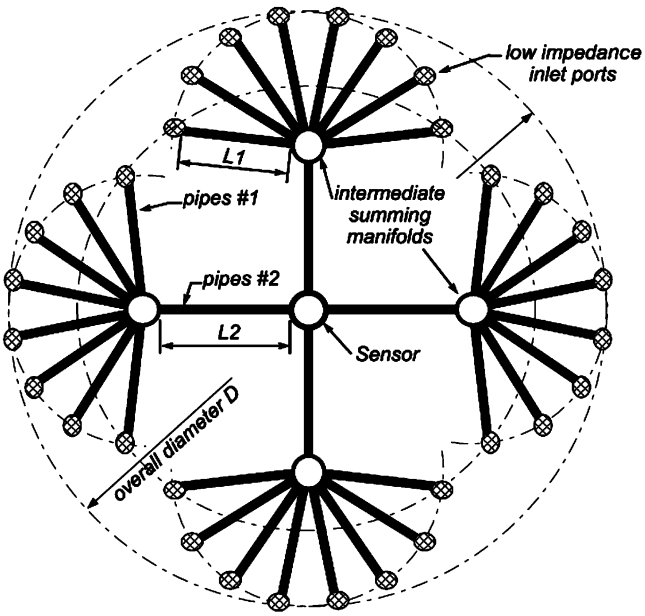


FIGURE 1 Plan view of a robust infrasonic noise-reducing system.

4 • Noise-Reducing Pipe Array Design and Performance

The pipe array is designed to provide maximum wind-noise reduction over a specified frequency range with minimum signal distortion. For example, the noise-reducing systems used in the CTBTO global infrasound monitoring network are designed to have maximum noise reduction in the 0.02–4.0 Hz frequency band. One way to optimize the design parameters is to conduct a series of experimental measurements. However, this method is impractical due to the large number of parameters that need to be optimized. The method used for design and optimization is therefore based on numerical simulations of pipe array performance (Alcoverro and Le Pichon, 2005). There are three parts to the simulation:

- evaluation of the acoustic frequency and phase responses based on an electro-acoustic model
- evaluation of the spatial filtering of acoustic waves as a function of the distribution of inlet ports and angle of incidence
- computation of noise reduction as a function of frequency based on an empirical wind turbulence model

4.1 Acoustic Frequency Response Simulation

The acoustic frequency response is calculated by introducing an electro-acoustic model with accurate electrical equivalents for all elements in the system, such as the inlet ports, pipes, manifolds, and the sensor. This model is illustrated in Figure 2 and the electrical equivalents for the elements in the system are given in Table 1. This calculation does not depend on the spatial distribution of the inlet ports.

The use of an electrical analogy for the acoustic system greatly simplifies the calculations. To first order, the pipe array and sensor are modeled as a linear system with multiple inputs (the

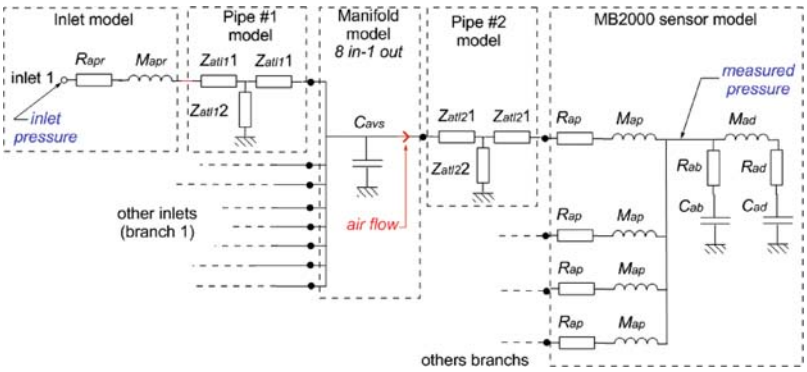


FIGURE 2 Electro-acoustic model of infrasonic noise-reducing system.

TABLE 1 Expressions for the electro-acoustic elements used in the pipe array model

Elements	Model	Values
Inlet port: *Radiation impedance* a: radius of measurement port ρ: air density k: acoustic wave number c: sound speed in air	See section on radiation impedances in Beranek (1986). $Z_{ar} = R_{apr} + j\omega \cdot M_{apr}$ $R_{apr} = Z_c \frac{(ka)^2}{1+(ka)^2}$ $M_{apr} = Z_c \frac{ka}{1+(ka)^2}$ with: $Z_c = \frac{\rho c}{\pi a^2}$, $k = \frac{2\pi f}{c}$	$a < 10\,mm, f = 10\,Hz$ $\rho = 1.225\,kg/m^3$ $\lvert Z_c \rvert > 1300\,k\Omega A^*$ $R_{apr} = 52E-6\,\Omega A^*$ $M_{apr} = 0.26\,\mathrm{kg/m^4}$ $\Rightarrow R_{apr} << \lvert Z_c \rvert$
Pipes 1 and 2: *Dissipative transmission line* l: length of pipes η: viscosity coefficient of air S: cross section of the pipe	$Zatli1 = Z_c \cdot \tanh\left(\frac{\Gamma \cdot l}{2}\right)$ $Zatli2 = \frac{Z_c}{\Gamma \cdot l}$ $Z_c = R_o \cdot \left[\left(1 + 0.369 \cdot r_v^{-1}\right) - j \cdot 0.369 \cdot r_v^{-1}\right]$ $\Gamma = k \cdot \left[1.045 \cdot r_v^{-1} + j \cdot \left(1 + 1.045 \cdot r_v^{-1}\right)\right]$ with $R_o = \frac{\rho \cdot c}{S}$, $r_v = \sqrt{\frac{2 \cdot \rho \cdot f \cdot S}{\eta}} \geq 1$	$\eta = 18.610^{-6}$ (Pa.s) at 20°C. These broad tube approximations (Keefe, 1984) are close to the exact wave parameters (Benade 1968).
Summation cavity: *Acoustic Compliance* V_s: cavity volume	$C_{avs} = \frac{V_s}{\rho \cdot c^2}$	$V_s = 0.310^{-3}\,\mathrm{m^3}$
Microbarometer		MB2000 type
Inlet pipes: *Mass + loses* l_p: length of inlets r_p: radius of inlets	$R_{ap} = \frac{8 \cdot \eta \cdot l_p}{\pi \cdot r_p^4}, M_{ap} = \frac{\rho \cdot l_p}{\pi \cdot r_p^2}$	$l_p = 30\,\mathrm{mom}$, $r_p = 5\,\mathrm{mm}$
Inner volume: *Acoustic compliance* V_{ab}: volume of cavity and loses R_{ab}	$C_{ab} = \frac{V_{ab}}{\rho \cdot c^2}, R_{ab}$	$V_{ab} = 0.610^{-3}\,\mathrm{m^3}$, $\mathrm{C_{ab}} = 4.3310^{-6}\,\mathrm{m^3/Pa}$, $R_{ab} = 500\,\Omega\mathrm{A}^*$
Sensitive bellows: *Mass spring oscillator* R_{ad}, M_{ad}, C_{ad}: sensor bellows	$Z_{ad} = R_{ad} + j\omega \cdot M_{ad} + \frac{1}{j\omega \cdot C_{ad}}$	$C_{ad} = 5.110^{-11}\,\mathrm{m^3/Pa}$ $M_{ad} = 2.8510^4\,\mathrm{kg/m^4}$ $R_{ad} = 2.6210^5\,\Omega\mathrm{A}^*$.

*ΩA denotes MKS acoustic ohms in Pa s/m^3

inlets) and a single output (the pressure at the sensor). For each input, the transfer function is calculated by assuming that all other inputs are connected to ground. The equivalent circuit equations for this system can be set up as follows. Each node in the circuit is represented by a connection between elementary acoustic elements. The acoustic flow q_v through each element depends only on the acoustic impedance $Z_{j,k}$ and the input and output pressures N_j, N_k.

$$q_{v_{(j,k)}} = \frac{N_j - N_k}{Z_{(j,k)}}$$

Since the sum of the acoustic flow is null at each node (Kirchoff's law),

$$\sum_j q_v(j,k) = 0,$$

the nodal equations for the equivalent circuit can now be written to give a set of linear equations where k is the specified node and j the adjacent nodes:

$$\sum_j \frac{\boldsymbol{N_j}}{\boldsymbol{Z_{(j,k)}}} - \boldsymbol{N_k} \cdot \sum_j \frac{1}{\boldsymbol{Z_{(j,k)}}} = 0.$$

These equations can be easily written in matrix form using an impedance matrix [Z] and nodal $\{N\}$ and inlet $\{e\}$ pressure vectors:

$$[Z] \cdot \{N\} = \{e\}$$

The vector $\{N\}$ is then calculated by matrix inversion. The calculation is performed for each frequency as follows:

$$\{N(\omega)\} = [Z(\omega)]^{-1} \cdot e(\omega)$$

The transfer function linking the pressure in the sensor cavity N_p to the excitation present at air inlet e_j, is obtained using an input vector $\{e\}$. All the components e_k of this vector are null, except for the component corresponding to air inlet j ($e_j \neq 0$):

$$\begin{Bmatrix} \vdots \\ N_j(\omega) \\ \vdots \\ N_p(\omega) \end{Bmatrix} = [Z(\omega)]^{-1} \cdot \begin{Bmatrix} 0 \\ \vdots \\ e_j(\omega) \\ \vdots \\ 0 \end{Bmatrix}$$

Finally, the transfer function $H_j(\omega)$ is given by:

$$H_j(\omega) = \frac{N_p(\omega)}{e_j(\omega)}$$

These calculations can be carried out very rapidly on a computer for each inlet and parametric analyses can then be undertaken to establish simple design equations. The overall response is given by the summation of the elementary inlets responses (superposition principle).

The response of the system takes the form of a high-order low-pass filter with multiple resonances. Figure 3 illustrates the influence of the design parameters on the frequency response. The high-frequency resonances in this system are due to the low impedance termination at the inlet ports and the impedance mismatch between the pipes and the manifolds.

The results presented in Figure 3a show that the lowest frequency resonance depends on the overall diameter of the pipe array. This resonance is generated by reflected waves inside the connecting pipes between the sensor and the intermediate summing manifold. Infrasonic signals with frequencies at or near a resonant frequency will be significantly distorted. Pipe arrays should therefore be designed to eliminate all resonances from the passband of interest. The lowest, or fundamental resonance frequency, f_r, in this pipe array can be approximated by:

$$f_r = 0.7\,c/D,$$

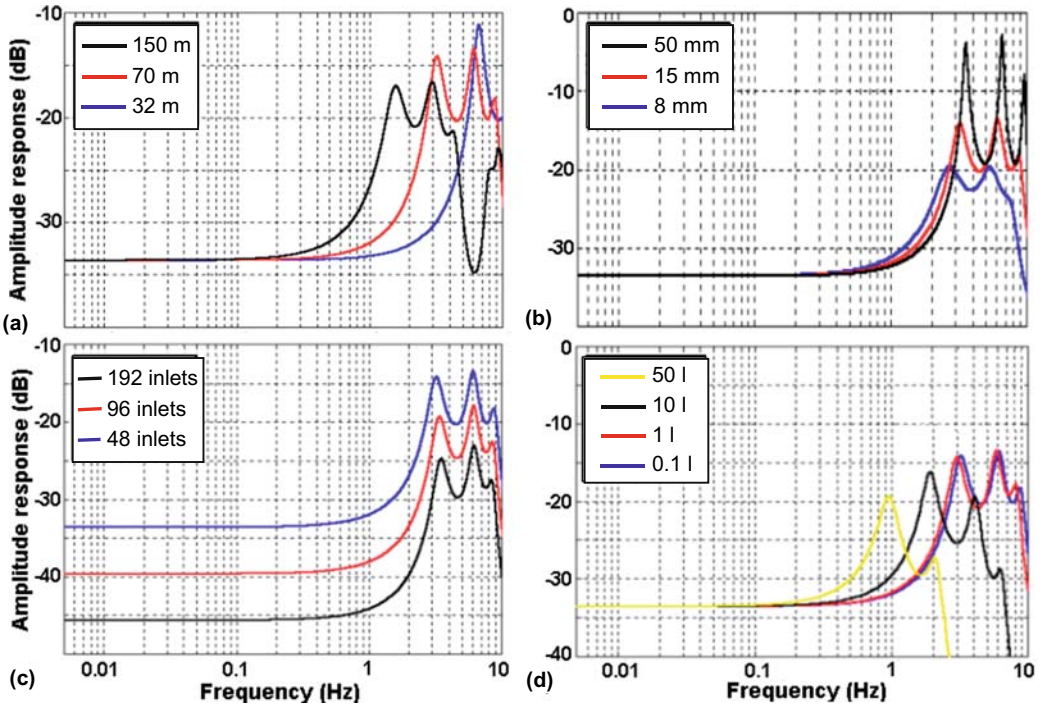


FIGURE 3 Frequency response for various pipe array parameters. (**a**) Array diameter. (**b**) Pipe inside diameter. (**c**) Number of inlets. (**d**) Volume of manifolds.

where c is the speed of sound and D is the diameter of the noise-reducing system.

Figure 3b illustrates the influence of the inside diameter of the connecting pipes on the response of a 70-m-diameter pipe array system. Arrays with pipes with large inside diameters have sharp high-amplitude resonances. In contrast, if the inside diameter of the pipes is small, then the viscothermal losses inside the pipe (resistance) are increased, and this, along with the associated manifold capacitance, dominates the morphology of the lowest resonance in the system. Increases in inside pipe diameter above 50 mm have no significant influence on resonance frequency. The results of these simulations indicate that 15 mm represents a reasonable choice for the inside diameter of the connecting pipes.

The simulation results presented in Figure 3c show that the shape of the frequency response is independent of the number of inlet ports. In this case, only the amplitude of the response at lower frequencies is affected. Since the gain of the elementary transfer function is equal to the inverse of the number of inlets, the response will be unity after summing.

The results illustrated in Figure 3d show that the response is quite sensitive to the volume of the manifolds. Manifolds with volumes of more than 1 l result in a significant decrease in the fundamental resonant frequency because the acoustic compliance of air inside the cavity coupled with the acoustic mass of the connecting pipe determine the position of the lowest frequency resonance.

In a well-designed pipe array, the acoustic frequency response is only limited at higher frequencies by the presence of resonances associated with reflections in the connecting pipes. The frequency of these resonances is determined primarily by the length of the pipes. These resonances can be eliminated by introducing capillaries with impedances that match the characteristic impedance of the pipe at all connections between the pipes and the manifolds (Hedlin et al., 2003). This solution works well, but the matching capillaries are sensitive to obstruction. Figure 4 shows the frequency response in amplitude and phase for 70-m-diameter pipe arrays with unmatching and matching termination impedances compared to an 18-m-diameter array with unmatched termination impedances. The properly terminated pipe array exhibits an extended flat frequency response but the phase rotation is much slower than that found for an uncompensated pipe array. This slower phase rotation may introduce a large group delay. If the impedance of the matching capillary changes because of moisture or dust, the group delay will also change

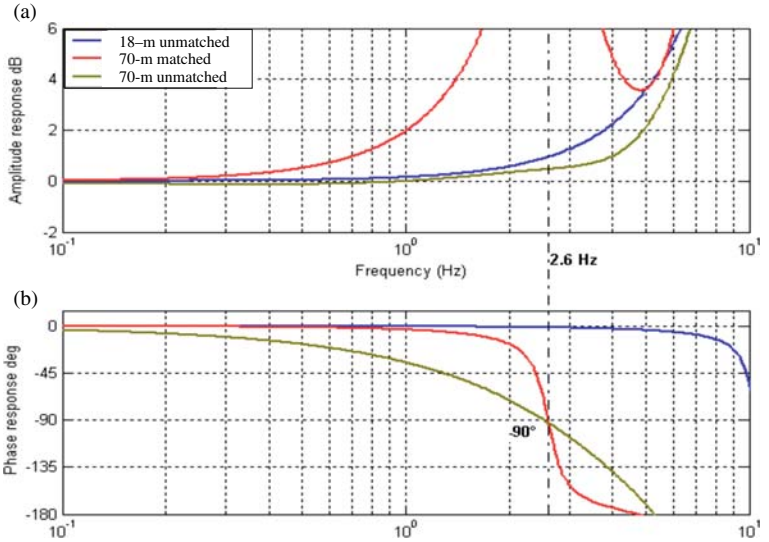


FIGURE 4 (**a**) Amplitude and (**b**) phase response of a 70-m-diameter pipe array with and without resonance-suppressing capillary impedances compared to 18-m-diameter uncompensated pipe array.

and this could introduce errors in the timing of signals, which could lead to signal detection problems (Alcoverro, 2002).

4.2 Spatial Filtering

Spatial filtering of the incident infrasonic wave leads to a further frequency limitation on wind-noise-reducing pipe arrays. The influence of spatial filtering on the signal generated by a wave propagating at c, with a trace velocity, v, depends on the diameter, D, of the pipe array and the angle of incidence, α, as shown on Figure 5. The integrated response is a low-pass comb filter with a first zero frequency, f_0, and a –3 dB attenuation frequency, f_{-3}, defined, for detection on a continuous circle, by:

In most cases, infrasonic signals arrive with low incidence angle (usually less than 30°). In this case, the cutoff frequency is lower than the first resonant frequency of the pipe array illustrated in Figure 1. By considering the –3 dB attenuation frequency, f_{-3} , of a 0° incidence wave, the high-frequency design limit, $f_{\max}$, is therefore given by

$$f_{\max} = 0.43\ c/D(\alpha = 0°)$$

4.3 Noise Reduction

The simulation of wind-noise reduction requires a turbulent wind-noise model. A physical wind-noise model is difficult to establish and so we have chosen to use an empirical method to

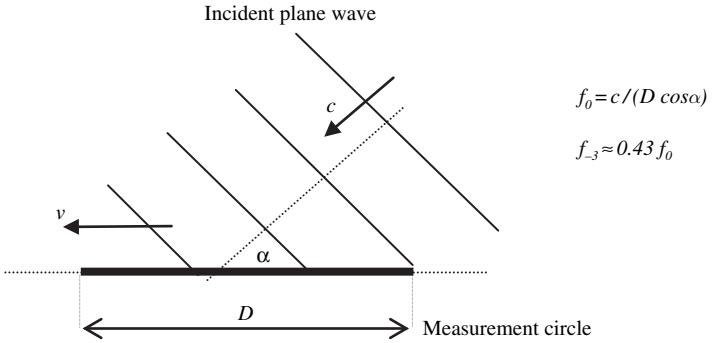


FIGURE 5 Incident plane wave impinging on a circular pipe array. In this example, the wave is detected at all points on the circumference of the measurement circle.

construct this model. Spatial coherency measurements of surface wind noise for wind speeds ranging from 0 to 7 m/s were carried out using an array of sensors. The correlation of the background noise was evaluated using standard coherence estimation procedures (Priestley, 1966; Mack and Flinn, 1971; McDonald and Herrin, 1975; Einaudi et al., 1989). For a given wind speed, the frequency, f_c, above which the coherence is lower than 0.5 has been extracted from measurements over 500-s time windows. We shall refer to this frequency as the decorrelation frequency. These observations show that advected turbulence is normally dispersive. The measured coherence decreases fairly rapidly as the distance between the sensors increases beyond the effective horizontal scale of the turbulent eddies in the atmospheric boundary layer. To a first approximation, f_c is proportional to the V/d ratio, where V is the average wind speed and d is the distance between the measurement points. A first-order least-squares fit to the data yields:

$$f_c \approx 0.11V/d.$$

The turbulent wind model generates wind signals by summing low-pass filtered normal random sequences. The filter frequencies are set to f_c to take into account the inlet configuration and average wind speed. The wind signals are convolved with the impulse response of each inlet as derived from the frequency analysis and the resulting background noise at the sensor is then calculated by summing over all of the convolved signals. The resulting noise spectrum at the sensor is divided by the ambient noise spectrum to determine the relative noise spectra as a function of frequency. Figure 6 shows examples of

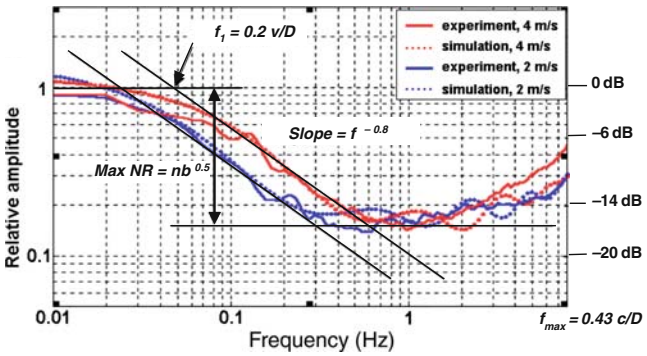


FIGURE 6 Reduction of wind-generated noise by a 16-m-diameter pipe array with 32 inlet ports. Simulated and experimental results are shown for wind speeds of 2 and 4 m/s.

simulated and measured relative noise spectra at the sensor for wind speeds of 2 and 4 m/s.

The simulated results are in good agreement with the observations. There is no noise reduction at very low frequencies because the decorrelation distance is larger than the diameter of the pipe array. Noise reduction is effective only at frequencies above f_1. This low cut-off frequency depends only on wind speed and pipe array diameter; i.e., f_1 is equivalent to the decorrelation frequency, *fc*. As the frequency increases, the relative noise level at the sensor decreases until the maximum reduction, as given by the square root of the number of inlets, is reached (−15 dB with 32 inlets). The slope of the curve above f_1 depends on the inlet distribution and is close to a power of −0.8 in frequency for the present design. The effective noise reduction at higher frequencies is constant. The apparent decrease in efficiency above 2 Hz is misleading. This increase is due to the acoustic resonance. However, the signal-to-noise ratio remains the same since the signal is also amplified by the resonances. In principle, there is no limitation on the degree of noise reduction at high frequencies, provided the pipe array has a sufficiently large number of inlets. But, as noted above, the useful passband is limited by spatial filtering to frequencies below the high frequency cut-off frequency, $f_{\max}$, which depends primarily on the diameter of the pipe array. Thus, the maximum noise reduction that can be achieved at the high-frequency cut-off depends only on the wind speed. These results are summarized in the simplified pipe array performance diagram shown in Figure 7.

The results presented here demonstrate that noise-reducing pipe arrays are very efficient at higher frequencies in relatively light winds. However, the noise level increases rapidly as the

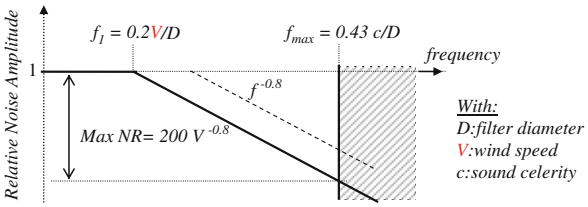


FIGURE 7 Example of simplified pipe array performance diagram and empirical relations. NR denotes the noise reduction.

wind speed increases. An estimate of the maximum possible noise reduction along with an estimate of the required number of inlet ports can be obtained by comparing the increasing level of wind-generated noise with the decreasing efficiency of the pipe array as a function of increasing wind speed. The results are sensitive to the geometrical configuration of the inlet ports and the structure of local wind-generated turbulence. These results show that a maximum noise reduction of about 35 dB will be obtained for wind speeds around 4.5–5 m/s. However, this optimum noise reduction can only be achieved by using more than 2,000 inlets. In a practical pipe array design, the number of inlets is usually limited to less than 150 in order to reduce cost and complexity, thus limiting the maximum noise reduction to about 22 dB.

5 • Conclusion

The wind-noise reducing pipe array described in this chapter represents a robust solution that provides efficient turbulent-wind-noise reduction using spatial pressure integration. The use of non-porous hoses and low-impedance inlet ports improves the performance stability and the resulting design is described by a simple, but realistic model. Frequency response and noise reduction are predicted with good agreement with experimental observations. The equations developed in this chapter can be used to design noise reduction systems for a wide range of infrasonic monitoring frequencies. The degree of noise reduction will be a negative power function of the wind speed. In contrast, wind noise increases as a positive power of wind speed. These two opposing factors will limit noise reduction to a maximum of about 35 dB for wind speeds around 4.5–5 m/s.

References

Alcoverro, B. (1998), "Acoustic filters design and experimental results," Proc. of the Infrasound Workshop, Commissariat à l'Energie Atomique, Bruyères-le-Châtel, France.

Alcoverro, B. (2002), "Frequency response of noise reducers," Proc. of the Infrasound Workshop, Royal Netherlands Meteorological Institute, Utrecht, Holland.

Alcoverro, B. and Le Pichon, A. (2005), "Design and optimization of a noise reduction system for infrasonic measurements using elements with low acoustic impedance", J. Acoust. Soc. Am. **117**, 1717–1727.

Beranek, L.L. (1986), "Acoustics," American Institute of Physics, Cambridge, MA.

Burridge, R. (1971), "The acoustic of pipe arrays," Geophys. J.R. astr. Soc. **26**, 53–69.

Christie, D.R., Vivas Veloso, J.A., Campus, P., Bell, M., Hoffmann, T., Langlois, A., Martysevich, P., Drmirovic, E., and Carvalho, J. (2001), "Detection of atmospheric nuclear explosions: the infrasound component of the International Monitoring System", Kernetechnik, 66, 96–101.

Daniels, F.B. (1959), "Noise reduction line microphone for frequencies below 1 cps", J. Acoust. Soc. Am. **31**, 529–531.

Einaudi, F., Bedard, A.J., and Finnigan, J.J. (1989), "A climatology of gravity waves and other coherent disturbances at the Boulder Atmospheric Observatory during March-April 1984", J. Atmos. Sci. **46**, 303–329.

Grover, F.H. (1971), "Experimental noise reducer for an active microbarograph array", Geophys. J.R. astr. Soc. **26**, 41–52.

Haak, H.W. and de Wilde, G.J. (1996), "Microbarograph systems for the infrasonic detection of nuclear explosions", Royal Netherlands Meteorological Institute Seismology Division, Scientific report WR 96-06.

Hedlin, M.A.H., Alcoverro, B., and D'Spain, G. (2003), "Evaluation of rosette infrasonic noise-reducer spatial filter", J. Acoust. Soc. Am. **114**, 1807–1820.

Mack, H. and Flinn, E.A. (1971), "Analysis of the spatial coherence of short-period acoustic-gravity waves in the atmosphere", Geophys. J.R. astr. Soc. **26**, 255–269.

McDonald, J.A., Douze, E.J., and Herrin, E. (1971), "The structure of atmospheric turbulence and its application to the design of pipe arrays", Geophys. J.R. astr. Soc. **26**, 99–109.

McDonald, J.A. and Herrin, E. (1975), "Properties of pressure fluctuations in an atmosphere boundary layer", Boundary-Layer Meteorol. **8**, 419–436.

Noel, S.D. and Whitaker, R.W. (1991), "Comparison of noise reduction systems", Los Alamos National Laboratory Report LA-12008-MS UC700.

Priestley, J.T. (1966), "Correlation studies of pressure fluctuations on the ground beneath a turbulent boundary layer", National Bureau of Standards Report No 8942, U.S. Dept. of Commerce.

81

Processing Infrasonic Array Data

John V. Olson and Curt A.L. Szuberla

Geophysical Institute, University of Alaska, Fairbanks, USA

1 • Array Processing

The subject of array processing has generated a vast literature. Several recent books have summarized the principal features of the subject and contain extensive bibliographies [2,10,15,16,28,31]. Early work on infrasound with a complete bibliography is summarized in a special issue of the *Geophysical Journal* of the RAS [6]. In this article we attempt to review a few of the basic procedures that are in current use within the infrasound community.

2 • Bearing Estimation

2.1 Traditional *f-k* Analysis

Consider M sensors in a plane located at $\boldsymbol{r}_i$ and a signal with a planar wavefront $s(x,t)$ propagating with a horizontal speed v in a direction $-\hat{k}_0$. The output of sensor i can be modeled as $x_i(t) = s(t+\boldsymbol{r}_i \cdot \hat{k}_0/v) + n(t)$ where $n(t)$ is the local noise signal. The simplest approach to beamforming is to sum the sensor outputs as $b(t) = \sum_{i=1}^{M} a_i x_i(t-\tau_i)$ where a_i is an arbitrary weight and τ_i is a time delay for sensor i. The idea is to choose τ_i so as to phase align the records such that $s(t)$ is in phase across all sensor records.

If this is done then $\tau_i = \boldsymbol{r}_i \cdot \hat{k}_0 / v$ and $b(t) = Ms(t) + \sum_{i=1}^{M} n_i(t - \tau_i)$. Assuming the $n_i(t)$ are uncorrelated with the signal and with each other then the beam power is $E[b^2] = M^2\sigma_s^2 + M\sigma_n^2$ where σ_s^2 is the signal variance and σ_n^2 is the noise variance. Thus, beamforming increases the signal-to-noise power by a factor of M.

Now if the beam signal, $b(t)$, is Fourier transformed the result is $B(f, \boldsymbol{k}) = \sum_{i=1}^{M} a_i X_i(f) exp(-j2\pi f \tau_i)$ where $X_i(f)$ is the transform of $x_i(t)$. This can be written in vector form as $B(f, \boldsymbol{k}) = \boldsymbol{A}^{\dagger}\boldsymbol{X}$ where $\boldsymbol{A}^{\dagger}$ represents the Hermitian conjugate of the steering vector $\boldsymbol{A} = a_1 exp(-j\omega\tau_1), \ldots, a_M exp(-j\omega\tau_M))^T$ and T is the transpose. The beam spectrum is then

$$P_B(f, \boldsymbol{k}) = \boldsymbol{A}^{\dagger}\mathsf{S}\boldsymbol{A} \tag{1}$$

where S is the spectral matrix. Recognizing that the $\tau_i = \boldsymbol{r}_i \cdot \boldsymbol{k}$ Equation (1) can be written in the traditional style of *f-k* analysis:

$$P_B(f, \boldsymbol{k}) = \sum_{i=1}^{M}\sum_{j=1}^{M} S_{ij}(f) e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_j - \boldsymbol{r}_i)} \tag{2}$$

In the case that $s(x, t)$ is a monochromatic plane wave with wave vector $\boldsymbol{k}_0$ and setting all the a_i to unity, one obtains

$$P_B(f, \boldsymbol{k}) = S(f) I(\boldsymbol{k} - \boldsymbol{k}_0)$$

where $I(\boldsymbol{k})$ is the spatial Fourier transform of the array, also called the *impulse response* of the array. The $\boldsymbol{k}$-space power can be seen to be the convolution of the array impulse response with the incoming wave spectrum. Further, $P_B(f, \boldsymbol{k})$ is characterized by a central maximum surrounded by side lobes, characteristic of the Bartlett estimator of Equation (1).

2.2 Array Aperture Function: Spatial Aliasing and Bandwidth

In the analysis of array data the geometry of the array plays an important role. Estimating the signal arrival times between pairs of signals is done using traditional cross-correlation calcualtions. However, the estimate of the cross-correlation function is affected by the geometrical arrangement of the array used to obtain the data.

Given a wave field $f(\boldsymbol{r}, t)$ that is sampled by sensors located at $\boldsymbol{r}_i, i = 1, M$ with each sensor weighted by $w(\boldsymbol{r}_k)$ the array output is

$$y(\boldsymbol{r}, t) = w(\boldsymbol{r}) f(\boldsymbol{r}, t). \tag{3}$$

Since a discrete set of sensors comprise the array, $w(\boldsymbol{r}) = \sum_{k=1}^{M} w_i \delta(\boldsymbol{r} - \boldsymbol{r}_i)$. In the case that all sensors are identical, $w_i = 1$. The Fourier transform of the array data is

$$Y(\boldsymbol{k}, \omega) = W(\boldsymbol{k}) * F(\boldsymbol{k}, \omega) \tag{4}$$

where $f * g$ represents the convolution of f and g. $W(\boldsymbol{k})$ is the aperture function and, for the case of a discrete array,

$$W(\boldsymbol{k}) = \sum_{i=1}^{M} w_i e^{j\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{5}$$

This aperture function is convolved with the k-space representation of the wave field and sets the limits on resolution of the array.

For symmetric arrays with uniform spacing $W(\boldsymbol{k})$ has side lobes away from the central lobe that are sometimes referred to as grating lobes. For multi-dimensional arrays the side lobes can be interpreted as images of the central lobe in analogy to the images of the frequency domain distribution of a function sampled in time. In the case of spatial sampling it can be seen that spatial aliasing is also possible for continuous signals. With this in mind it is essential that the sensor bandwidth be limited to frequencies such that the associated wavelength is less than half the sensor spacing. For non-uniform array geometries the regions of the spectrum that are subject to spatial aliasing become more complicated to describe.

2.3 The Co-array

When the array signals, $y(\boldsymbol{r}_i, t)$, are used to estimate the correlation function of the wave field, $R_f(\boldsymbol{r}, \tau)$, the sensor weighting function plays an important role. With the correlation function of the array data defined as $R_y(\tau) = E[y_{i1}(t)y^*_{i2}(t+\tau)]$ one finds

$$R_y(\tau) = w_{i1}w^*_{i2}R_f(\boldsymbol{r}_{i2} - \boldsymbol{r}_{i1}, \tau) \tag{6}$$

The sum of $w_{i1}w^*_{i2}$ over all indices results in the co-array function for discrete arrays

$$c(\boldsymbol{r}) = \sum_{i1,i2} w_{i1}w^*_{i2}. \tag{7}$$

It can be seen that $c(\boldsymbol{r})$ is the inverse Fourier transform of $|W(\boldsymbol{k})|^2$, the squared magnitude of the array aperture function. A more detailed description of the co-array is given by Johnson and Dudgeon [10] and Van Trees [31].

2.4 Least-Squares Beamforming

A plane wave traversing an array of M sensors with velocity v from azimuth θ arrives at each sensor with a time delay τ_i as discussed above. Using the data one can compute the cross-correlation to estimate the time delay for the signal to appear between two sensors. Taken pairwise there are $M(M-1)/2$ combinations of sensors and for each pair the time of flight between sensors is

dependent upon sensor spacing and the direction of travel of the wave. The $M(M-1)/2$ equations take the form $\boldsymbol{\tau} = X\boldsymbol{f}$ where $\boldsymbol{\tau} = (\tau_{12}, \tau_{13}, \cdots, \tau_{(M-1)M})^T$, $X = (\boldsymbol{r}_{12}, \boldsymbol{r}_{13}, \cdots, \boldsymbol{r}_{(M-1)M})^T$, and $\boldsymbol{f} = (\frac{1}{v}\sin\theta, \frac{1}{v}\cos\theta)^T$. Here, $\tau_{ij} = \tau_i - \tau_j$ and $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$. Since there are $M(M-1)/2$ equations and only two unknowns the system is overdetermined and a least squares estimate of the solution is called for. Assuming there are errors present in the determination of the elements of $\boldsymbol{\tau}$ the relationship above can be written to show the errors explicitly: $\boldsymbol{\tau} = X\boldsymbol{f} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon}$ is a vector of the errors. The least-squares solution to this equation is $\hat{\boldsymbol{f}} = (X^TX)^{-1}X^T\boldsymbol{\tau}$. The solution, $\hat{\boldsymbol{f}}$, can be shown to be unbiased with a variance given by $(X^TX)^{-1}\sigma_\epsilon^2$ where σ_ϵ^2 is the variance of the error vector $\boldsymbol{\epsilon}$.

Finally, the signal velocity and azimuth of arrival can be extracted from $\hat{\boldsymbol{f}}$ as $\hat{v} = 1/|\hat{\boldsymbol{f}}|$ and $\hat{\theta} = \tan^{-1}(\hat{f}_1/\hat{f}_2)$. Approximate expressions for the variances associated with these estimates can be made [19]; however, the full expression depends upon array geometry [29]. The Cramer–Rao bounds on azimuth and slowness estimates have been investigated for varying signal-to-noise conditions by Harris [8]. Shumway [25] has also investigated the bounds and included the effects of imperfect signal correlation.

2.5 High-Resolution, Non-linear and Signal-Subspace Methods

It is well known that the mathematical description of spectral estimation is identical with that of bearing estimation [30]. The *f-k* beamform obtained in Section 2.1 is equivalent to the ordinary Bartlett spectral estimator and is characterized by a broad beam pattern. Several high-resolution spectral estimation methods have been applied to the problem of bearing estimation. These include the maximum-liklihood method introduced by Capon [4], non-linear methods [13], and the signal subspace methods that include MUSIC [23] and ESPRIT [20] approaches. Johnson [10] shows that some of these methods can lead to biased estimates and that there is sensitivity to residual correlations between signal and noise. Van Trees [31] gives an extensive description of these approaches.

2.6 Near-Source Range and Bearing Estimation

When sources lie close to an array the wavefront has a finite curvature and several methods have been described to use measurements of the wavefront curvature to estimate the location of the source.

Geometric techniques that rely on time-of-arrival information at each sensor are equivalent to those which estimate wavefront curvature. Techniques that do not depend upon estimates of wave amplitude are particularly useful for infrasound work, as individual sensors often exhibit variations in gain due to terrain

and vegetation owing to their emplacement. In these methods the complete set of time delays from the co-array are estimated to give time-of-arrival information at each sensor in vector form as $\boldsymbol{\tau}_c$. Assuming that the array spans the space in which a source is to be localized, a point source will give rise to a unique set of time-of-arrival information. Thus, each point in space is characterized by a unique value of the vector $\boldsymbol{\tau}_r$. It is a straightforward computational task to minimize, in some optimal sense, the distance between the vectors: $||\boldsymbol{\tau}_c - \boldsymbol{\tau}_r||$. The point at $\boldsymbol{r}$ is taken to be the source location. This technique is accurate for high signal-to-noise ratios; however, even in cases of relatively high noise, pre-processing the signal is effective in producing accurate localizations. The technique is generally effective for signal sources lying no more than 10–15 array apertures from the center of the array.

A second class of range and bearing estimators have been introduced using signal subspace concepts such as those employed in the MUSIC algorithm [9]. There, the covariance matrix is formed from array data and then decomposed into its eigenvectors and eigenvalues. When signals are contained in the data their presence will be signaled by eigenvalues of the covariance matrix that are significantly larger than the noise covariance. The eigenvectors associated with the non-signal, or noise-like, variances are then assembled into a matrix, U, that defines the noise subspace power. Signal vectors, $\boldsymbol{s}(\boldsymbol{r}, \boldsymbol{\theta})$ are, by definition, orthogonal to this subspace. The psuedo power is calculated as the reciprocal of the projected power as

$$P_M = \frac{1}{\boldsymbol{s}^{\dagger}\mathrm{U}\mathrm{U}^{\dagger}\boldsymbol{s}}. \tag{8}$$

Typically these techniques give results that have high resolution for nearby, continuous sources.

Ultimately the ability to estimate the range and bearing of a signal source depends upon the accuracy with which the time delays are determined that represent the passage of the signal across the array. When the cross-correlation function is used to estimate time delays it can be shown that the quality of the estimate degrades as the signal-to-noise level decreases below unity. In addition, the accuracy of the time delay is somewhat limited by the sample rate for digital data. The Cramer–Rao lower bound on the estimation of time lags determined over M sensors using all pairs can be shown to be [28,12]

$$\sigma_{\tau}^2 = \frac{T}{4\sqrt{M(M-1)/2}f_B S_{NR}} \tag{9}$$

where T is the duration of the signal, M is the number of sensors, f_B is the signal bandwidth and S_{NR} is the signal-to-noise ratio. It is clear from this equation that for a given signal, enhancing the signal-to-noise ratio is important. One technique used to enhance the signal-to-noise ratio is the application of an Eckart filter [28]. We will suggest another technique based upon Samson's pure-state filter in Section 3.3.

3 • Signal Detection and Enhancement

3.1 *F*-Statistic Detection

Fisher's f-statistic arises in the comparisons of sets of means in the analysis of variance. When samples are drawn from several processes the F-statistic is used to test the null hypothesis that the sample variances are equal. The test is defined as a ratio of variances and it is this ratio that is characterized by the F-distribution function. The ratio of variances can also be shown to be the signal-to-noise power estimate when applied to time series.

3.1.1 Time-Domain Application

Following Spiegel [26], consider an array of values obtained from M sensors of N time intervals: $x_{ij}, i = 1, M; j = 1, N$. The observations can be arranged in a matrix where each row represents the output from one sensor and each column represents the M sensor outputs at one moment in time. Then let $\bar{x}$ represent the sample mean over all MN samples and let $\bar{x}_j$ be the mean over each row. The total variation then is $V = \sum_{i,j}(x_{ij} - \bar{x})$ and can be shown to be composed of two portions: the variation over sensors (rows) and the variations between sensors (columns). That is, $V = V_W + V_B$ where$V_W = \sum_{i,j}(x_{ij} - \bar{x}_j)^2$ and $V_B = \sum_{i,j}(\bar{x}_j - \bar{x})^2$.

If we assume that the $x_{ij} = m + \epsilon_{ij}$ where m is the mean and ϵ_{ij} is normally distributed noise with variance σ_n^2 then we can test the hypothesis that all sensor signals have the same mean and variation. To do this we use the fact that $E[V_W] = M(N-1)\sigma_n^2$ and $E[V_B] = (M-1)\sigma_n^2 + N\sum_i \alpha_i^2$ where $\alpha_i = m - m_j$, the difference between the mean of the entire sample to the mean of each sensor output. Thus if the null hypothesis is true then the $\alpha_i = 0$ and $\hat{S}_B^2 = E[V_B]/(M-1) = \sigma_n^2$ and $\hat{S}_W^2 = E[V_W]/M(N-1) = \sigma_n^2$. Then V_B/σ_n^2 and V_W/σ_n^2 are χ^2 distributed with $M-1$ and $N(M-1)$ degrees of freedom, respectively. Note that if the null hypothesis is not true then $\hat{S}_B^2 > \sigma_n^2$ while $\hat{S}_W^2 = \sigma_n^2$. The F-test measures the ratio $\hat{S}_B^2/\hat{S}_W^2$. If the ratio becomes large then the hypothesis can be rejected. The F-distribution characterizes the ratio and is characterized by $M-1$ and $N(M-1)$ degrees of freedom.

Melton[14] first proposed the use of the F-statistic as an appropriate signal detector. In signal processing terms the ratio $\hat{S}_B^2/\hat{S}_W^2$

is just the signal-to-noise power ratio, P_{SNR}. Melton shows further that $P_{SNR} = (F-1)/M$ where F is the F ratio. Once a threshold is set for F, say 90%, this relationship gives the signal-to-noise power required for significant detection. Subsequently, several authors have shown the effectiveness of the F-statistic as an event detector in array data [24,27,1,7].

3.1.2 Frequency Domain Application

Shumway [24] showed that a measure of the Fisher statistic can be put in the form

$$F_f(\omega, \boldsymbol{k}) = (M-1)\frac{P(\omega, \boldsymbol{k})}{P(\omega) - P(\omega, \boldsymbol{k})} \tag{10}$$

where $P(\omega, \boldsymbol{k})$ is the power at $(\omega, \boldsymbol{k})$ and $P(\omega)$ is the total power at ω. This is the ratio of the power at a point in ω-k space to the difference between that power and the total power at that frequency. Smart and Flinn [27] point out that the use of the Fisher statistic for signal detection allows one to distinguish between energy arriving through the main array lobe from energy arriving through the side lobes.

In the search for signals impinging on an array the data are formed into a beam through particular choices of time delays applied to each signal and then the F-statistic is computed on the lagged array data. In this way, at a particular frequency, the entire velocity-azimuth space can be swept for signals. While this is an effective method for surveying the data for incoming signals it is computationally intensive since the number of beams required to cover the velocity-azimuth space may be large.

3.2 PMCC Detection

The Progressive Multi-Channel Correlation Method (PMCC) was originally designed as a seismic event detector applied to data from arrays of seismic instruments [3]. It has been developed and applied to infrasound analysis as a real-time detector and analysis tool [11]. The method is based upon sequential analysis of windowed signals from an array. Cross-correlation estimates are computed over subsets of three sensors from the larger array. Time delays between signal arrivals at sensors are computed from the cross-correlations and a closure relationship is developed in the form $\tau_{ijk} = \Delta t_{ij} + \Delta t_{jk} + \Delta t_{ki}$ where Δt_{ij} is the time delay between sensors i and j. The analysis begins with records from the smallest groups of sensors and progresses to larger-scale groups. In this way, by using data from each triad for which the closure relationship holds, increasingly accurate estimates can be made of the signal azimuth and velocity. The method is applied across frequency bands and then across time with a moving window.

In order to mitigate against false alarms a consistency parameter is used to evaluate the number of sensors participating in a detection. Clusters of points in the frequency–azimuth–velocity space are judged to be associated through a weighting scheme. Thus, in addition to the sizing of the initial data window appropriately for the signals present, one must set a threshold value for consistency and a weighting parameter for cluster identification. An interactive software package, WinPMCC, has been developed that implements the PMCC method in a user-friendly style [11].

3.3 Pure State Filtering and Signal Enhancement

The pure-state filter described here is based upon a frequency-domain estimator developed by J. Samson [21,22]. The estimator is constructed so as to give a measure of signal to noise at each frequency estimate called P^2. By construction, $0 \leq P^2 \leq 1$ with $P^2 = 0$ defining uncorrelated, white noise and $P^2 = 1$ defining pure states of information [5]. The application of the filter to infrasound has been described by Olson [17,18] where it led to an increase in signal-to-noise levels by 10–15 dB.

The pure-state estimator is constructed from the spectral matrix, S, derived from the data themselves and is expressed in terms of the invariants of S. This makes the estimates invariant to the orientation of the microphone array (but not its geometry). The estimator is given by

$$P^2 = \frac{M\mathrm{Tr}[\mathsf{S}^2] - \mathrm{Tr}[\mathsf{S}]^2}{\mathrm{Tr}[\mathsf{S}]^2} \tag{11}$$

where M is the number of sensors in the array and Tr[S] is the trace of S.

The filter is constructed from the data as follows. Given M data sequences of length N we can construct a vector $\boldsymbol{x}(t) = [x_1(t), \ldots, x_M(t)]^T$ where $x_k(t)$ is the time series derived from the kth sensor and $t = n\tau, n = 1, N$ with τ the sample spacing; $\boldsymbol{x}^T$ is the transpose of $\boldsymbol{x}$. The Fourier transform of $\boldsymbol{x}(t)$ is $\boldsymbol{X}(f)$ and the spectral matrix is formed from the outer product of $\boldsymbol{X}(f)$ with its Hermitian conjugate: $\mathsf{S(f)} = <\boldsymbol{X}(f)\boldsymbol{X}^{\dagger}(f)>$. S is an $M \times M$ Hermitian matrix by construction. The angular brackets designate ensemble averaging to reduce variance in the spectral estimates.

With S(f) in hand then $P^2(f)$ can be calculated. Since it is constrained to $0 \leq P^2(f) \leq 1$ it can be applied as a multiplicative filter in the frequency domain to $\boldsymbol{X}(f)$. When inverted to the time domain a filtered time series, $\boldsymbol{x}'$, results for which the signal

content is preserved and the noise-like elements of the spectrum have been suppressed:

$$\boldsymbol{x}'(f) = \mathrm{IDFT}[P^2(f) \cdot \boldsymbol{X}(f)] \tag{12}$$

where IDFT represents the inverse discrete Fourier transform operation.

Acknowledgments The authors would like to acknowledge the useful discussions with many individuals in the Geophysical Institute of the University of Alaska, especially Dr. Charles Wilson and Dr. Ken Arnoult. Support for this effort was provided by the Geophysical Institute, University of Alaska.

References

1. R. R. Blandford, An automatic event detector at the Tonto Forest seismic observatory, Geophysics, **39**, 633, (1974)
2. M. Brandstein and D. Ward (eds.), *Microphone Arrays, Signal Processing Techniques and Applications*, Springer-Verlag, Berlin, (2001)
3. Y. Cansi, An automatic seismic event processing for detection and location: The P.M.C.C. method, Geophys. Res. Lett., **22**, 1021, (1995)
4. J. Capon, Maximum likli-hood spectral estimation, *Nonlinear Methods of Spectral Analysis, vol. 34 of Topics in Applied Physics*, S. Haykin, ed., Springer-Verlag, Berlin, (1983)
5. U. Fano, Description of states in quantum mechanics by density matrix and operator techniques, Rev. Mod. Phys., **29**, 74, (1957).
6. E. A. Flinn (ed.), Special issue on infrasonics and atmospheric acoustics, Geophys. J. R. astr. Soc., **24**, (1971)
7. L. G. Evers and H. W. Haak, Listening to sounds from an exploding meteor and oceanic waves, Geophys. Res. Lett., **28**, 41, (2001)
8. D. B. Harris, Comparison of the direction estimation and performance of high-frequency seismic arrays and three-component stations, Bull. Seis. Soc. Am., **80**, 1951, (1990).
9. Y.-D. Huang and M. Barkat, Near-field multiple source localization by passive sensor array, IEEE Trans. Ant. Prop., **39**, 968, (1991).
10. D. H. Johnson and D. E. Dudgeon, *Array Signal Processing*, Prentice Hall, New Jersey, (1993).
11. A. Le Pichon, and Y. Cansi, PMCC for infrasound data processing, Inframatics Newsletter (http://www.inframatics.org), No. 2, 1, June (2003)
12. N. Levanon, *Radar Principles*, John Wiley & Sons, New York, (1988)
13. Application of the maximum-liklihood method and the maximum-entropy method to array processing, *Nonlinear Methods of Spectral Analysis, vol. 34 of Topics in Applied Physics*, S. Haykin, (ed.), Springer-Verlag, Berlin, (1983)
14. B. S. Melton and L. F. Bailey, Multiple signal correlators, Geophysics, **22**, 565, (1957)

15. R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*, Wiley-Interscience, New York, (1980)
16. P. S. Naidu, *Sensor Array Signal Processing*, CRC Press, New York, (2001)
17. J. V. Olson, Noise suppression using data-adaptive polarization filters: Applications to infrasonic array data, J. Acoust. Soc. Am., **72**, 1456, 1982.
18. J. V. Olson, The application of the pure-state filter to infrasound array data, Inframatics Newsletter (http://www.inframatics.org), No. 7, September (2004)
19. J. V. Olson, Least-squares estimation of azimuth and velocity of plane waves using infrasound array data, Geophysical Institute Report, (1996)
20. R. H. Roy and T. Kailath, ESPRIT-estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust., Speech and Sig. Proc., **37**, 995, (1989)
21. J. C. Samson, Descriptions of the polarization states of vector processes: Applications to ULF magnetic fields, Geophys. J. R. astr. Soc., **34**, 403, (1973)
22. J. C. Samson and J. V. Olson, Generalized Stokes vectors and generalized power spectra for second-order stationary vector processes, SIAM J. Appl. Math., **40**, 137, (1981)
23. R. O. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas and Propagat., **AP-34**, 276, (1986)
24. R. H. Shumway, On detecting a signal in N stationarily correlated noise series, Technometrics, **13**, 499, (1971)
25. R. H. Shumway, S.-E. Kim, and R. R. Blandford, Nonlinear estimation for time series observed on arrays, Ch. 7 in *Asymptotics, Non Parametrics and Time Series*, S. Ghost and M. Puri (eds.), (Marcel Dekker, 1999).
26. M. R. Spiegel and L. J. Stephens, *Statistics*, 3rd edn, McGraw-Hill, New York, (1999)
27. E. Smart and E. A. Flinn, Fast frequency-wavenumber analysis and Fisher signal detection in real-time infrasonic array data processing, Geophys. J. R. astr. Soc., **26**, 279, (1971)
28. D. C. Swanson, *Signal Processing for Intellgent Sensor Systems*, Dekker, New York, (2000)
29. C. A. L. Szuberla and J. V. Olson, Uncertainties associated with parameter estimation in atmospheric infrasound arrays, J. Acoust. Soc. Am., **115**, 253, (2004)
30. C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, New Jersey, (1992)
31. H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, Wiley-Interscience, New York, (2002)

82

Infrasound Propagation

Rodney W. Whitaker[1] *and David E. Norris*[2]

[1]*Los Alamos National Laboratory, PO Box 1663, Los Alamos, NM 87544, USA*
[2]*BBN Technologies, 1300 N. 17th St., Arlington, VA 22209, USA*

1 • Introduction

This chapter provides an overview of issues relevant to infrasound propagation and models that have been employed to understand low-frequency atmospheric wave dynamics and interpret recorded infrasonic signals. Each section is covered at a fairly high level, but references are included to guide the reader who wishes to explore an area in greater detail. A comprehensive literature review has been compiled recently for infrasound in 1997 [1], so the reference list presented here is more representative than comprehensive of the body of research to date. That compilation includes many references to historically important early papers. The chapter concludes with the authors' views on where near-term advances are likely and the outlook for the future.

Since the late nineteenth and early twentieth centuries we have known that low frequency sound could propagate to long ranges, hundreds of kilometers and more. Quite often the sources were military cannons or explosions that were heard far away [2]. With the current knowledge of atmospheric structure and sound propagation, we can understand how sound propagates to such long ranges. Of particular interest here is the regime known as infrasound, sound waves in the frequency range of a few hundredths of a hertz to 20 Hz. At infrasonic frequencies,

especially below a few hertz, attenuation due to absorption can be negligible such that geometric spreading is the dominant loss mechanism [3]. Measurement ranges can extend from several hundred kilometers to thousands of kilometers for a variety of man-made (e.g. mining activity, factories, explosions) and natural (e.g. meteors/bolides, volcanoes, storms) sources.

Infrasound signals are typically measured with four to eight highly sensitive microbaragraphs (sensors) comprising a two-dimensional planar array, where each sensor is fitted with a wind-noise cancellation filter. Measured amplitudes can range from a few hundredths of a pascal (Pa) to tens of pascal. Arrivals from distant sources can be assumed to travel over the array as planar waves. Thus, array processing of the digitally recorded waveforms can provide a variety of outputs, including direction of arrival (back azimuth) of the correlated energy, trace velocity (ground sound speed for horizontal arrivals and infinite for vertical arrivals), and measures of correlation and/or coherence. Two or more arrays recording the same signal can quickly provide a source location estimate through the intersecting back azimuths.

Array processing can be improved with propagation studies and models that provide a basis for a more complete understanding of the observed signals: from observed amplitudes to travel times to azimuth deviations to predications of propagation for future experiments. This is especially true for sound propagation in a dynamic medium that can move with velocities that are a significant fraction of the sound speed and have significant velocity gradients on length scales comparable with the wavelengths of interest.

Infrasound research and development (R&D) has at times been directed by an interest in monitoring man-made explosions, atmospheric nuclear explosions, in particular from the late 1950s to the mid–1970s. A view of the field in 1971 is available for reference [4]. From the mid–1970s to 1994, the amount of infrasound R&D declined dramatically, although a few organizations maintained small, but active, programs and expertise. Renewed efforts began in 1996, driven by the recently negotiated international Comprehensive Nuclear-Test-Ban Treaty (CTBT). Part of the verification regime of the CTBT includes a global network of 60 infrasound stations to monitor for atmospheric explosions, many of which are now operational. Thus there has been a renewed interest of infrasound R&D in array operations and observations, signal processing and analysis, source characterization, and propagation modeling.

Several excellent books covering acoustics, atmospheric waves and computational methods include [5–9]. A recent review of long-range sound propagation can be found in [10].

2 • Relevant Physics and Issues

Acoustic propagation is a wave phenomenon, and atmospheric acoustic waves propagate via coupled fluctuations in pressure, velocity, and density. The propagation can be classified by comparing the acoustic pressure amplitude to the background ambient pressure. For large amplitudes, nonlinear acoustic theory must be employed to describe the propagation; small-scale fluctuations can be described with linear acoustics theory.

Sound propagating from large impulsive sources, such as explosions, generally makes the transition from being strongly nonlinear in the near field, to weakly nonlinear in a transition region, to linear in the far field. In infrasound, impulsive sources have been modeled by defining the acoustic field at a sufficient range from the source such that linear theory can be applied. For example, good results have been obtained using a Pierce blast wave [11] at 1 km range to model nuclear explosions [12].

Successful infrasonic modeling is composed of two key components: accurate characterization of the relevant environmental variables (e.g. wind velocity, temperature, air composition) and adequate application of acoustic models that capture the relevant fundamental physics of a propagating wave front.

2.1 Fundamental Equations and Propagation Modes

Three fundamental equations describe linear acoustic propagation, where we assume the acoustic pressure, density, and velocity (p', ρ', $\boldsymbol{v}'$) deviate only slightly from the ambient values (p_0, ρ_0, $\boldsymbol{v}_0$). These equations are the linearized Eulers equation, the linearized continuity equation, and the state equation for an adiabatic process. For a homogeneous medium (p_0, ρ_0, and $\boldsymbol{v}_0$ do not vary with position) and quiescent medium (p_0 and ρ_0 do not vary with time and $\boldsymbol{v}_0 = 0$), they take the form

$$\rho_0 \frac{\partial \boldsymbol{v}'}{\partial t} + \nabla p' = 0, \tag{1}$$

$$\frac{\partial \rho'}{\partial t} + \rho_0 \nabla \cdot \boldsymbol{v}' = 0, \tag{2}$$

$$p' = c^2 \rho', \tag{3}$$

where c is the static sound speed, which to first order is proportional to the square root of temperature. Through simple manipulation, the linear wave equation can be derived:

$$\nabla^2 p' - \frac{1}{c^2}\frac{\partial^2 p'}{\partial t^2} = 0. \tag{4}$$

Sound propagation in the Earth's atmosphere is significantly affected by winds, especially for infrasound where the propagation can extend beyond 100 km in altitude. The refractive effects of wind can be included in the wave equation by replacing c with an effective sound speed

$$c_{eff} = c + \boldsymbol{v}\cdot\boldsymbol{n}, \tag{5}$$

where $\boldsymbol{v}$ is the wind vector and $\boldsymbol{n}$ is the ray normal, their dot product being the wind projection along the propagation path.

The presence of wind also means we have a non-stationary medium that the sound propagates through. Sound is carried with the wind as it propagates, so that relative to the fixed reference frame on the ground, there will be translation of the wave's position over time. Moving media can be accounted for by reformulating the fundamental equations and resulting wave equation [13]. These effects are typically negligible for short-range propagation, but for long-range infrasound, horizontal winds perpendicular to the propagation path can reach 50 m/s or greater, resulting in significant ground offsets compared to propagation in a stationary medium. For rays, the effect is quantified by calculating an azimuthal deviation of the received ray vector. Note that refraction in the horizontal plane from horizontal wind gradients can also contribute to azimuthal deviation.

Many of the sources of interest will be at or near the ground. In such cases, energy from the sources can be turned, ducted, in the atmosphere where the sound speed plus wind speed in the direction of propagation exceeds the sound speed on the ground. This may occur in the stratosphere at altitudes of 40–50 km and in the thermosphere at around 100 km. Thus we speak of stratospheric returns and thermospheric returns [14]. There may also be ducting in the troposphere (below 10–15 km in altitude) under certain conditions such as when the jet stream is in the direction of propagation. For some sources, a surface Lamb wave may be generated. These waves are coupled to the ground in that they propagate at the ground–air interface and exponentially decay with altitude. Large explosions (such as atmospheric nuclear explosions) can generate a clearly defined Lamb wave followed by other arrivals [15].

Low-frequency absorption is negligible except in the thermosphere where the density decreases such that longer atomic mean free paths result in significant thermal loss. In infrasound, absorption is generally confined to the upper turning points of thermospheric paths. The absorption can be quantified using low-frequency, high-altitude models based on specification of temperature, density, and air composition [16].

2.2 Atmospheric Models

Atmospheric characterizations must cover a domain from the ground up to heights of 140 km or greater. Two widely used empirical models–Horizontal Wind Model (HWM) [17] and the Extended Mass Spectrometer - Incoherent Scatter Radar (MSIS) [18,19]–were developed at the National Aeronautics and Space Administration (NASA) and are currently maintained by the Naval Research Laboratory (NRL). HWM provides zonal and meridional wind components, and MSIS provides temperature, density, and atmospheric composition. HWM and MSIS are based upon over 10 years of atmospheric measurements. They have global coverage from the ground into the thermosphere and capture the long-term average seasonal and diurnal trends. Figure 1 illustrates typical temperature, wind, and effective sound speed profiles for both upwind and downwind propagation. It should be emphasized that local wind and temperature profiles for a specific date and time can exhibit non-trivial offsets [20].

In recent years, an effort has been made to improve upon HWM and MSIS by assimilating near-real-time atmospheric data into the lower height regions. These data are obtained from numerical

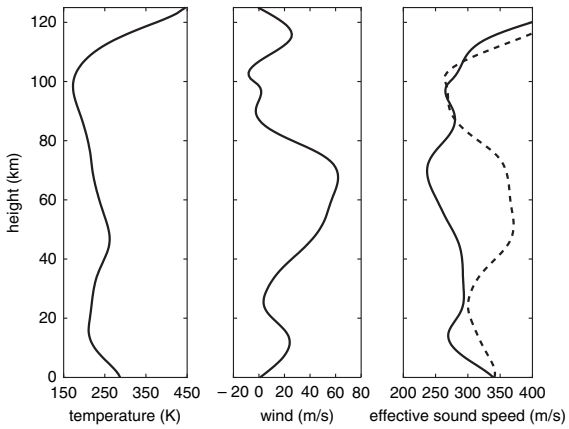


FIGURE 1 Typical winter, mid–latitude profiles of temperature and wind from MSIS and HWM climatological models. Resulting effective sound speed is shown for upwind (*solid*) and downwind (*dashed*) directions.

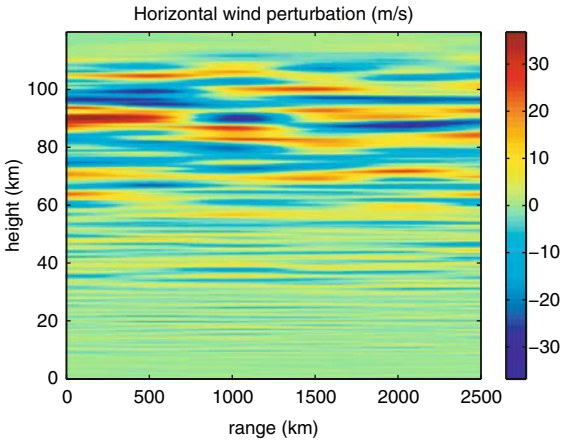


FIGURE 2
Realization of wind perturbations using spectral gravity wave model.

weather models, which use in situ measurements and prognostic models to output global gridded datasets every 4–6 hours. As an example, NRL-G2S uses output from multiple weather models as near-real-time data sources over the first 60 km, which is then merged with HWM and MSIS through the synthesis of new empirical model coefficients [21].

Smaller scale structure in the atmosphere not predicted by climatological or near-real-time models is also relevant, as it can lead to transient ducts, diffraction, and scattering. Atmospheric gravity waves are of interest because their spatial scales are of the same order as infrasonic wavelengths and thus they are suspected of having the greatest influence on propagation variability. Dominant vertical spatial scales are a few kilometers, horizontal scales are hundreds of kilometers, and wind perturbation magnitudes can exceed 50 m/s. Using a spectral model [22], gravity waves have been modeled stochastically [23]. An example realization of wind perturbations associated with gravity waves is given in Figure 2. The gravity wave effects result in thin layers of coherent wind anomalies.

2.3 Atmospheric Variability

Evolution of the atmosphere plays a dominant role in infrasound. Because stratospheric winds are seasonal [24], there are seasonal effects on measured infrasound signals. For example, in mid-northern latitudes, these winds are directed from west to east in the winter and from east to west in the summer. Measured amplitudes for a given source will be larger downwind than upwind (with regard to the stratospheric winds). This was well recorded

in studies of explosion signals [25], and has been observed in early observations of microbarom signals [26]. More recently, year-long observations from signals generated by the Concord aircraft showed systematic variability in azimuthal deviation of 5°. Using ray models, this variability was correlated with seasonal cross winds along the propagation paths [27]. On a global scale, NRL-G2S models have been used to examine the spatiotemporal variability in infrasonic duct formation [21].

Atmospheric variability, both temporally and spatially, is also significant over a signal path. For propagation distances under several hundred kilometers, averaged profiles of temperature and wind can be used in a range-independent atmospheric realization for propagation modeling. For longer range, however, the variability in the atmosphere becomes significant and range-dependent atmospheric realizations will likely improve modeling results [28]. Temporally, winds and temperature fields are not stationary in the hour it takes infrasound to propagate approximately 1000 km. This time evolution of the atmospheric may also be important for consideration in specific scenarios [29].

3 • Propagation Models

Several propagation models have been used for infrasonic modeling. They can be loosely classified under the headings of geometric, continuous wave, and broadband. Applications of time-domain, finite-difference (TDFD) solutions of the full linearized equations have also been made.

When considering infrasound propagation modeling in the dynamic atmosphere, one should maintain realistic expectations of model predictions, even allowing for range dependence. It will be difficult to reproduce every fine scale waveform variation in an observed signal, though agreement in amplitude, duration, and frequency content are more reasonable goals.

In general, geometric models, such as ray theory, are used to visualize the approximate sound path and predict arrival times. Continuous wave models, such as parabolic equation (PE), account for diffraction effects at specified frequency, and they are used to predict amplitude fields. Broadband models, such as normal mode and time-domain parabolic equation (TDPE), are used to predict the full waveform time series, from which arrival times and amplitudes can be identified. They are more applicable for longer-range propagation, since errors in geometric approaches can accumulate over range.

3.1 Ray

Ray theory, or geometric acoustics, is based on representing propagation with ray paths. The derivation is based on applying a high-frequency limit to the wave equation, producing the eikonal and transport equations. In a stationary medium these equations are respectively,

$$|\nabla\tau(\boldsymbol{x})|^2 = \frac{1}{c^2(\boldsymbol{x})}, \tag{6}$$

and

$$2\nabla\tau\cdot\nabla A_0 + (\nabla^2\tau)A_0 = 0. \tag{7}$$

Here, $\boldsymbol{x}$ is an arbitrary coordinate system, c the local sound speed, and A_0 the amplitude. The *eikonal*, $\tau(\boldsymbol{x})$, defines where all the wavefront peaks arrive at position $\boldsymbol{x}$ at time $t = \tau(\boldsymbol{x})$.

These equations are valid under two assumptions: the amplitude varies slightly over distances comparable to a wavelength and the wavefront radii of curvature is substantially larger than a wavelength. For a given coordinate system and environment, these equations can be straightforwardly solved using standard numerical techniques [5,8]. The initial conditions for a ray solution are the initial ray launch angle and amplitude. If a solution is sought for rays between a source and receiver location, then an eigenray can be found by searching over a fan of launched rays that met the receiver boundary conditions.

Ray models are very useful in visualizing the path the sound takes in a refractive environment, and the travel time along this path can be easily computed. Because the solution of ray equations is computationally efficient, they can be applied to three-dimensional, range-dependent, non-stationary realization of the atmosphere. Thus, horizontal wind effects such as azimuthal deviation can be predicted. Amplitude predictions based upon the transport equation have seen fewer applications. At caustics, regions where adjacent rays cross, a singularity in the solution exists. In addition, zero amplitude is predicted in ray shadow zones, since diffraction and scattering effects are not modeled.

In infrasound, a three-dimensional, range-dependent, moving media Hamiltonian ray solution, HARPA, has enjoyed widespread application [30]. As an example, a fan of rays from HARPA are given in Figure 3 for both downwind and upwind conditions.

Because of their computational efficiency, ray models have been used to create travel time tables for use in infrasound localization. Seasonal travel time tables have been precomputed at

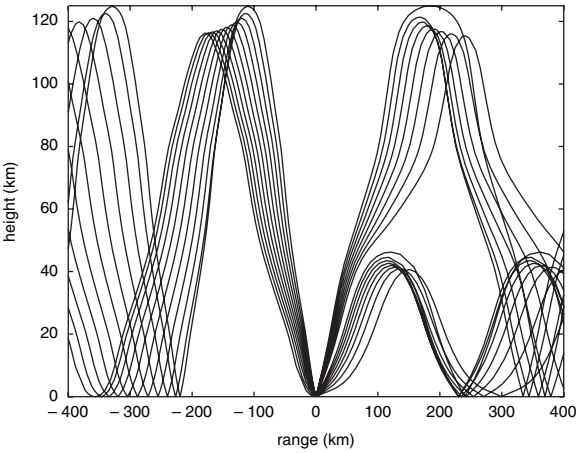


FIGURE 3 A fan of rays launched both upwind and downwind illustrating the dominant propagation ducts for upwind and downwind scenarios in Section 2.2.

International Monitoring Station (IMS) locations over a grid of azimuthal directions for use in source localization [31] and evaluation of atmospheric characterizations. Ray-based travel times have also been computed to evaluate the network performance of the IMS [32].

A variation from the standard ray model is the Gaussian beam approach, which mitigates the problem of amplitude predictions at ray caustics and shadow zones. This model is based upon associating with each ray a Gaussian beam intensity normal to the ray path [33,34]. The Gaussian beamwidth and curvature are computed from differential equations that are solved along with the ray equations. A simplified variation on this approach, Gaussian ray bundles [35], has been applied to modeling recorded waveforms from rocket launches [36].

Tau-p is another ray-based model that has been applied to infrasonic modeling. It was developed in seismology [37] and reformulated in the atmosphere where wind effects are accounted for [38]. Assuming a stratified atmosphere, the time a ray spends traveling through a vertical layer is expressed as a function of a ray parameter that is inversely proportional to phase velocity. The Tau-p model enables fast computation of ray travel time, range, and transverse offset as driven by horizontal winds normal to the propagation path. Tau-p travel times have been used in localizing infrasonic events such as the 2001 Pacific bolide [39].

3.2 Continuous-Wave Parabolic Equation

PE models have enjoyed widespread use in modeling outdoor sound propagation, including long-range infrasound. Continuous-wave PE models are used to predict transmission loss at a single frequency, and they are generally applied over a range-dependent, vertical slice of the atmosphere. There exist numerous derivations and implementation of PE models. For detailed derivations, the reader is referred to refs. [8,9,40]. Here we summarize the foundations of the split-step Fourier (SSF) PE derivation.

We start with the Helmholtz equation in cylindrical coordinates for pressure $p(r,z)$ as a harmonic point source of frequency ω

$$\frac{\partial^2 p(r,z)}{\partial r^2} + \frac{1}{r}\frac{\partial p(r,z)}{\partial r} + \frac{\partial^2 p(r,z)}{\partial z^2} + k_0^2 n(r,z)^2 p(r,z) = 0, \qquad (8)$$

where $k_0 = \omega/c_0$ is the reference wavenumber, $n(r,z) = c_0/c(r,z)$ the index of refraction, and c_0 the reference sound speed. Implicit in this equation is the assumption of uncoupled azimuth, which is valid when variation in the horizontal plane is small.

We express the pressure as an envelope function, $\psi(r,z)$, multiplied by the Hankel function. Using the far-field approximation ($k_0 r \gg 1$) for the Hankel function and substituting into Eq. (8), we obtain the elliptic wave equation

$$\frac{\partial^2 \psi(r,z)}{\partial r^2} + 2ik_0\frac{\partial \psi(r,z)}{\partial r} + \frac{\partial^2 \psi(r,z)}{\partial z^2} + k_0^2(n^2-1)\psi = 0, \qquad (9)$$

where $p(r,z) \sim (1/\sqrt{r})\psi(r,z)\exp(ik_0 r)$. The wide-angle solution is derived by separating Eq. (9) into outgoing and incoming wave components. Neglecting backscatter and assuming weak range dependence, the one-wave outgoing wave equation takes the form

$$\frac{\partial \psi(r,z)}{\partial r} = ik_0(Q-1)\psi(r,z), \qquad (10)$$

where

$$Q = \sqrt{n^2 + \frac{1}{k_0^2}\frac{\partial^2}{\partial z^2}}. \qquad (11)$$

This first-order partial differential equation in r can be solved as an initial-condition problem, where the solution over z is stepped forward in range:

$$\psi(r+\Delta r, z) = \exp(ik_0(Q-1))\psi(r,z). \qquad (12)$$

Different PE implementation are distinguished by how they implement an approximation for the operator Q and numerically solve the resulting equation. The validity of the approximation is measured by an inherent phase error in the $\psi(r, z)$ prediction, a phase error that increases with propagation elevation angle.

Early PE application to infrasound focused on transforming existing underwater models [41]. Since that time, SSF approximations have gained the widest acceptance due to their numerical efficiency. Whereas in ocean applications complex boundary conditions (elastic or multi-layer) necessitate more numerically sophisticated approaches, infrasound boundary conditions can be expressed simply and thus the SSF approach applied. It should also be noted that absorption models can easily be integrated into the SSF PE [42].

To solve for the field at range $r + \Delta r$, $\psi(r, z)$ is first transformed to vertical wavenumber space, k_z, using an Fast Fourier Transform (FFT) algorithm. It is then multiplied by an operator that accounts for diffraction effects. The result is inverse Fourier transformed and multiplied by a second operation that accounts for refraction effects [8].

PE predictions can provide accurate prediction of attenuation due to refraction and diffraction, including regions within shadow zones and near the ground. This predictive capability is gained at the cost of higher computational loads that increase with source frequency. Figure 4 gives a PE prediction at 2.0 Hz for the sample scenario given in Section 2.2, using a wide angle SSF approximation [43].

It is instructive to compare the PE solution with the ray solution shown in Figure 3. The ray density matches well with the higher-energy regions in the PE plot. Note, however, the spread of energy that extends beyond the rays due to diffraction.

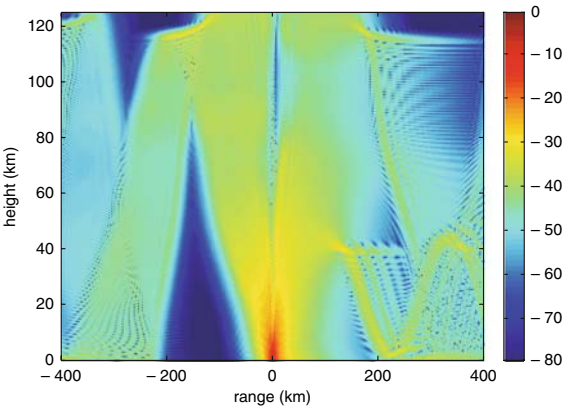


FIGURE 4 PE transmission loss predictions (in decibel) at 2 Hz illustrating diffraction effects not modeled by ray theory, using upwind and downwind scenarios in Section 2.2.

In addition, the first bound range for the stratospheric energy matches well, but there is a mismatch in the thermospheric energy ranges. This difference may in part be due to inherent phase errors in the PE model, which has elevation angle accuracy up to about 35°.

3.3 Normal Mode

The atmosphere has temperature-defined and wind-defined ducts in the thermosphere and stratosphere at altitudes of 100–120 km and about 50 km, respectively. These have been known for some time, and energy refracted in these regions is known as thermospheric or stratospheric returns or signals. Normal mode techniques provide a natural computational approach for this ducted propagation.

Continuous-wave normal mode models are generally derived from the Helmholtz equation in cylindrical coordinates in a range-independent environment. Separation of variables is used to express the solution as a sum of eigenfunctions in r and z

$$p(r,z) = \sum \Phi_m(r)\Psi_m(z). \tag{13}$$

The solution for $\Phi_m(r)$ is a Hankel function, and the asymtopic form can be substituted as an approximation in the far field, giving

$$p(r,z) \sim \frac{1}{\sqrt{r}} \sum \Psi_m(z_s)\Psi_m(z)\frac{\exp(ik_{rm}r)}{\sqrt{k_{rm}}}, \tag{14}$$

where z_s is the source height.

The solution for $\Psi_m(z)$ is a classical Sturm-Liouville eigenvalue problem. The eigenvalues are discrete horizontal wavenumbers, k_{rm}, at modes $m = 1, 2, 3, \ldots$. The dispersion equation relates k_{rm} at each mode to angular frequency ω. Its solution is typically expressed in a dispersion curve, which plots ω versus k_{rm} or ω versus the phase velocity $v_m = \omega/k_{rm}$. Numerical solutions center around solving the transcendental characteristic equation for the eigenvalues k_{rm} [8].

The normal mode equations can be simplified through approximations to the eigenfunction equation. The high-frequency Wentzel, Kramers, Brillon (WKB) approximation has been used extensively in the infrasound community. Geometrically, it can be interpreted as expressing each mode as the sum of upward and downward traveling plane waves whose path can be described with rays normal to the wavefront. The dispersion equation takes the form

$$k_{zm}(z) = \sqrt{\frac{\omega^2}{c^2(z)} - k_{rm}^2}, \tag{15}$$

where k_{zm} is the vertical wavenumber of mode m. k_{zm} goes to zero at the ray turning points, beyond which it becomes imaginary and the mode decays exponentially. The WKB approximation is valid if k_{zm} changes slowly over the distance of a vertical wavelength.

For realistic atmospheres, Pierce and his students developed a normal mode code in which winds could easily be incorporated [44,12,11]. Their formulation was explicitly for an explosion source modeled as an ideal blast wave. The atmosphere is modeled as many horizontal, isothermal layers, each of which may have a different temperature and horizontal wind velocity. Adopting Cartesian coordinates, they Fourier analyzed the linearized equations of continuity and motion in the horizontal and time, and solved the resulting system, of two coupled ordinary differential equations, for the vertical dependencies of the density and velocity components of each mode. The two dependent variables, called potentials, were related to pressure and velocity and were continuous across atmospheric layer boundaries. The modes were defined as functions of phase velocity and angular frequency by application of the bottom boundary condition. The code was successful in predicting the major features in observed signals from atmospheric nuclear explosions [45,12]. An earlier insightful discussion of normal modes was given by Haskell [46]. An excellent separate presentation of the Pierce approach can be found in Chapter 9.

Figure 5 shows an example calculation with a version of the Pierce normal mode code illustrating the waveform synthesis for a given range and time interval. Here we show pressure as a

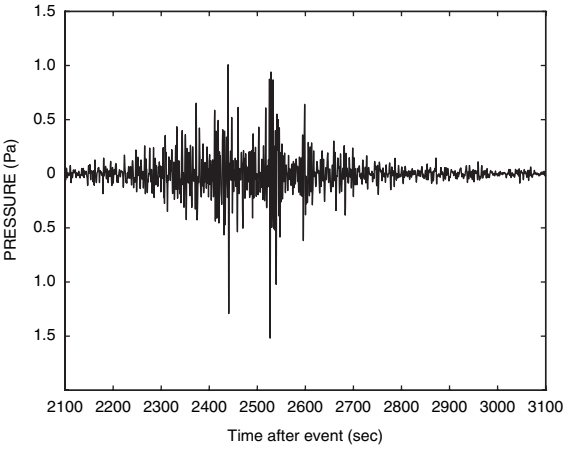


FIGURE 5 Sample normal mode waveform.

function of time after the event at a range of 750 km from a few thousand tons of explosive. Measured wind data were used in the calculation.

3.4 Time-Domain Parabolic Equation

The TDPE model is more flexible than the normal mode model, but at the price of increased computational loading. It can accommodate range-dependent atmospheres, absorption models, and realistic boundary conditions.

There are two major categories of TDPE solution: Fourier synthesis and finite difference. Fourier synthesis involves running a continuous wave model (such as PE) at each frequency bin that defines the source waveform. An inverse Fourier transform of the synthesized spectrum is then computed to derive the received waveform [43,47]. The main advantage of Fourier synthesis is that all capabilities of the continuous wave model are maintained. The main disadvantage is the high computational loading, especially for high bandwidth sources and for acoustic arrivals that are widely dispersed in time. In the latter case, the frequency step must be made small enough that spatial aliasing of the acoustic energy will not occur. The second approach uses finite-difference methods to solve the propagation equations directly in the time domain [48]. Finite-difference approaches potentially offer increased computational efficiency, but the stability constraints on the numerical implementation must be carefully enforced.

Figure 6 gives an example of a Fourier synthesis TDPE prediction for an impulsive source over a 0–2 Hz bandwidth. The waveform solution is computed at a 700 km range over all heights. Plots of waveform amplitude versus time can be obtained by taking an image slice at the height of interest. For the upwind case, only thermospheric arrivals are observed with

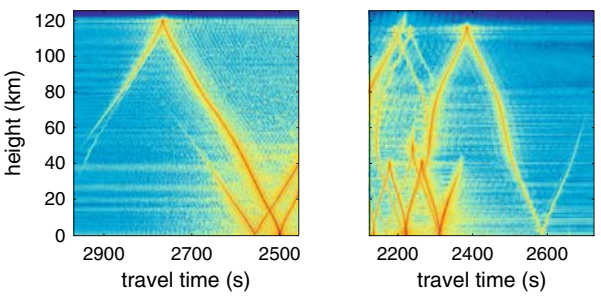


FIGURE 6 TDPE waveform prediction for upwind (*left*) and downwind (*right*) scenarios in Section 2.2. Source is 1 kT Pierce blast wave with 0–2 Hz bandwidth.

group velocities (horizontal range divided by travel time) in the range of 0.27 km/s. Downwind, both stratospheric and thermosphere arrivals are present. The stratospheric arrivals can be identified by their faster group velocity (in range of 0.31 km/s), greater travel time change with height, and disappearance at the upper duct boundary near 40 km.

3.5 Acousto-gravity Waves

Our discussion will only touch on acoustic-gravity waves, which are covered in detail in [49,7,6]. They occur at lower frequencies than infrasound and encompass wavelengths of sufficient size that the basic equations must include gravity. In these waves, wave motion occurs both in the direction of propagation and transverse to it. Some authors refer to these waves as buoyancy waves.

Acousto-gravity waves have been modeled using both PE [50] and normal mode approaches [12]. Following Beer [6] for the isothermal windless atmosphere, the dispersion relation may be written as

$$\omega^4 - \omega^2 c^2 (k_x^2 + k_z^2) + (\gamma - 1) g^2 k_x^2 + i\omega^2 \gamma g k_z^2, \tag{16}$$

where ω is the angular frequency, c is the speed of sound, g is the acceleration of gravity, γ is the ratio of specific heats, and k_x, k_y, and k_z are wavenumbers in the cartesian directions. Let H be the atmospheric scale height and define the isothermal acoustic cutoff frequency as $\omega_a = c/(2H)$ and the isothermal Vaisala-Brunt frequency, $\omega_b = ((\gamma - 1)g^2/c^2)^{1/2}$. Then the solutions of the dispersion relation follow the well-known structure of acoustic waves for $\omega > \omega_a$ and gravity waves for $\omega < \omega_b$, [6, Section 2.3]. Our treatment here is concerned with the former case of acoustic waves.

3.6 Empirical Models

Empirical models refer to relations often derived from observations and associated fits to the observations that yield an equation for pressure versus range. If there were no losses and no winds, then we would have the geometric relation $p \alpha R^{-1}$, or geometric spreading. In the ANSI Blast Standard [3] an exponent of -1.2 is suggested as a simple way to account for losses and is based on observed data and calculations. These are the simplest empirical models.

Because of the obvious effect on signal amplitude of stratospheric winds [25], some normalization is necessary and is presented in [51], based on the wind speed at 50 km in the direction of propagation. This empirical normalization for stratospheric returns was derived from observations of stratospheric signals and is called the wind-corrected amplitude (wca).

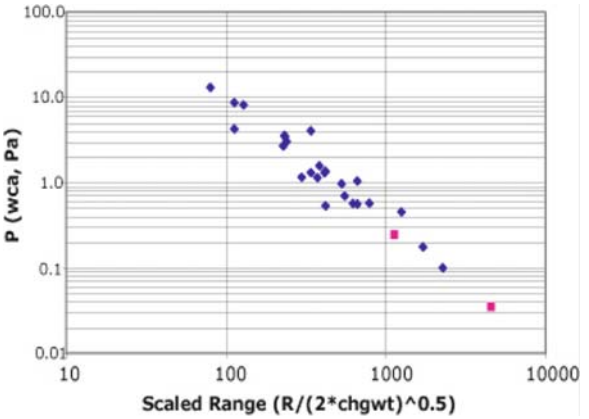


FIGURE 7 Observed high-explosive data. Wind-corrected amplitude vs. scaled range is shown.

A very good relation was found for the wca versus scaled range for a set of explosion data [52,53,54], where scaled range is the range in kilometers divided by the explosion charge weight to the 1/3 or 1/2 power. Figure 7 displays these data, where the regression fit is given by

$$P_{wca} = 5.95 \times 10^4 R_{sc}^{-1.41} \tag{17}$$

with a squared correlation of 0.93. Here P_{wca} is the wca in Pa and R_{sc} is the scaled range using the 1/2 power. The events shown as diamonds are from a series of large-scale high-explosive tests at White Sands Missile Range (WSMR), done under Department of Defense sponsorship. There were eight tests from 1981 to 1993 ranging from 24 tons to 4,880 tons of ammonium nitrate and fuel oil (ANFO). The ranges of infrasonic measurements were from 250 to 5,300 km. The squares are for a 19 ton test, Watusi, at the Nevada Test Site. The data in Figure 7 span almost two orders of magnitude in scaled range.

Studies of other explosion datasets [55,56] show similar relations.

4 • Summary and Future Direction

In this chapter, we have endeavored to give an overview of the issues relevant to infrasonic propagation and many of the atmospheric characterizations and propagation models that have been employed in its study. In this section, we review some of the latest research trends and look at what might be expected in the future.

As touched upon in Section 2.2, current atmospheric modeling fidelity is limited at the higher altitudes of the mesosphere and lower thermosphere (MLT), and for smaller spatial and temporal scales. Atmospheric models coupled with near-real-time updates can now successfully characterize the climatological trends and synoptic scale variability up to approximately 60 km altitude. However, we must still rely on climatological models above this height region. Gravity wave structure that dominates in the mesosphere has been modeled spectrally, but smaller scale variability below the mesosphere has yet to be addressed. One area identified for future research is the height region from 10–20 km, where finer scale structure (such as Kelvin-Helmholtz waves) may direct infrasonic energy out of elevated ducts and down to the ground.

The atmosphere is continually evolving, and we would like to characterize how this variability affects our confidence bounds in propagation predictions. Initial efforts to compute these bounds have used spectral gravity wave models and ray tracing to perform Monte Carlo simulations [29]. These efforts have met limited success due to the over-sensitivity of the rays to perturbations. This area would benefit from future research that applies broadband propagation models, such as the PE or TDPE, to atmospheric realizations that include fine-scale structure.

It was noted that wide-angle PE solution has an inherent phase error that becomes significant for launch elevations above 35°. The PE Padé approximation that uses a fifth-order rational function provides very wide angle prediction with minimal phase error [57]. A split-step version of this approximation also makes it more numerically tractable [58]. Application of this PE and TDPE version in infrasound may improve predictions of thermospheric returns with large upper turning heights.

We have limited our overview to linear acoustic propagation, although nonlinear effects of certain infrasonic source are significant, such as finite amplitude wavefront steepening [59]. Energetic impulse sources or explosions, which are good sources of long-range infrasound signals start out as shock or blast waves that become distant oscillatory infrasound signals. Initial efforts to model the full acoustic evolution, from strong shock, to weak shock, to linear wave have been completed [60]. A key element of this modeling is the use of the nonlinear progressive wave equation (NPE) in the weak shock region [61]. The NPE has been used elsewhere to model shock distortion [62], and as computational capabilities continue to expand, this research path is likely to produce increased understanding of the relevant physical processes in the early stages of wave expansion. As

an example, NPE predictions out to first bounce distances, ~ 200–250 km, could further quantify when propagation region transitions occur.

Computational resources are becoming increasingly more affordable and powerful, opening the possibility to consider calculations that would have been unthinkable just a few years earlier. TDFD solutions of the full set of fundamental equations in three dimensions are still ambitious and probably not achievable soon, but two-dimensional problems out to reasonable distances may be. Even these would require a few hundred million cells but would be an excellent starting point.

For infrasound, some simplifications could be employed, such as assuming a perfectly reflecting bottom boundary, avoiding complicated details of finite impedance. The TDFD would provide waveform results (amplitude versus time) throughout the whole domain and with proper visualization one can follow the evolution of all the signals and the signal progression at a particular point. Range dependence of the environment can be included. These still will be computationally challenging and demanding calculations but now is the time to really begin to start their application and to develop accurate and efficient programs.

Early work in this area is shown in Figure 8, which is output from a simplified TDFD model for an impulse source (Eric Chael, 2003, personnel communication). The calculation used 0.25 km zoning, an ideal reflecting bottom boundary, absorbing boundaries on the top and right, and used no winds or dissipation. Pressure contours are shown at 200, 400, and 600 seconds. Modeling this simplified problem illustrates the potential of such techniques. Another very ambitious TDFD code development project is under way for inhomogeneous moving media, including winds and turbulence; rigid, free surface or partially absorbing

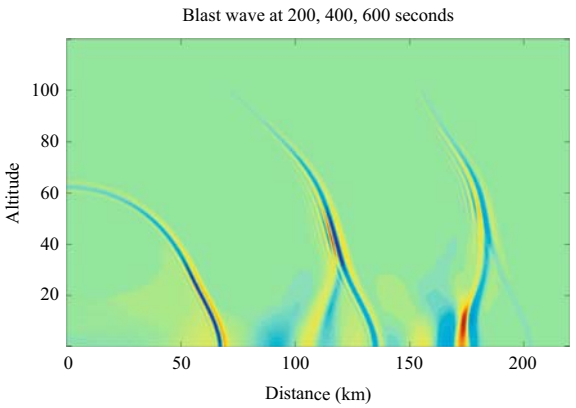


FIGURE 8 Example of simple TDFD wave calculation.

bottom boundaries; high-order difference schemes and a quasi-wavelet turbulence model [63]. The code has been applied to low-frequency propagation out to 500 km range [64]. Our expectation is this and future research projects will continue to push the limits of computational capabilities in capturing as much of the relevant physics as possible in propagating infrasonic waves.

Although not discussed in detail, there exists a rich variety of impulsive and continuous sources of infrasound, both man-made and natural. The propagation models used today are generally formulated for modeling point sources. However, many sources have non-trivial spatial extent, such as microbaroms, or non-zero velocity, such as meteors (bolides). We expect future research effort to begin incorporating these source characteristics into predictions. As an example, the theoretical basis for meteor-generated infrasound has been established [65], and the opportunity exists to leverage this work in modeling recorded signals from such events.

Finally, the operation of the global infrasound network under the CTBT will contribute to our understanding of infrasound propagation by providing a rich set of observations against which propagation models can be tested and applied. This interplay will improve both areas and open unanticipated areas of research.

Acknowledgments This work was supported in part by the US Department of Energy and BBN Technologies. The cooperation of the Springer editors is appreciated.

References

1. J. M. McKisic, Infrasound and the infrasonic monitoring of atmospheric nuclear explosions: A literature review, Technical Report A942933, TRACOR, Rockville, Md., 1997.
2. R. K. Cook, *Strange Sounds in the Atmosphere. Part I*, Sound **1**, 12–16, 1962.
3. American National Standards Institute ANSI S2.20-1983 (R-1989), *Estimating Air Blast Characteristics for Single Point Explosions in Air, with a Guide to Evaluation of Atmospheric Propagation and Effects*, 1983.
4. Special Infrasound Issue, Geophys. J. R. Astr. Soc. **26**, 1971.
5. A. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications*, Acoustical Society of America, Woodbury, 1989.
6. T. Beer, *Atmospheric Waves*, Wiley, New York, 1974.
7. E. Gossard and W. Hooke, *Waves In The Atmosphere: Atmospheric Infrasound and Gravity Waves - their Generation and Propagation*, Elsevier Scientific Publishing Company, New York, 1975.
8. F. Jensen, W. Kuperman, M. Porter, and H. Schmidt, *Computational Ocean Acoustics*, American Inst. Physics, New York, 1994.

9. E. Salomons, *Computational atmospheric acoustics*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.
10. S. Kulichkov, "Long-range propagation of sound in the atmosphere, a review," Izv. Atm. and Ocean Phys. **28**, 253, 1994.
11. A. Pierce, C. Moo, and J. Posey, Generation and propagation of infrasonic waves, Technical Report AFCRL-TR-73-0135, Air Force Cambridge Research Laboratories, Bedford, Mass., 1973.
12. A. Pierce and J. Posey, Theoretical prediction of acoustic-gravity pressure waveforms generated by large explosions in the atmosphere, Technical Report AFCRL-70-0134, Air Force Cambridge Research Laboratories, Bedford, Mass., 1970.
13. V. E. Ostashev, *Acoustics in moving inhomogeneous media*, E & FN Spon, London, 1997.
14. T. Georges and W. Beasley, "Infrasound refraction by upper-atmospheric winds," J. Acoust. Soc. Am. **61**, 28–34, 1977.
15. D. O. ReVelle and S. Kulichkov, On Lamb wave propagation from small surface explosions in the atmospheric boundary layer, Technical Report LA-UR-98-3957, Los Alamos National Laboratory, Los Alamos, New Mex., 1998.
16. L. Sutherland and H. Bass, "Atmospheric absorption in the atmosphere up to 160 km," J. Acoust. Soc. Am. **115**, 1012–1032, 2004.
17. A. Hedin et al., "Empirical wind model for the upper, middle, and lower atmosphere," J. Atmos. Terr. Phys. **58**, 1421–1447, 1996.
18. J. Picone, A. Hedin, D. Drob, and A. Aiken, "NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues," J. Geophys. Res. **107**, 1468 2002.
19. J. Picone, A. Hedin, S. Coffey, J. Lean, D. Drob, H. Neal, D. Melendez-Alvira, R. Meier, and J. Mariska, The Naval Research Laboratory program on empirical models of the neutral upper atmosphere, in *Astrodynamics: Advances in the Astronautical Sciences, Vol. 97*, edited by F. Hoots, B. Kaufman, P. Cefola, and D. Spencer, American Astronautical Society, San Diego, 1997.
20. D. Drob and J. Picone, Statistical performance measures of the HWM-93 and MSISE-90 empirical atmospheric models and the relation to infrasonic CTBT monitoring, in *Proceedings of the 22nd Annual Seismic Research Symposium*, New Orleans, La., 2000.
21. D. Drob, J. Picone, and M. Garcés, "The global morphology of infrasound propagation," J. Geophys. Res. **108**, 4680–4692, 2003.
22. C. Gardner, Scale-independent diffusive filtering theory of gravity wave spectra in the atmosphere, in *The Upper Mesosphere and Lower Thermosphere: A Review of Experiment and Theory, Geophysical Monograph 87*, edited by F. Hoots, B. Kaufman, P. Cefola, and D. Spencer, American Geophysical Union, Washington, D.C., 1995.
23. D. Norris and R. Gibson, InfraMAP enhancements: Environmental/propagation variability and localization accuracy of infrasonic networks, in *Proceedings of the 24th Seismic Research Review*, Ponte Vedra Beach, Fl., 2002.

24. W. L. Webb, *Structure of the Stratosphere and Mesosphere*, Academic Press, New York, 1966.
25. J. Reed, Climatology of airblast propagations from Nevada test site nuclear airbursts, Technical Report SC-RR-69-572, Sandia National Laboratory, Albuquerque, New Mex., 1969.
26. W. Donn and D. Rind, "Natural infrasound as an atmospheric probe," Geophys. J. R. Astr. Soc. **26**, 111–133, 1971.
27. A. LePichon, M. Garcés, E. Blanc, M. Barthélémy, and D. Drob, "Acoustic propagation and atmosphere characteristics derived from infrasonic waves generated by the Concorde," J. Acoust. Soc. Am. **111**, 629–641, 2002.
28. R. Gibson and D. Norris, Infrasound propagation modeling: near-real-time environmental characterization and variability studies, in *Infrasound Technology Workshop*, De Bilt, The Netherlands, 2002.
29. D. Norris and R. Gibson, Integration of enhanced propagation, environmental variability, and network performance models into the InfraMAP software tool kit, Technical Report (Draft), Defense Threat Reduction Agency, Alexandria, Va., 2003.
30. M. Jones, J. Riley, and T. Georges, A versatile three-dimensional Hamiltonian ray-tracing program for acoustic waves in the atmosphere above irregular terrain, Technical report, NOAA Wave Propagation Laboratory, Boulder, Co., 1986.
31. D. Brown, A. Gault, R. Geary, P. Caron, and R. Burlacu, The Pacific infrasound event of April 23, 2001, in *Proceedings of the 23rd Seismic Research Review*, Jackson Hole, Wy., 2001.
32. M. OḂrien and G. Shields, Infrasound source location using time-varying atmospheric models, in *Proceedings of the 26th Seismic Research Review*, Orlando, Fl., 2004.
33. M. Porter and H. Bucker, "Gaussian beam tracing for computing ocean acoustic fields," J. Acoust. Soc. Am. **82**, 1349–1359, 1987.
34. V. Cerveny, M. Popov, and I. Psencik, "Computation of wave fields in inhomogeneous media – Gaussian beam approach," Geophys. J. R. Astr. Soc. **70**, 109–128, 1982.
35. H. Weinberg and R. Keenan, "Gaussian ray bundles for modeling high-frequency propagation loss under shallow-water conditions," J. Acoust. Soc. Am. **100**, 1421–1431, 1996.
36. G. D'Spain, M. Hedlin, J. Orcutt, W. Kuperman, C. deGroot-Hedlin, L. Berger, and G. Rovner, Application of physics-based underwater acoustic signal and array-processing techniques to infrasound source localization, in *Proceedings of the 24th Seismic Research Review*, Ponte Vedra Beach, 2002.
37. R. Buland and C. Chapman, "The computation of seismic travel times," Bull. Seismol. Soc. Amer. **73**, 1271–1302, 1983.
38. M. Garcés, R. Hansen, and K. Lindquist, "Travel times for infrasonic waves propagating in a stratified atmosphere," Geophys. J. Int. **135**, 255–263, 1998.
39. M. Garcés, C. Hetzer, K. Lindquist, R. Hansen, J. Olson, C. Wilson, D. Drob, and M. Hedlin, Infrasonic source localization of the April

23, 2001, bolide event, in *Proceedings of the 23rd Seismic Research Review*, Jackson Hole, Wy., 2001.
40. K. Smith and F. Tappert, The University of Miami parabolic equation model, Technical Report MPL 432, Scripps Institute of Oceanography, San Diego, 1993.
41. G. Norton and W. Kinney, Adaption of an underwater acoustic propagation model to infrasonic atmospheric propagation, in *Proceedings of the 21st Seismic Research Symposium*, Las Vegas, 1999.
42. D. Norris and R. Gibson, InfraMAP propagation modeling enhancements and the study of recent bolide events, in *Proceedings of the 23rd Seismic Research Review*, Jackson Hole, Wy., 2001.
43. F. Tappert, J. Spiesberger, and L. Boden, "New full-wave approximation for ocean acoustic travel time predictions," J. Acoust. Soc. Am. **97**, 2771–2782, 1995.
44. A. Pierce and W. Kinney, Computational techniques for the study of infrasound propagation in the atmosphere, Technical Report AFGL-TR-76-56, Air Force Geophysics Laboratories, Hanscom AFB, Ma., 1976.
45. A. D. Pierce, J. W. Posey, and E. F. Iliff, "Variation of nuclear explosion generated acoustic-gravity wave forms with burst height and with energy yield," J. Geophys. Res. **76**, 5025–5042, 1971.
46. N. Haskell, "Asymptotic approximation for the normal modes in sound channel wave propagation," J. Appl. Phys. **22**, 157–168, 1951.
47. L. Nghiem-Phu and F. Tappert, "Parabolic equation modeling of the effects of ocean currents on sound transmission and reciprocity in the time domain," J. Acoust. Soc. Am. **78**, 642–648, 1985.
48. M. Collins, "The time-domain solution of the wide-angle parabolic equation including the effects of sediment dispersion," J. Acoust. Soc. Am. **84**, 2114–2125, 1988.
49. C. J. Nappo, *An Introduction to Atmospheric Gravity Waves*, Academic Press, San Diego, 2002.
50. J. Lingevitch, M. Collins, and W. Siegmann, "Parabolic equations for gravity and acousto-gravity waves," J. Acoust. Soc. Am. **105**, 3049–3056, 1999.
51. J. Mutschlecner, R. Whitaker, and L. Auer, An empirical study of infrasonic propagation, Technical Report LA-13620-MS, Los Alamos National Laboratory, Los Alamos, New Mex., 1999.
52. R. Whitaker, Infrasonic monitoring, in *Proceedings of the 17th Seismic Research Symposium*, Scottsdale, 1995.
53. M. Davidson and R. W. Whitaker, Miser's gold, Technical Report LA-12074 MS, Los Alamos National Laboratory, Los Alamos, New Mex., 1992.
54. R. Whitaker, T. Sandoval, and J. Mutschlecner, Recent infrasound analysis, in *Proceedings of the 25th Seismic Research Review*, Tucson, AZ, 2003.
55. J. Stevens, D. Adams, G. Baker, H. Xu, J. Murphy, I. Divnov, V. Bourchik, and I. Kitov, Infrasound scaling and attenuation relations from Soviet explosion data and instrument design criteria from

experiments and simulations, in *Proceedings of the 21st Seismic Research Symposium*, Las Vegas, 1999.
56. D. Clauter and R. Blandford, Capability modeling of the proposed international monitoring system 60-station infrasonic network, in *Proceedings of the infrasound workshop for CTBT monitoring*, Santa Fe, 1998.
57. M. Collins, "Applications and time-domain solution of higher-order parabolic equations in underwater acoustics," J. Acoust. Soc. Am. **86**, 1097–1102, 1989.
58. M. Collins, "A split-step Padé solution for the parabolic equation method," J. Acoust. Soc. Am. **93**, 1736–1742, 1993.
59. M. Hamilton and D. Blackstock, editors, *Nonlinear Acoustics*, Academic Press, San Diego, 1998.
60. F. van der Eerden, E. Védy, and E. Salomons, Prediction of shock waves over a large distance: A hybrid method, in *Proceedings of the 11th long range sound propagation symposium*, Fairlee, Vermont, 2004.
61. B. McDonald and G. Orris, High-angle formulation for the nonlinear progressive-wave equation (NPE) model, Technical Report NRL/FR/7140–98-9859, Naval Research Laboratory, Washington, D.C., 1998.
62. A. Piacsek, "Atmospheric turbulence conditions leading to focused and folded sonic boom wave fronts," J. Acoust. Soc. Am. **111**, 520–529, 2002.
63. V. E. Ostashev, D. K. Wilson, L. Liu, D. Aldridge, N. Symons, and D. Marlin, "Equations for finite-difference, time-domain simulation of sound propagation in moving inhomogeneous media and numerical implementation," J. Acoust. Soc. Am. **117**, 503–517, 2005.
64. N. P. Symons, D. F. Aldridge, D. H. Marlin, D. K. Wilson, E. G. Patton, P. P. Sullivan, S. L. Collier, V. E. Ostashev, and D. P. Drob, 3d staggered grid finite-difference simulation of sound refraction and scattering in moving media, in *Proceedings of the 11th International Symposium on Long Range Sound Propagation*, Fairlee, Vermont, 2004.
65. D. ReVelle, "On meteor-generated infrasound," J. Geophys. Res. **81**, 1217–1230, 1976.

PART XV
SEISMOLOGY

Max Deffenbaugh

ExxonMobil Research and Engineering Company, Annandale, NJ, USA

83 Structure and Elastic Properties of the Earth

Rebecca Saltzer 1523

Why Seismology is Studied • Structure and Composition of the Earth • Seismic Waves • Elastic Properties of the Earth • Commercial Seismology

84 Seismic Wave Propagation

Xianyun Wu and Ru-Shan Wu 1535

Introduction • Medium Models • Reflections of Plane Wave

85 Seismic Data Acquisition

Christine E. Krohn 1545

Introduction • The Acquisition System • Signal-to-Noise Quality • Sampling for Migration Imaging

86 Signal Models in Seismic Processing

A.J. Berkhout 1559

Introduction • The Data Matrix • Source, Transfer and Receiver Matrices • Accuracy and Tractability in Seismic Signal Models

87 Deconvolution and Optimal Filtering in Seismology

Ramesh Neelamani 1571

Introduction • Deconvolution with Known Source Signatures • Deconvolution with Unknown Source Signatures • Conclusions

88 Seismic Migration Imaging

Michael Fehler 1585

Introduction • The Migration Problem • Time Migration • Depth Migration

89 Geophysical Parameter Estimation

Max Deffenbaugh 1593

Introduction • Model Selection • Information Assessment • Estimator Selection • Estimator Validation • Conclusion

90 Earthquake Localization

Michael Fehler 1627

Introduction • The Location Problem • Geiger's Method • Grid Search • Methods to Reduce Uncertainty

91 Seismic Tomography

Robert L. Nowack and Cuiping Li 1635

Introduction • Seismic Travel-Time Tomography • Examples of Travel-Time Tomography • Extensions of Travel-Time Tomography

92 Array Signal Processing

Christine Thomas 1655

Introduction • Slowness, Backazimuth and the Complete Slowness Vector • Array Methods • Array Configurations

83

Structure and Elastic Properties of the Earth

Rebecca Saltzer

ExxonMobil Upstream Research Company, Houston, TX, USA

1 • Why Seismology is Studied

Seismology is rooted in both academic curiosity and practical need. Of academic interest are the history, structure and composition of the Earth as well as the processes that create continents and ocean basins. Of commercial importance are ore and mineral deposits, such as diamonds, which tend to be found beneath the oldest and thickest continental regions (cratons), and gas and oil reserves which are found in the shallow crust. Seismic tomography can help illuminate the thick cold areas where diamonds are likely to reside and reflection seismology can do the same for identifying and assessing potential hydrocarbon reserves in the crust. Seismology can also be applied to practical, societal problems such as trying to predict when and where an earthquake might occur or a volcano might erupt and determining whether an explosion in a foreign country is a nuclear test or simply an earthquake or mine blast.

2 • Structure and Composition of the Earth

In the last century, the study of seismic waves has allowed scientists to determine the structure and density of the Earth as well as the physics of earthquake ruptures. From travel time studies, we

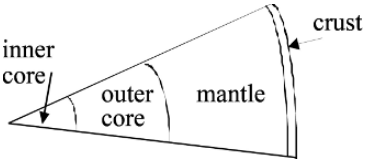


FIGURE 1 Slice of the Earth.

now know that the Earth consists of three major layers: the solid inner core, the liquid outer core and the rocky mantle. Riding on the top of the mantle is the crust, which is ~6 km thick in oceanic regions and generally between 30 and 50 km thick on the continents (Figure 1).

2.1 Inner Core

As the Earth cooled over its history, the denser elements settled toward the center. The core of the Earth is approximately half the radius of the Earth and about twice as dense as the mantle. Consisting of iron-rich and nickel-rich alloys, it represents 32% of the mass of the Earth. The inner, most dense part of the core has a radius of 1,222 km and a density of about 13 g/cm^3. It is solid, despite being hot, because of the extremely high pressures near the center of the Earth (~3290–2640 kbars). There is mounting seismic evidence (e.g., Song and Richards, 1996) that the inner core rotates in an easterly direction relative to the mantle. A flurry of papers have debated this issue in the last decade, but in recent years a consensus has emerged that the inner core has a super-rotation of a few tenths of a degree per year.

2.2 Outer Core

The outer core is compositionally similar to the inner core but with a density of just ~10 g/cm^3. The boundary between the two is defined by the transition of the core material from a solid to a liquid state that occurs ~1,221 km from the center of the Earth at a pressure of 2,640 kbar and a temperature ~4,300 C. Thermal convection, chemical convection, precession, combined with Coriolis forces, cause circulation of the fluid outer core around the solid inner core. This fluid motion in the presence of the Earth's magnetic field induces electrical currents, which in turn sustain the magnetic field, a process called the geodynamo. Of the terrestrial planets, Earth and Mercury have a magnetic field. Mars has a magnetic field in patches that is thought to be a remnant of a previous dynamo that became extinguished when the planet cooled sufficiently for the liquid core to solidify. Venus has none, which is taken as evidence that the planet lacks a solid core.

2.3 Mantle

In contrast to the fluid outer core, the rocky mantle behaves as a solid over the short time scales relevant to seismic wave

propagation. The mantle comprises minerals that are less dense than the core (~3.4 g/cm^3), such as olivine, pyroxene and garnet.

The bottom 200 km of the mantle (where the mantle meets the outer core) is a region called D″ (pronounced "D double-prime"). This odd name derives from an earlier nomenclature, where the layers of the Earth were identified by letters of the alphabet from A to G. The lower mantle (D) was later divided into subregions D' and D". It is unclear whether D" is compositionally distinct from the mantle or simply a portion of the mantle that is strongly affected by heat flow from the core. It has attracted interest from Earth scientists because it exhibits a large amount of seismic scattering and appears to be laterally variable in both velocity and thickness.

Within the mantle there are two major seismic discontinuities. One occurs at ~670 km depth and another at ~410 km depth. Both are believed to be the result of mineralogical phase changes (the atoms rearrange into denser structures) but debate on the cause of both continues. The mantle is often subdivided into an upper and lower half at the 670 km discontinuity. Cooling of the Earth is believed to cause convection within the mantle on very long time scales. Scientists debate whether the upper and lower halves of the mantle convect separately (hence the reason to subdivide them into two halves) or together as one huge convective region.

At ~200 km depth is the elusive Lehmann discontinuity, which was first observed in the early 1960s by Inge Lehmann while analyzing seismograms from Nevada Test Site explosions. Subsequent studies have found conflicting and varied evidence for the presence of the discontinuity suggesting that one exists, but that its presence is sporadic. Further complicating matters, there is no obvious mineralogical change known to occur near the appropriate depth and pressure so the reason for such a discontinuity is unclear. Many intriguing suggestions have been made and it remains an open research topic.

The top of the mantle is defined by the Mohorovicic discontinuity (or Moho for short). First observed in 1909, it represents the boundary between the crust and the mantle. The discontinuity occurs because there is a compositional change between the crust and the denser mineral assemblage of the mantle.

The crust and uppermost mantle together form a rigid, high–velocity zone called the lithosphere. At ~150 km depth below continents and ~50 km depth below oceanic regions is a plastic low-velocity zone within the mantle called the asthenosphere. This is thought to be the boundary layer between the near surface

where heat is transported by conduction and the deep interior where heat is transported by convection.

2.4 Crust

The Earth's crust tends to consist of light, alumino-silicate minerals such as quartz, feldspar and amphiboles with densities of $\sim 2.8\,g/cm^3$. It represents just 0.4% of the Earth's mass. The oceanic crust is generally 5–15 km thick whereas the continental crust is 30–50 km thick. Most of the continental regions of the Earth have permanent seismometers recording continuously the hundreds of earthquakes that occur daily. From studies of these earthquakes, we know that earthquakes tend to occur along well-defined belts of seismicity at the edges of tectonic plates. These plates are spreading apart in some regions, such as the Mid-Atlantic Ridge, where a new crust is being formed and are sinking down in other regions, such as the Western Pacific Rim in what are called subduction zones. Along the western coast of the US, the Pacific and North American plates are sliding past one another, which gives rise to a large number of earthquakes in California.

Additional details about the composition and rheology of the Earth can be found in Anderson (1989), Stacey (1992) and Turcotte and Schubert (1982).

3 • Seismic Waves

Energy sources such as earthquake faulting or explosions cause elastic deformations (strains) and internal forces in the rock (stresses). Except in the immediate vicinity of the source, the stresses and strains are linearly related to one another. Hooke's law expresses this relationship by a tensor of elastic moduli.

If the medium is isotropic, the tensor relating the stresses and strains contains just two independent elastic moduli (λ and μ) called the Lamé constants, greatly simplifying the wave equation. The solution to the wave equation for waves traveling through the Earth (body waves) yields two independent wave types: compressional waves (also called "primary" or P waves) and shear waves (also called "secondary" or S waves). The P waves travel faster than the S waves and have particle motion only in the direction of wave propagation. The S waves have particle motion perpendicular to the direction of propagation in either a vertical or horizontal plane. Basic seismology texts such as Lay and Terry (1995) or Shearer (1999) give a clear and comprehensive derivation of these equations.

There are additional solutions to the wave equation that occur at the surface of the Earth where the surface tractions

vanish. Because the boundary conditions are different at the free surface, the horizontally polarized S waves (SH) will constructively interfere to produce a wave that travels along the surface called a Love wave. P waves and vertically polarized S waves (SV) will also get trapped at the surface to form a wave with retrograde particle motion that is called a Rayleigh wave. Surface waves are generally of much higher amplitude than body waves, travel more slowly, show less amplitude-decay with distance (since they spread out over a 2D surface rather than a 3D volume), and are dispersive (i.e., the velocities are frequency-dependent). They tend to be generated by shallower earthquakes and often are also far more destructive.

Because surface waves travel along great-circle paths on the Earth's surface, they are often treated mathematically with a spherical coordinate system. The solutions to the equation occur at discrete (resonant) frequencies that are similar to standing waves called normal modes. For the Earth, there are two distinct modes: spheroidal, which are analogous to P/SV and Rayleigh wave motion, and toroidal, which are analogous to SH and Love wave motion. Toroidal modes involve no radial motion and are sensitive only to the shear velocity, whereas spheroidal modes have both radial and horizontal motion and are sensitive to both compressional and shear velocities. Spheroidal modes at very low frequencies are also sensitive to gravity and provide the best direct seismic constraints on the Earth's density structure.

The lateral variations in normal mode eigen-functions are described in terms of spherical harmonics, which provide an orthogonal set of basis functions on a spherical surface. Computing synthetic seismograms by summing normal modes is standard practice in surface-wave and long-period body-wave seismology. Dahlen and Tromp (1998) provide a comprehensive theoretical discussion of this subject.

3.1 Identification of Seismic Phases

The path that a seismic wave follows through the Earth and the wave type for each leg of the path are denoted by a sequence of letters. Compressional wave legs in the mantle, outer core, and inner core are identified by uppercase P, K and I, respectively. Shear wave legs in the mantle and inner core are identified by uppercase S and J respectively. Lowercase p and s indicate compressional and shear waves that leave the focus in an upward direction. Lowercase letters m, h, i indicate reflections at the Mohorovicic discontinuity, outer core, and inner core respectively. So, for example, pPKP indicates a compressional wave that initially travels upward, reflects off the surface of the Earth,

then passes through the mantle, outer core and mantle again without reflecting, while SKiKP indicates an arrival that begins as a shear wave traveling downward through the mantle, converts to a compressional wave at the outer core boundary, reflects off the inner core and remains a compressional wave as it travels back up through the outer core and mantle. LQ and LR are used to identify Love and Rayleigh waves. Free oscillations of the Earth are identified by ${}_nS_m$ for Spheroidal modes and ${}_nT_m$ for torsional modes, where the subscript m gives the angular order and n the overtone degree (number of radial node surfaces).

4 • Elastic Properties of the Earth

4.1 Velocity

The inner core has a P wave velocity of about 11 km/s and an S wave velocity of about 3.6 km/s. The outer core has a P wave velocity of about 8 km/s, and cannot sustain S waves because it is a fluid. Velocities in the mantle are greater than in the outer core, ranging from 8 to 14 km/s for P waves and from 4 to 7 km/s for S waves. Typical P wave velocities in the crust are ~6 km/s while S wave velocities are about half that at ~3 km/s.

Figure 2 shows a model of the Earth's velocity structure (PREM) that was developed by Dziewonski and Anderson (1981). PREM is just one of a number of Earth models that have been developed by researchers at different universities. All of the models are grossly similar, but they do differ in detail depending on the particular set of seismic observations that were used to derive them and the mathematical method used to infer the structure. Examples of other models include IASPEI (Kennett and Engdahl, 1991), which, has been adopted as the global reference model for the International Data Centre in Vienna, SP6 (Morelli and Dziewonski, 1993), and AK135 (Kennett, Engdahl and Buland, 1995).

TABLE 1 Table of earth properties from PREM (1981) global earth model

Region	Radius km	Density kg/m^3	Bulk modulus GPa	Shear modulus GPa	Vp km/s	Vs km/s
Inner core	0–1221	12774–13088	1343–1425	156–176	11–11.3	3.5–3.7
Outer core	1221–3480	9903–12166	644–1304	0	8–10	0
D″	3480–3630	5500	644	291	13.7	7.3
Lower mantle	3630–5700	4400–5500	255–641	120–290	10.2–13.7	5.5–7.3
Upper mantle	5700–6345	3375–4000	130–250	68–120	8–10	4.5–5.5
Crust	6345–6371	2600–2900	50–70	26–44	5.8	3.2

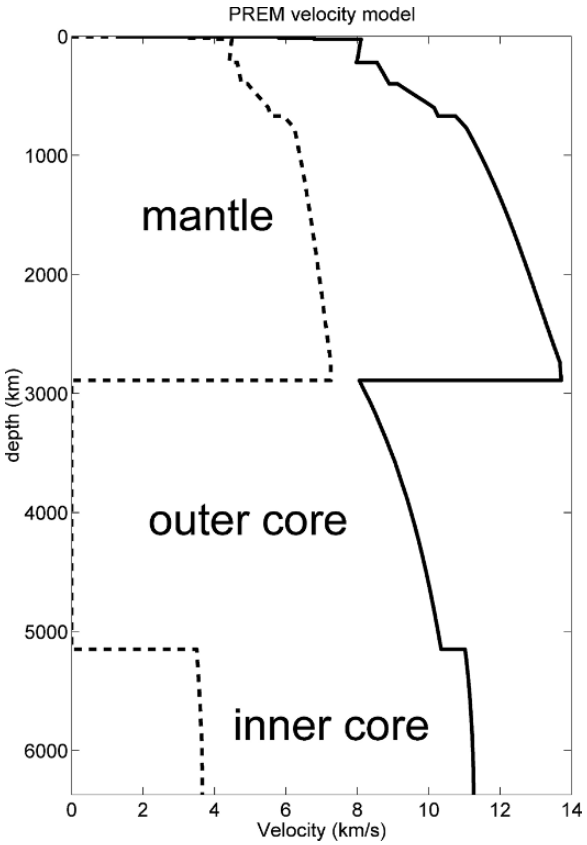


FIGURE 2 PREM (1981) Earth reference model showing the P wave (*solid line*) and S wave (*dashed line*) velocities as a function of depth. The outer core does not support S waves.

4.2 Anisotropy

The Earth is generally taken to be isotropic, that is, the velocity does not depend on the direction of the propagation. However, that is simply a mathematical convenience. Layered solids or aligned, anisotropic crystals can give rise to velocity variations that are a function of propagation direction. S waves, in particular, exhibit another interesting feature in anisotropic media. A third, independent wave is generated as a result of the S wave splitting into a "fast" and a "slow" wave with orthogonal planes of particle motion. This phenomenon is termed "shear wave splitting". In general, the propagation direction of these waves is not perpendicular to their wavefronts, so the particle motions differ from isotropic behavior (the group velocity and phase velocity vectors do not point in the same direction).

Studies of the core have found evidence that the inner core is anisotropic; waves travel 3%–4% faster in directions parallel to the north–south axis than in directions parallel to the equatorial plane). It is thought that perhaps non-uniformity of the core (due to the Earth's rotation) or flattening (due to an excess of iron that has a particular crystal structure) may be responsible; however, its cause remains an open debate. Papers that propose different theories for the anisotropy in the core include Stacey (1992); Karato (1993); and Kumazawa et al., (1992).

Mantle shear wave splitting has been observed in a number of places around the world. One of the great debates in the Earth science community is whether the anisotropy causing this splitting is due to aligned olivine crystals at the base of the plates sliding around on the Earth's surface or is a result of older processes that deformed and shaped the continents. Olivine is one of the major minerals in the mantle. It has orthotropic symmetry, which means it takes nine elastic constants to fully describe the elastic response of that mineral (as compared with just two in the isotropic case). Consequently, olivine crystals exhibit different velocities depending on the direction of wave propagation. In the mantle, olivine is thought to have a tendency to align in a particular direction when subjected to stress, such as the slow motion of tectonic plates over the surface of the Earth. If this alignment occurs on a large enough scale, then it is possible for seismic waves to be sensitive to the anisotropy. Recent studies have found frequency dependence (e.g., Marson-Pidgeon and Savage, 1997) associated with the shear wave splitting measurements, suggesting that variability in the anisotropy, as a function of depth, may be adding additional confusion to the whole question. A review paper by Silver (1996) discusses anisotropy in the mantle.

In the crust, layering, such as might occur when there are a series of depositional layers, or fracturing can give rise to anisotropic behavior.

4.3 Attenuation

In addition to discontinuities and geometrical spreading effects, another factor affecting amplitudes is energy loss (attenuation) due to anelastic processes. Intrinsic attenuation is generally attributed to internal friction during wave propagation from irreversible processes such as mineral dislocations and shear heating along grain boundaries. It is generally represented (inversely) by the parameter Q, which stands for "quality factor", a term borrowed from engineering literature. The strength of intrinsic attenuation is defined in terms of the fractional energy

loss per cycle:

$$Attenuation = \frac{1}{Q(\omega)} = -\frac{\Delta E}{2\pi E}$$

where ω is the frequency, E is the peak strain energy, and ΔE is the fractional loss of energy per cycle of oscillation. In the Earth, shear attenuation (Q_μ) is observed to be much stronger than bulk attenuation (Q_κ) and Q_κ is assumed to be infinite except in the inner core. Shear attenuation is strongest in the upper mantle and inner core. In the PREM Earth model, $Q_\mu = 600$ near the surface, drops to 80 in the upper mantle and then slowly increases to 312 for the lower mantle. Shear waves do not propagate in the fluid outer core, but in the inner core $Q_\mu = 85$. Because Q>>1 in most seismic applications, the amplitude of a seismic wave is often approximated as:

$$A(x) = A_0 e^{-\omega x/2VpQ}$$

where x is the distance traveled along the propagation direction. It can be seen from this equation that high frequencies will attenuate more rapidly than low frequencies (holding all else constant). Physically, this occurs because high-frequency waves undergo more oscillations than low-frequency waves do for a given distance traveled. Seismic waves that have traveled through attenuating regions will appear broader than waves that have not.

In intervals where Q is not frequency-dependent, the velocity dispersion may be expressed as:

$$Vp(\omega) = Vp(\omega_0)\left(1 + \frac{1}{\pi Q}\ln\frac{\omega}{\omega_0}\right)$$

where ω_0 is the reference frequency. For additional details on attenuation in the Earth, see Aki and Richards (2002).

5 • Commercial Seismology

Commercial seismology is focused on the shallowest portions of the crust. In oil and gas exploration, a prospect or "play" is created by the fortunate congruence of six elements over geologic time: source, maturation, migration, reservoir, trap, and seal. Source refers to the availability of abundant organic matter that is preserved in the absence of oxygen (reducing conditions) so that it is not consumed by aerobic bacteria. Often this happens when the organic matter is deposited along with mud or clay particles that seal it away from oxygen-rich water. The

maturation or chemical transformation of the organic matter into hydrocarbons requires that its burial history place it in a certain temperature regime for an appropriate period of time. For example, 50 million years at 100°C is about the center of the "oil window" where organic material will be transformed into oil. At higher temperatures (150°C) or substantially longer times (500 MY), more natural gas than oil would be formed. At still higher temperatures (260°C) the organic material would be carbonized and destroyed as potential source material. Geologists use seismic images of the subsurface to evaluate when and where appropriate conditions existed for creation of source material and to reconstruct the burial history of that material. Temperature regimes are often estimated from burial depth based on a thermal gradient for the sedimentary basin (roughly 30°C per kilometer of depth). Times are often determined from geologic unconformities seen on seismic images when these can be associated with known unconformities observed in other wells or in outcrop studies.

The source rocks where hydrocarbons are produced tend to be relatively impermeable so that the hydrocarbons cannot be directly extracted at economically viable rates. However, over geologic time scales the hydrocarbons migrate slowly upward toward the surface since they are lighter than the surrounding ground water. The migration pathways often include faults, fractures and higher-permeability rocks such as sandstones. A large enough volume of high-permeability hydraulically connected rock can become an economically viable oil or gas reservoir if there is a trap that causes the hydrocarbons to pool in the reservoir rock instead of continuing their migration to the surface. Potential traps include anticlines and faults or salt bodies that seal off the up-dip end of a permeable body of rock. A trap will only hold hydrocarbons if it has a good seal, which usually means an impermeable body of rock immediately above the trap that prevents the hydrocarbons from escaping. If the sealing rocks are weak or there are leaky faults that penetrate them, then the gas may escape and leave just the oil behind. Sometimes both escape leaving just ground water behind. Potential migration pathways, reservoirs and trapping structures are identified in seismic images by seismic interpreters.

In addition to revealing the structure of the subsurface so that the geologic history and the "plumbing" of the reservoir system can be determined, seismic data are also used to estimate rock properties (e.g., what is permeable sand vs. impermeable shale, how strong is the seal and what is the porosity and/or permeability of the reservoir rock) and fluid properties (e.g., is this reservoir full of oil, natural gas or ground water). One of the most

common techniques is to study the energy reflected at various incidence angles from subsurface layers. A P wave incident on an interface separating two solids will generally produce reflected P waves and SV waves (vertically polarized S waves), which vary in amplitude with the incident angle. Changes in the properties of the Earth from one layer to the next give rise to changes in the amplitude vs. angle characteristics of the reflected waves. These changes can be in rock type, such as from sand to shale or vice versa, or in the fluid contained within the rock, or in the porosity of the rock. By inverting the seismic data, geophysicists infer the velocity and density of the fluid-saturated reservoir rocks in an effort to determine rock properties and to discriminate between rocks with significant hydrocarbon accumulations and rocks filled with just water. Generally, some assumptions about the rocks and fluids must be made in order to constrain these sorts of inversions. For example, a relationship between velocity and porosity in a rock or between velocity and density in a fluid is often assumed. In practice, a well drilled in a nearby location or a geologically similar location often serves as an "analog". Relationships based on log data or rock samples from the analog are assumed to apply at the reservoir of interest. Additional relationships may be estimated based on analytical models. Mavko et al. (1998) is an excellent source for lab-measured rock properties. Gassman's equation (Gassman, (1951) is often used to estimate the velocity of a fluid-filled porous rock given the velocity, density and porosity of a dry rock. A good discussion of geophysical oil exploration methods is given in Sherriff and Geldart (1995).

References

Aki, Keiiti and Paul Richards, Quantitative Seismology, University Science Books, 2002.

Anderson, Don L., Theory of the Earth, Blackwell Scientific Publications, 1989.

Dahlen, F. A. and Jeroen Tromp, Theoretical Global Seismology, Princeton University Press, 1998.

Dziewonski, A, M. and D. L. Anderson, Preliminary reference earth model, Physics of the Earth and Planetary Interiors, 25, 297–356, 1981.

Gassman, F., Elastic Waves through a packing of spheres, Geophysics, 16, 673–685, 1951.

Karato, S., Inner core anisotropy due to the magnetic field-induced preferred orientation of iron, Science, 262, 1708–1711, 1993.

Kennett, B. L. N. and Engdahl, E. R., Traveltimes for global earthquake location and phase identification, Geophysical Journal International, 105, 429–465, 1991.

Kennett, B. L. N. and Engdahl E. R. Buland R., Constraints on seismic velocities in the Earth from traveltimes, Geophysical Journal International, 122, 108–124, 1995.

Kumazawa, M. Yoshida, and M. Yutani, Growth tectonics of the inner core (Abstract E3). Abstract volume of the Third SEDI Symposium, Mizusawa, Japan, 1992.

Lay, Thorne and Terry C. Wallace, Modern Global Seismology, Academic Press, 1995.

Marson-Pidegon, K. and M. K. Savage, Frequency-dependent anisotropy in Wellington, New Zealand,Geophys. Res. Lett., 24, 3297–3300, 1997.

Mavko, G., Mukerji, T. and J. Dvorkin, The Rock physics Handbook, Tools for Seismic Analysis in Porous Media, Cambridge University Press, 1998.

Morelli, A. and Dziewonski, A., Body wave traveltimes and a spherically summetric P- and S-wave velocity model, Geophysical Journal International, 112, 178–194, 1993.

Shearer, Peter M., Introduction to Seismology, Cambridge University Press, 1999.

Sheriff, Robert E. and Lloyd P. Geldart, Exploration Seismology, Cambridge University Press, 1995.

Silver, P. G. Seismic anisotropy beneath the continents: Probing the depths of geology, Annu. Rev. Earth Planet Sci., 24, 385–432, 1996.

Song, X. and P. G. Richards, Seismological evidence for differential rotation of the Earth's inner core, Nature, 382, 221–224, 1996.

Stacey, Frank D., Physics of the Earth, Brookfield Press, Brisbane, Australia 1992.

Turcotte, D. L., and G. Schubert, Geodynamics, Applications of Continuum Physics to Geological Problems, John Wiley & Sons, 1982.

Udías, Agustín, Principles of Seismology, Cambridge University Press, 1999.

84

Seismic Wave Propagation

Xianyun Wu[1] *and Ru-Shan Wu*[2]

[1]*ExxonMobil Upstream Research Company, Houston, TX, USA*
[2]*University of California, Santa Cruz, CA, USA*

1 • Introduction

A seismic wave is a mechanical disturbance or energy packet that can propagate from point to point in the Earth. Seismic waves can be generated by a sudden release of energy such as an earthquake, volcanic eruption, or chemical explosion. There are several types of seismic waves, often classified as body waves, which propagate through the volume of the Earth, and surface waves, which travel along the surface of the Earth. Compressional and shear waves are the two main types of body wave and Rayleigh and Love waves are the most common forms of surface wave.

Seismic waves carry a great deal of information about both the seismic source and the physical properties of the rock they penetrate. For example, naturally generated seismic waves might provide information on the characteristics of an earthquake source and the Earth's internal structure. These waves can be recorded by global, regional, or local network stations. The recorded data or seismograms can be used to quantitatively analyze the mechanisms, locations, and magnitudes of the earthquakes and build models of the Earth's interior structure. Artificially generated seismic waves provide information about the configuration of rock layers for oil exploration and, on a smaller scale, information as to the rigidity of shallow layers for

engineering purposes. Properties of oil reservoirs are determined using seismic waves generated by sources the surface of the earth and by sources in wells which penetrate the reservoir. The detection of underground nuclear explosions and their discrimination from natural and other man-made sources is another application of seismic waves.

2 • Medium Models

Seismic wave propagation in the Earth is extremely complicated, and it is impossible to exactly model it. Various assumptions and approximations must be made to quantitatively study the excitation and propagation of seismic waves and their relation to physical properties of the Earth. This chapter will present a brief description of the theory of seismic wave propagation in various idealized physical models such as the homogeneous, isotropic elastic model, the transversely isotropic (TI) model, the weakly heterogeneous model, and the poro-elastic model. There have been a great number of books published on this subject. The authors recommend the following books for those readers who are interested in studying this topic in detail: White (1965), Aki and Richards (1980), Ben-Menahem and Singh (1981), Kennett (1983), Hanyga (1985), Bourbiè et al., (1987), Berkhout (1987), Lay and Wallace (1995), Sato and Fehler (1998), Udias (1999), Carcione (2001), and Červenỳ (2001).

2.1 General Anisotropic Elastic Media

The fact that the stresses acting on an elementary volume of an elastic solid can be expressed as a linear combination of the strains was established experimentally for many substances in the seventeenth century, and the relation is now known as Hooke's law. This provides the basis for the description of static and dynamic deformation of physical media. In general, the strain–stress relation can be expressed as:

$$\sigma(\mathbf{x}) = \mathbf{c}(\mathbf{x}) : \varepsilon(\mathbf{x}) = \mathbf{c}(\mathbf{x}) : \frac{1}{2}(\nabla\mathbf{u} + \mathbf{u}\nabla) \tag{1}$$

where $\sigma(\mathbf{x})$ is the stress tensor (dyadic) at the position $\mathbf{x}$ in 3D, $\mathbf{c}(\mathbf{x})$ is the elastic constant tensor of the medium, $\varepsilon(\mathbf{x})$ is the strain field, $\mathbf{u}(\mathbf{x})$ is the displacement vector, $\mathbf{u}\nabla$ stands for the transpose of $\nabla\mathbf{u}$, and ":" stands for double inner product of tensors. In equation (1) the elastic tensor $\mathbf{c}(\mathbf{x})$ determines the constitutive relation of stress and strain and is a rank 4 tensor.

For a general anisotropic medium, $\mathbf{c}$ has 21 independent elastic constants (Carcione, 2001). For the homogeneous monoclinic, orthorhombic, and TI media, $\mathbf{c}$ has 12, 9, and 5 independent

elastic constants, respectively. These material symmetries are sufficient to describe most geological systems at all scales. For example, the TI medium may represent a finely layered medium, the orthorhombic medium may represent two sets of cracks with crack intersecting at 90°, or a vertical set of cracks in a finely layered medium, and the monoclinic medium may represent two sets of cracks with crack normals other than 0° or 90°. For a homogeneous and isotropic medium, $\mathbf{c}$ has only two independent elastic constants: the Lamè constants λ and μ. $\mathbf{c}$ can be a function of $\mathbf{x}$, e.g., in an inhomogeneous medium.

Generally, the equation of wave motion in a linear, heterogeneous, anisotropic medium can be written as (Aki and Richards, 1980):

$$\rho(\mathbf{x})\frac{\partial^2 \mathbf{u}}{\partial t^2} = \nabla \cdot \sigma(\mathbf{x}) + F(\mathbf{x}) \tag{2}$$

where $\rho(\mathbf{x})$ is the density of the medium and $F(\mathbf{x})$ is body force. If both $\rho(\mathbf{x})$ and $\mathbf{c}(\mathbf{x})$ are independent of $\mathbf{x}$, the medium is referred to as homogeneous.

2.2 Homogeneous, Isotropic Media

The simplest mathematical model for a rock is a homogeneous, isotropic media. Then equation (2) becomes

$$\rho\frac{\partial^2 \mathbf{u}}{\partial t^2} = F(\mathbf{x}) + (\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu\nabla \times (\nabla \times \mathbf{u}) \tag{3}$$

Equation (3) governs the propagation of two types of waves, namely, the compressional wave and the shear wave. The compressional wave is also called the primary wave or P wave, because it is normally the first signal to arrive in seismograms. The shear wave is also called the secondary or S wave, since it follows the P wave as the second arrival. Because they can propagate independently in an unbounded medium, P and S waves are also called body waves. The P and S waves differ in their propagation speeds and particle motions.

2.2.1 Compressional Waves

The compressional wave causes an arbitrary element of the medium to change its volume (contraction and expansion) without undergoing rotation. Thus, a compressional wave can propagate through solid rock and fluid. It travels at a speed of $\alpha = \sqrt{(\lambda + 2\mu)/\rho}$ ($\mu = 0$ for a fluid). In engineering, its speed is given as $\alpha = \sqrt{(K + 4\mu/3)/\rho}$ where K is the medium's bulk modulus. The particle motion associated with the wave is parallel to the direction of wave propagation, so the P wave is a longitudinal wave. The P wave has various applications. In seismic

exploration, the reflected P waves are used to image subsurface structure, which is part of the routine procedure in oil/gas exploration and reservoir evaluation. The amplitude of the reflected P wave is widely used as an indicator of hydrocarbons or for rock/reservoir property estimation using amplitude versus offset (AVO) analysis. In formation evaluation, the P wave travel time is used to determine formation porosity and lithology.

2.2.2 Shear Waves

Shear waves can cause an arbitrary element of the medium to have rotational distortion without a change in volume. The S wave travels at a speed of $\beta = \sqrt{\mu/\rho}$, making it slower than the compressional wave. Its particle motion is perpendicular to the direction of wave propagation and therefore it is also called a transversal or tangential wave. Note that shear waves can only propagate through an elastic solid ($\mu > 0$). Like the P wave, S wave measurements also have various applications. For example, S wave velocities can be used to estimate rock stiffness and determine anisotropy, and so on. Combining P and S wave velocities with density information, the bulk and shear moduli can be derived.

2.2.3 Surface Waves

Surface waves travel along the Earth's surface rather than through the volume of the Earth. Rayleigh and Love waves are the most common types of surface wave. In the expression of the reflection coefficients given above, there is a pole, which corresponds to the well-known Rayleigh wave mode that propagates along the surface at a speed little slower than that of the S wave. The pole occurs at the zero of the Rayleigh function (the denominator in equation (12)).

The Rayleigh wave is the most prominent arrival in the seismograms for near-surface sources. In exploration seismology, it is also called "ground roll" and is often muted before imaging (migration) of the deep reflection events. As a surface wave at a homogenous half-space, it has the following main properties: (1) its propagation speed, c_R, is a few percent less than the shear-wave velocity (for media with Poison ratio of 0.25, c_R is 0.92 times the S wave velocity); (2) c_R is independent of frequency, hence the wave is not dispersive; (3) the particles move in an elliptical trajectory (retrograde near the surface, but becoming prograde below a certain depth); (4) the amplitude decreases exponentially with depth (fundamental mode).

Love waves are horizontally polarized surface waves and are trapped guided waves due to post critical reflections from the surface above and some deeper reflector below, often the Moho. They exist only around the interface with a step increase in

velocity. They travel somewhat slower than the Rayleigh waves and their velocities are frequency dependent (i.e., dispersive). Both Rayleigh and Love waves are widely studied and used for seismological analyses.

For multi-layered media (the real earth), the Rayleigh and Love waves become more complicated. The Rayleigh wave in this case has higher modes and becomes dispersive. The dispersion curves (phase velocity vs. frequency) observed in earthquake seismology are widely used for velocity inversion for the structure beneath the seismic station. In exploration seismology, the dispersive Rayleigh wave is called a pseudo Rayleigh wave.

2.3 TI Media

Many crustal and mantle rocks exhibit a wide range of anisotropic features. The simplest anisotropic medium is TI, which has a symmetry with respect to the z-axis. In the plane perpendicular to the z-axis, all physical properties are the same, while they vary as a function of angle (θ) with respect to the axis. Daley and Hron (1977) gave a clear derivation of the directional dependence of the three phase velocities:

$$V_P(\theta) = \left\{\left[c_{33} + c_{44} + (c_{11} - c_{33})\sin^2(\theta) + D(\theta)\right]/2\rho\right\}^{1/2}, \quad (4)$$

$$V_{SV}(\theta) = \left\{\left[c_{33} + c_{44} + (c_{11} - c_{33})\sin^2(\theta) - D(\theta)\right]/2\rho\right\}^{1/2}, \quad (5)$$

and

$$V_{SH}(\theta) = \left\{\left[c_{66}\sin^2(\theta) + c_{44}\cos^2(\theta)\right]/\rho\right\}^{1/2} \quad (6)$$

with

$$D(\theta) = \Big\{(c_{33} - c_{44})^2 + 2\left[2(c_{13} + c_{44})^2 - (c_{33} - c_{44})(c_{11} + c_{33} - 2c_{44})\right] \sin^2(\theta) + \left[(c_{11} + c_{33} - 2c_{44})^2 - 4(c_{13} + c_{44})^2\right]\sin^4(\theta)\Big\}^{1/2}$$

where $c_{ij}(i, j = 1, \cdots, 6)$ are the elements of the elastic constant tensor **c**. Figure 1 shows the velocities of P, SV, and SH waves versus wave propagation angle θ (Xu, et al., 2005.) The shear wave will split into two events (slow S and fast S) when its propagation direction has a certain angle with respect to the symmetry axis.

Thomsen (1986) introduced a set of parameters, ε, γ, and δ (Thomsen's parameters), to express the anisotropy properties of a rock (see Figure 1). The detailed relations between P and S wave velocities and Thomsen's parameters can be found in Thomsen's paper.

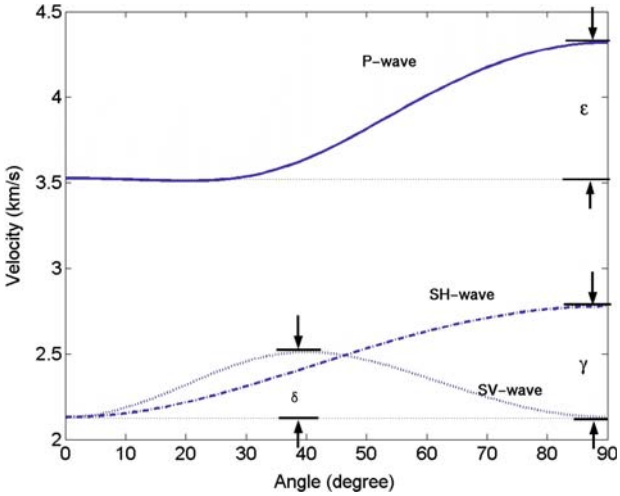


FIGURE 1
Velocities of P, SV, and SH waves versus incidence angle.

2.4 Heterogeneous Media

If $\mathbf{c}(\mathbf{x})$ and/or $\rho(\mathbf{x})$ in equation (1) are a function of position ($\mathbf{x}$), the medium is heterogeneous. Wavefields propagating in a heterogeneous medium will be scattered by the heterogeneities. The scattering characteristics of waves are dependent on the type of incident wavefield (P or S waves) and the type of perturbations. A detailed study of elastic wave scattering in heterogeneous media can be found in Wu and Aki (1985). If the parameters of the elastic medium and the total wavefield can be decomposed as $\rho(\mathbf{x}) = \rho_0 + \delta\rho(\mathbf{x})$, $\mathbf{c}(\mathbf{x}) = \mathbf{c}_0 + \delta\mathbf{c}(\mathbf{x})$, and $\mathbf{u}(\mathbf{x}) = \mathbf{u}^0(\mathbf{x}) + \mathbf{U}(\mathbf{x})$ where ρ_0 and $\mathbf{c}_0$ are the parameters of the reference medium, $\delta\rho(\mathbf{x})$ and $\delta\mathbf{c}(\mathbf{x})$ are the corresponding perturbations, $\mathbf{u}^0(\mathbf{x})$ is the incident field and $\mathbf{U}(\mathbf{x})$ the scattered field, then equation (2) can be rewritten as:

$$\rho_0 \frac{\partial^2 \mathbf{U}(\mathbf{x})}{\partial t^2} - \nabla \cdot \left[\mathbf{c}_0 : \frac{1}{2} (\nabla \mathbf{U} + \mathbf{U} \nabla) \right] = Q(\mathbf{x}) \tag{7}$$

with

$$Q(\mathbf{x}) = -\delta\rho(\mathbf{x}) \frac{\partial^2 \mathbf{u}(\mathbf{x})}{\partial t^2} + \nabla \cdot [\delta\mathbf{c}(\mathbf{x}) : \varepsilon(\mathbf{x})] \tag{8}$$

where $Q(\mathbf{x})$ is the equivalent body force due to scattering. If the heterogeneity exists only within a volume V, then the scattered field can be derived as (Wu, 1994):

$$\mathbf{U}(\mathbf{x}_1) = \int_V dV \left\{ -\delta\rho(\mathbf{x}) \frac{\partial^2 \mathbf{u}(\mathbf{x})}{\partial t^2} \cdot G(\mathbf{x}_1, \mathbf{x}) - [\delta\mathbf{c}(\mathbf{x}) : \varepsilon(\mathbf{x})] : \nabla G(\mathbf{x}_1, \mathbf{x}) \right\} \tag{9}$$

where $G(\mathbf{x}_1, \mathbf{x})$ is the displacement Green's function in the reference medium. For isotropic media,

$$\delta\mathbf{c}(\mathbf{x}) : \varepsilon(\mathbf{x}) = \delta\lambda(\mathbf{x})\,|\varepsilon(\mathbf{x})|\,\mathbf{I} + 2\delta\mu(\mathbf{x})\varepsilon(\mathbf{x}) \tag{10}$$

and

$$\begin{aligned} G(\mathbf{x}_1, \mathbf{x}) = {} & \frac{1}{4\pi\rho_0\alpha_0^2 r} e^{ik_\alpha r}\left[\hat{r}\hat{r} + \left(\frac{1}{ik_\alpha r} + \frac{1}{k_\alpha^2 r^2}\right)(\mathbf{I} - 3\hat{r}\hat{r})\right] + \\ & \frac{1}{4\pi\rho_0\beta_0^2 r} e^{ik_\beta r}\left[(\mathbf{I} - \hat{r}\hat{r}) + \left(\frac{1}{ik_\beta r} + \frac{1}{k_\beta^2 r^2}\right)(\mathbf{I} - 3\hat{r}\hat{r})\right] \end{aligned} \tag{11}$$

where $k_\alpha = \omega/\alpha_0$, $k_\beta = \omega/\beta_0$, α_0, and β_0 are the P and S wave velocities in the reference medium, respectively, $r = |\mathbf{x}_1 - \mathbf{x}|$, $\hat{r} = (\mathbf{x}_1 - \mathbf{x})/r$, and $\mathbf{I}$ is the unit tensor.

Replacing $\mathbf{u}(\mathbf{x})$ in the right-hand side of equation (9) by the incident field $\mathbf{u}^0(\mathbf{x})$, equation (9) becomes the well-known Born approximation solution. The Born approximation is used only for weakly inhomogeneous media and short propagation distances. However, recent progress has led to the development of iterative perturbation solutions, in which the Born approximation is replaced by the De Wolf approximation (multiple forward scattering and single backscattering approximation). The solution can be implemented in the form of a one-way marching algorithm in strongly heterogeneous media (Wu, 1994, 1996).

2.5 Porous Media

The real world is more complicated than the elastic model. Many earth materials are composite media consisting of an elastic porous skeleton and a pore fluid.

Gassmann (1951a,b) derived expressions for the elastic constants of a saturated porous medium on the assumption that the fluid and the solid skeleton move as a unit, and described the interaction between the two constituents. In other words, relative motion between fluid and skeleton has negligible influence on seismic wave propagation in fluid-saturated rocks.

Biot (1956) presented a comprehensive theory (called Biot theory) for fluid-filled rocks, deriving from basic principles a pair of differential equations that describe the coupled motion of the fluid and skeleton and discussing the properties of plane sinusoidal waves, which are the solutions to the equations. Fluid motion relative to the solid skeleton is taken into account, with flow being governed locally by Darcy's law. At low frequencies, viscous forces dominate, and the fluid moves as a unit with the solid; at high frequencies, inertial forces dominate, and a mass-coupling effect between solid and fluid occurs. The transition

between the two is governed by a critical frequency $f_c = \eta\phi/2\pi\kappa\rho_f$, where η is viscosity, ϕ is porosity, κ is permeability, and ρ_f is the density of the pore fluid. Biot theory predicts that there exist three types of body wave that can independently propagate in porous media. They are referred to as the fast compressional wave, shear wave, and slow compressional wave. The fast compressional wave is analogous to conventional compressional wave in an elastic medium. The corresponding solid and fluid displacements are in phase and approximately equal. The shear wave in a porous medium is equivalent to the shear wave in an elastic solid. The fast compressional and shear waves are weakly dispersive and slightly attenuated, whereas the slow compressional wave is mainly dominated by the pressure wave of the pore fluid and exists only in porous media. The displacements of the fluid and solid associated with the propagation of the slow compressional wave are out of phase, which causes the slow compressional wave high dispersion and strong dissipation. Biot theory provides a theoretical framework to investigate the effects of reservoir properties such as porosity and permeability on seismic wave attributes, which cannot be provided by traditional elastic wave theory. A distinct feature of the Biot theory is the prediction of the slow compressional wave, which was observed in the laboratory by Plona (1980). Biot theory is a particularly well-accepted theoretical model of wave propagation in porous media. Recommended books for wave propagation in porous media are Bourbiè et al. (1987) and White (1983).

3 • Reflections of Plane Wave

Waves that have planar wave fronts and uniform amplitude are called plane waves. An arbitrary wave front can be expressed as a superposition of a series of plane waves. When a steady-state plane wave (P wave or S wave) is incident on a free surface (solid–air or fluid–air interfaces), the incident, reflected, and transmitted waves must satisfy the traction-free boundary condition. Since the particle motion of S waves has two degrees of freedom, there are two kinds of polarizations of S wave relative to the free surface. When the particle motion is parallel to the surface, it is called an SH wave; when the particle motion is in a plane vertical to the surface, it is the SV wave. At the surface, P and SV waves are coupled so that P–SV, SV–P conversions can happen. On the contrary, the SH wave is decoupled from the P and SV waves and cannot convert into P and SV waves. For the P–SV case, assume the incident angle of the P wave and the reflection angle of the converted SV wave with respect to the normal to the free

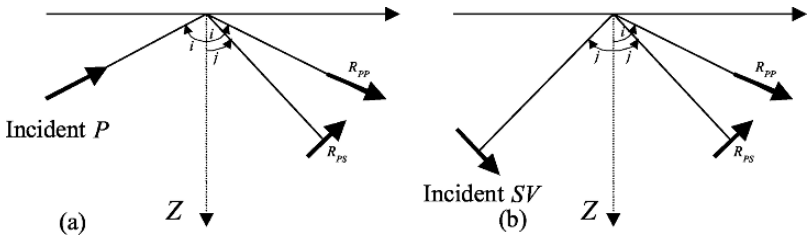


FIGURE 2 Notation for P and SV wave incident on a free surface.

surface are i and j. Figure 2 shows the notation and sign conventions for the coefficients of reflection and conversion due to P or SV incidence (Aki and Richards, 1980). Then the reflection and conversion coefficients are expressed as

$$R_{PP} = \frac{-\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha}\frac{\cos j}{\beta}}{\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha}\frac{\cos j}{\beta}}, \quad R_{PS} = \frac{4\frac{\alpha}{\beta}p\frac{\cos i}{\alpha}\left(\frac{1}{\beta^2} - 2p^2\right)}{\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha}\frac{\cos j}{\beta}},$$

$$R_{SP} = \frac{4\frac{\beta}{\alpha}p\frac{\cos j}{\beta}\left(\frac{1}{\beta^2} - 2p^2\right)}{\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha}\frac{\cos j}{\beta}}, \text{ and } R_{SS} = \frac{\left(\frac{1}{\beta^2} - 2p^2\right)^2 - 4p^2 \frac{\cos i}{\alpha}\frac{\cos j}{\beta}}{\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha}\frac{\cos j}{\beta}} \tag{12}$$

where subscripts PP and PS indicate the reflected P wave and reflected SV wave for P wave incidence, and SP and SS indicate the reflected P wave and SV wave for SV incidence. p is called the ray parameter (horizontal slowness) defined as $p = \sin i/\alpha = \sin j/\beta$, where α and β are the velocities of compressional and shear waves in the half space.

For the case of SH wave incidence on a free surface, there is no S–P conversion, since the particle motion is parallel to the surface. Only R_{SH} needs to be calculated. For the general case of a P–SV wave incident on an elastic interface, the P–SV reflection and transmission coefficients can be organized into an R–T matrix:

$$\begin{bmatrix} R_{PP} & R_{SP} & T_{PP} & T_{SP} \\ R_{PS} & R_{SS} & T_{PS} & T_{SS} \end{bmatrix}$$

where R and T denote reflection and transmission coefficients, respectively. These coefficients can be calculated with the Zoeppritz equations (Aki and Richards, 1980).

References

Aki, K. and Richards, P.G., 1980, Quantitative seismology: Theory and methods, I and II.

Ben-Menahem, A. and Singh, S.J., 1981, Seismic waves and sources, Springer-Verlag.

Biot, M.A., 1956, Theory of propagation of elastic waves in fluid-saturated porous solid, I: Low frequency range, J. Acoust. Soc. Am., 28, 168–178.

Berkhout, A.J., 1987, Applied seismic wave theory, Elsevier.

Bourbiè, T., Coussy, O., and Zinszner, B., 1987, Acoustics of porous media, Gulf Publishing Co.

Carcione, J.M., 2001, Wave fields in real media: Anisotropic, anelastic and porous media, Vol. 31, Elsevier Ltd.

Červenỳ, V., 2001, Seismic ray theory, Cambridge Univ. Press.

Daley, P.F. and Hron, F., 1977, Reflection and transmission coefficients for transversely isotropic media, Bull. Seis. Am., 67, 661–675.

Gassmann, F., 1951a, Ueber die Elastizität poröser Medien, Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, 96, 1–23.

Gassmann, F., 1951b, Elastic waves through a packing of spheres: Geophysics, 16, 673–685.

Hanyga, A., 1985, Seismic wave propagation in the Earth, Elsevier.

Kennett, B.L.N., 1983, Seismic wave propagation in stratified media, Cambridge University Press.

Lay, T. and Wallace, T., 1995, Modern global seismology, Academic Press.

Plona, T.J., 1980, Observation of a second bulk compressional wave in a porous medium at ultrasonic frequencies, Appl. Phys. Lett., 36, 259–261.

Sato, H. and Fehler, M.C., 1998, Seismic wave propagation and scattering in the heterogeneous Earth, Springer-Verlag, New York.

Thomsen, L., 1986, Weak elastic anisotropy, Geophysics, 51, 1954–1966.

Udias, A., 1999, Principles of seismology, Cambridge University Press.

White, J.E., 1965, Seismic waves: Radiation, transmission, and attenuation, McGraw-Hill Book Company.

White, J.E., 1983, Underground sound: application of seismic waves: Elsevier, Amsterdam.

Wu, R.S. and Aki, K., 1985, Scattering characteristics of waves by an elastic heterogeneity, Geophysics, 50, 582–595.

Wu, R.S., 1994, Wide-angle elastic wave one-way propagation in heterogeneous media and an elastic wave complex-screen method: J. Geophys. Res., 99, 751–766.

Wu, R.S., 1996, Synthetic seismograms in heterogeneous media by one-return approximation, Pure and Applied Geophys., 148, 155–173.

Xu, S., Wu, X., Huang, X., and Yin, H., 2005, (OTC), Evaluation of anisotropic rock properties in sedimentary rocks from well logs.

85

Seismic Data Acquisition

Christine E. Krohn

ExxonMobil Upstream Research Company, Houston, TX, USA

1 • Introduction

The seismic reflection method involves energizing seismic sources at discrete locations and recording signals at numerous receiver locations. The spatial sampling of the resulting seismic wavefield must be sufficient to image reflections and also to enhance signal and suppress noise (Pritchett, 1990). In order to determine the optimal source parameters and spatial sampling, it is important to obtain information about the imaging objectives, the seismic properties, and the data quality issues for the area. Analyses of previously recorded data are invaluable. Computer models, ranging from simple layers to representations of complex geology, are often constructed to test acquisition parameters and sampling.

Seismic data may be acquired for either 2D or 3D imaging. In 3D seismic acquisition, sources and receivers are deployed over an area of the Earth's surface and processed to produce an image of the underlying volume of the subsurface. The number of active receivers can be over 2,000 and cover an area of 6–12 km^2. In 2D seismic acquisition, sources and receivers are deployed along a line on the surface and processed to image the underlying vertical plane in the subsurface, assuming that subsurface properties vary only within the plane. This assumption is never satisfied in practice, so there is always some distortion in a 2D image. Most seismic surveys acquired by the oil industry over potential or discovered oil fields are 3D. Because 2D is less costly, it is commonly used for reconnaissance of large areas.

TABLE 1 3D Acquisition and processing costs for a 5×5 km block in millions of dollars (Costs can vary greatly for different parts of the world.)

Land open area	0.33
Land coastal woods to wetlands	0.60
Land transitional zone – marsh	0.83
Jungle	2.10
Marine	0.26

The goal of seismic acquisition design is to optimize the survey parameters to achieve the technical objectives with the least acquisition effort and cost. Acquisition is the largest cost element in seismic projects, typically several times the subsequent processing costs. Sample costs for acquisition and processing for different environments are shown in Table 1. In general, more sophisticated analytical methods, such as time-lapse methods for reservoir monitoring, require higher data quality and hence greater acquisition effort and cost than less sophisticated uses of the data, such as mapping large structures. Cost and quality issues are complex and often uncertain trade-offs are required (Houck, 2004). Acquisition parameters can limit fundamental resolution and data quality, and deficiencies usually cannot be fixed by processing. After the parameters are determined, a detailed implementation plan that addresses logistical, safety, and environmental issues must be developed for the specific site.

2 • The Acquisition System

Source and receiver technologies are different for land and for marine. For land, the most common sources are seismic vibrators (Baeten and Ziolkowski, 1990), which are truck or buggy deployable, weak, but highly repeatable sources. A metal plate is lowered to the ground, the weight of the truck is applied, and hydraulic actuators impart a swept frequency sine wave with duration of 6–12 seconds. Cross correlating the recorded data with the reference sweep compresses the radiated signal into a wavelet a few tens of milliseconds long. The cost of the survey is directly related to the time spent sweeping and listening per source interval in the survey. Dynamite is used when the area is not accessible by vibrators. The cost is related to the number and depth of the holes; the best signals are obtained when the dynamite is in competent rock below the weathering. For marine, the principal sources are air-gun arrays

(Dragoset, 2000; Caldwell and Dragoset, 2000). An air-gun array typically consists of several subarrays containing clusters of different-sized guns. Firing the guns causes a sudden release of high-pressure compressed air into the water. The array is designed to cancel the subsequent bubble oscillations to give a clean pulse. The bandwidth is shaped by the ghost reflection from the water surface with shallower deployment resulting in higher frequencies but with noisier data. Modeling is used to evaluate the array and deployment depths. Selected source parameters for an oil exploration survey are given in Table 2.

Most of the cost of seismic surveys is associated with the source effort while receivers are relatively inexpensive. Typically, a large number of receivers are deployed for each shot. For marine, shooting is "off-end" with the source behind the vessel followed by receiver streamers (Figure 1). For land acquisition, there is more flexibility and receivers are placed at different directions from the source (Figure 2). In 2D acquisition, having receivers on both sides of the source is called a "split spread". Receivers may also be placed on the ocean bottom or in wells for specialized surveys. Land recording systems can handle 3,000–5,000 receiver stations and marine vessels typically pull 6–12 cables, each with 400–500 receiver stations. Each recorded channel consists of a short array of single sensors. Hydrophones, used in marine streamers, are piezoelectric transducers, which are sensitive to pressure. Geophones, used on land, are directional electrical–magnetic transducers, which respond to the particle velocity. Conventionally, vertical geophones are used to record compressional wave energy, but increasingly three-component units with one vertical and two horizontal sensors are used to record and image both compressional and shear wave energy. Geophone signals and hydrophone signals are theoretically in phase, but a matching filter may be needed to combine the signals in processing due to differences in sensor frequency response. Newer high-technology three-component microelectromechanical (MEM) receivers, which respond to particle acceleration, are being developed (Dragoset, 2005). Land acquisition has a high cost because it involves the manual deployment of equipment including the receivers, digitizers, batteries, and cables. In comparison, the use of a large number of streamers towed from a single vessel has dramatically decreased the cost of marine surveys. Selected receiver parameters are given in Table 3.

Signals are digitized and recorded on digital media and positioning information is added. For exploration purposes, acquisition is typically performed with a 2-ms sample interval and with a 24-bit recording system. In these systems, instrument

TABLE 2 Source parameters

Parameter	Examples	Comments
Land vibrator		
Number of sweeps N_S	4	Combining N sweeps with phase rotations of $360°/N$ can suppress harmonic ghosts.
Number of vibrators N_V	1–4	The increase in signal-to-random noise ratio is $N_v\sqrt{N_s}$. Also, multiple vibrators form a source array and can suppress ground roll noise.
Sweep type and range	linear 6–100 Hz	5–6 Hz is the low frequency limit for vibrators. Frequencies are typically swept linearly with a constant amplitude or with a 6–8 dB /octave increase.
Sweep length	4–12 s	The length of the sweep and number of sweeps increase signal strength.
Listen time	6 s	The length of the correlated record; longer than 2-way time for the target.
Production rate	300–500 stations/day	Estimated from (the sweep time plus the listening time) multiplied by the number of sweeps and the time to move to the next station.
Land dynamite		
Charge size	1–20 lbs	Smaller charges used for shallow high-resolution surveys; larger charges used for deeper penetration.
Hole depth	60–200 ft	Drilling tools range from hand-held augers to truck-mounted units. Larger rigs are needed for deeper depths; the best data achieved when dynamite is in competent rock.
Marine air guns		
Air volume and pressure	3000–5000 in^3 2000 psi	Compressed air is limited by compressor capacity on vessel; source strength depends on specific array design and depth.
Peak-to-peak output	100 bar-m	Can be modeled using array design and individual gun signatures.
Peak-to-bubble ratio	12:1	Indicates the ability of array to sharpen pulse.
Depth of source and streamer	6-m source 8-m streamer	The gun output is reduced at depth; surface ghosts limit frequency bandwidth. Shallower tows are more susceptible to weather noise but provide higher frequencies.
Source interval	25 m	Boat speed must be at least 4–6 knots to avoid tangling the cables, limiting the source interval for a given record length.

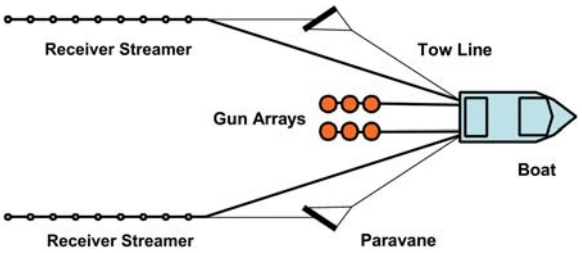


FIGURE 1 Marine 3D seismic acquisition. A seismic recording vessel can tow multiple receiver streamers and gun arrays; two each are shown in this illustration. Paravanes hold the streamers at a preset lateral distance from the boat. This arrangement allows the efficient collection of 3D data and the ability to record multiple imaging lines in each vessel pass.

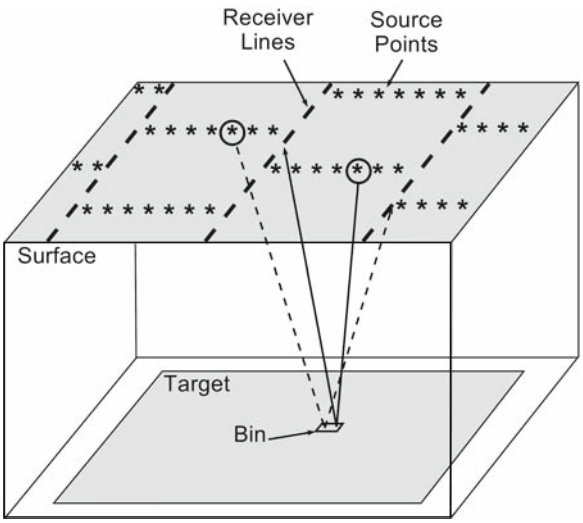


FIGURE 2 Land 3D seismic acquisition. With land acquisition, receivers are laid out in lines with typically four to six lines active for each source point. Source points are oriented along lines, which range from parallel to perpendicular to the receiver lines. Consider a bin of specified area on a flat target. The number of source and receiver pairs whose midpoints fall within a bin is called the fold. The pairs differ in their offset, the distance between source and receiver, and their azimuth, the direction between the source and receiver. With a given acquisition geometry, variations in fold and the offset and azimuth distributions from bin to bin can cause amplitude artifacts in the migrated image.

TABLE 3 Receiver parameters

Parameter	Examples	Comments
Geophones (land or ocean bottom)		
Frequency	10 Hz	Resonances of internal springs limit the low frequency response and lower frequency geophones are less rugged. Other specifications include distortion, sensitivity, and spurious resonances.
Type	Vertical; horizontal	Geophones are directional; vertical geophones are standard. Horizontal is needed to record shear waves.
Cases	Land or marsh	5–cm spikes are common; electrical leads may come from the top (marsh) or side (less wind noise).
Number per group	6–12	These sensors form an array. Signals are combined electronically and recorded on a single channel.
Group or station interval	25–50 m	Determined by desired CMP bin and need to sample or suppress ground roll.
Array length and shape	Inline or areal 50 m	Typically, the array length equals the station interval.
Hydrophones (marine)		
Group interval and array length	12.5 m	Standard group interval in many cables.
Steamer type	solid; oil filled	Solid streamers attenuate bulge waves within the cable.

noise is generally much smaller than environmental noises. Positioning information is critical for imaging and for locating wells. Positioning data for land consists of surveyed coordinates for the receiver and source locations. In the marine environment, global positioning system (GPS) units are located on the air-gun arrays and on the streamer tail buoys. Acoustic transponders and compasses on the streamer are used to locate the streamer hydrophones by interpolating between the GPS locations. The accuracy for marine positioning should be within 10 m laterally and within 1 m vertically for standard oil exploration. Higher accuracy is needed for repeat surveys and time-lapse measurements.

3 • Signal-to-Noise Quality

The signal-to-noise quality of the seismic data is enhanced by increasing the source effort, which includes the number and size of the sources, and by increasing data redundancy. The

total source energy per unit area must be sufficient to obtain a target signal that is above the noise for all frequencies within the seismic bandwidth. Seismic data from the area are analyzed to determine the maximum frequency that can be recovered from the target, and source effort tests are often performed at the start of each survey. The source needs to contain the frequencies of interest, and source effort must be sufficient to overcome environmental noise. Once the target can be illuminated, further increases in source size are generally not advantageous, because even though reflection energy increases, source-generated noise also increases.

After source effort, the chief method to enhance signal and suppress noise is through data redundancy, both from common-midpoint (CMP) stacking of differently sampled data and through field arrays. Together this approach has been called the stack array, and the ideal stack array can pass the desired wavenumbers corresponding to the primary reflections but attenuate others (Anstey, 1986). An ideal stack array is achievable in 2D acquisition, but in 3D, compromises are made to reduce cost.

In marine acquisition, short hydrophone arrays are used to suppress cable noise along the streamers. The air-gun arrays

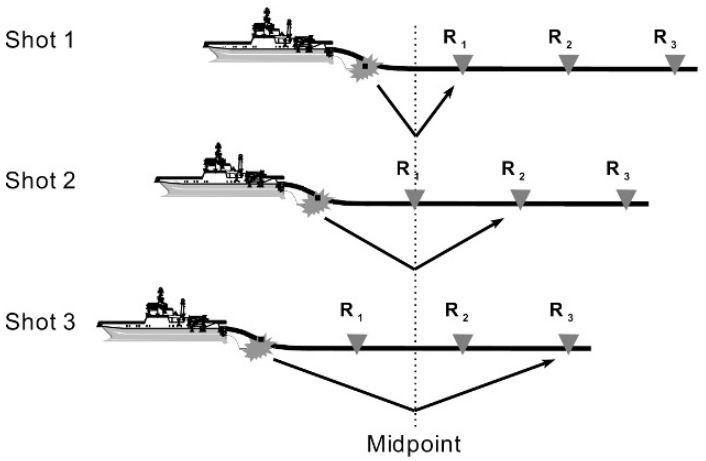


FIGURE 3 Two-dimensional marine CMP acquisition. As the boat moves forward, the air guns are fired at regular intervals, and the data are recorded by receivers (R_1, R_2, R_3, …) in a long streamer. The data are later sorted (or binned) into gathers where the source and receiver are centered on the same midpoint location. Each midpoint bin contains data with reflections from approximately the same vertical line through the subsurface.

suppress bubble oscillations and focus the downgoing energy. On land, both source and receiver field arrays are used to suppress horizontally propagating ground roll for wavelengths smaller than the receiver station interval. Typically, a receiver array consists of 6–12 receivers spread out over the station interval, and a vibrator source array often is composed of 4 vibrators. Longer wavelengths will be suppressed by the CMP stack (Ongkiehong and Askin, 1988). Areal and cross-line arrays may be needed in some areas to suppress cross-line noise. Arrays are most useful for 2 D acquisition where they can be oriented along the source-to-receiver azimuth to suppress source-related noises. Arrays are less effective for wide azimuth 3 D acquisition.

The CMP method, illustrated in Figures 2 and 3, involves stacking seismic traces with nearly the same midpoints but with different offsets, the distances between the sources and the receivers. The recorded traces are sorted in processing by CMP bin and offset. Traces with the same CMP bin (but differing

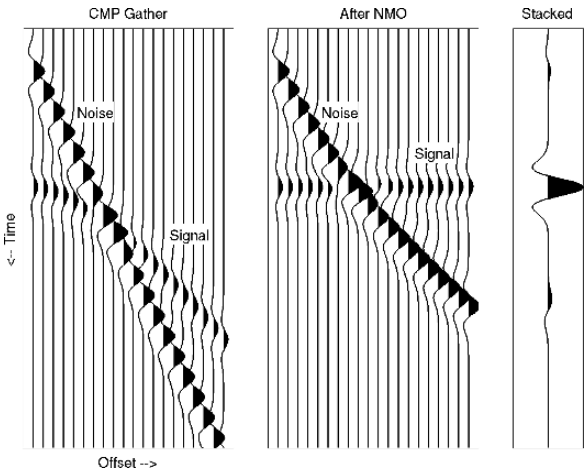


FIGURE 4 The CMP method. After sorting the data recorded in Figure 3 into midpoint gathers, signals corresponding to the same reflector are time aligned by the process of normal moveout (NMO), assuming a value for the average velocity to the reflector. After the desired signals are aligned, the traces are averaged. This process enhances the desired reflections and suppresses other events, such as water bottom multiples, which do not have the expected moveout with offset. Missing offsets make the stack less effective in removing the noise and cause amplitude variations. Offsets may be missing because of obstructions on the surface or because of compromises in the acquisition sampling. Most 3D acquisition geometries, as shown in Figure 2, will have variations in the offset distribution from bin to bin.

offsets) are shifted and stretched in time to align the desired reflections before averaging or stacking the traces (Figure 4). Stack may be performed before or after migration depending on the amount of lateral variation in geology. The amount of noise suppression by the stacking process will be related to the fold or the number of different traces and their offset distribution. Typically, the CMP bin size is half of the receiver interval, and the bins cover a region up to half the maximum useable offset. The fold is calculated as follows

$$\text{Fold} = \frac{\text{number of receiver stations} \times \text{station interval}}{2 \times \text{source interval}} \quad (1)$$

For full-fold acquisition, the source station interval is the same as the receiver interval. The best offset distribution is obtained by shooting every shot at positions located half-way between the receiver stations for split-spread land surveys. Full-fold shooting is expensive, and it is common to increase the distance between source locations. For example, dynamite surveys may have shots every fourth station. Offshore, two source arrays (Figure 1) are common and are alternately fired. This flip-flop shooting doubles the distance between source intervals for each CMP line. There is a trade-off in expense and in potential noise suppression. In

TABLE 4 Geometry parameters for S/N quality

Parameter	Examples	Comments
Maximum offset	110% of target depth	Longer offsets are muted in processing for shallow depths.
Receiver interval	25 m	Twice the CMP bin interval.
Number of receiver stations	2 times max offset divided by interval	Split spread shooting includes offsets on both sides of the source. Marine cables are often 5–7 km long.
Number of active receiver lines per shot	6–12	Determined by equipment, objective and terrain.
Receiver line spacing	200–600 m	Larger line spacing limits sampling of shallow data and has a larger azimuth distribution. It is difficult to tow marine cables too close together without tangling.
Shot interval	25–50 m	Skipping shot stations drops fold. Cross-line shot spacing is twice cross-line bin size.
In-line and cross-line fold	60:3	Greater fold results in greater noise suppression. Cross-line fold is much smaller than in-line fold.
Source pattern	zigzag, brick, slant	Different patterns have different offset and azimuth patterns, which can vary from bin to bin.

three dimensions, it is difficult and expensive to fully implement a full-fold stack array. Variations in fold from bin to bin for 3D geometries or missing data because of obstacles can cause variations in the stacked reflection amplitudes and in the noise suppression of 3D images. These artifacts must be evaluated considering the interpretation objectives (Cordsen, et al., 2000). Selected geometry parameters for signal-to-noise ratio quality are shown in Table 4.

4 • Sampling for Migration Imaging

The survey design should consider the technical objectives of the survey in terms of the desired resolution, the dips to be imaged, and the need for amplitude fidelity. These needs will determine the aperture, bin width, line direction and survey geometry required for imaging with migration. Migration is a processing procedure that moves reflectors from their recorded positions up dip to their actual positions and collapses the Fresnel zone. It is migration and the imaging of diffractions that determines the lateral resolution of the processed data. In survey design, these parameters can be estimated assuming post-stack migration and zero-offset or stacked traces.

To properly image a given target depth over a given image area, we must extend both the record time and the survey area to capture dipping reflections and diffractions (Figure 5). This distance and additional record length is called the migration aperture. The optimal survey size is determined by ray-tracing models using a velocity-depth curve for the area and Snell's law. Any information about complex structures and dips should be included. The aperture can add appreciably to the size of the survey area, especially with steeply dipping reflectors. Even for flat geology with small faults or terminations, the aperture should include the hyperbolic tails of diffractions out to 40° of dip so that migration can collapse the diffractions and properly image the terminations (Figure 5).

The CMP bin size and the receiver and source intervals are determined by the need to record without aliasing the largest dips at the maximum frequency for imaging (Figure 6) (Liner and Underwood, 1999). For example, the maximum bin size that will prevent aliasing of dipping reflectors and diffraction tails up to a given angle is:

$$\text{Maximum bin size} = \frac{\text{interval velocity}}{(4 \times \text{frequency} \times \text{SIN(dip angle)})} \qquad (2)$$

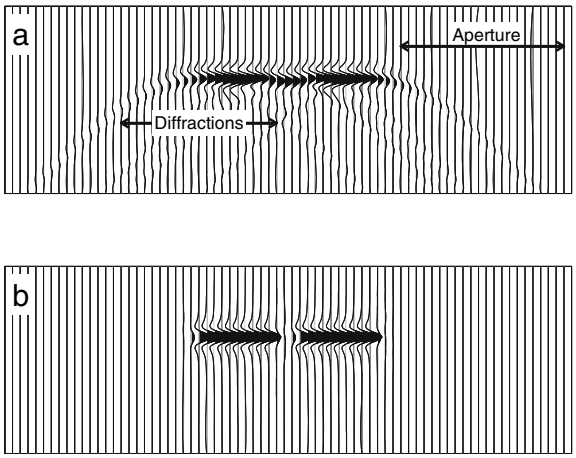


FIGURE 5 Lateral resolution and migration. Changes in reflector amplitude or terminations produce diffractions with dipping tails (**a**). Migration collapses the diffractions resulting in sharper reflector terminations and an improvement in lateral resolution. After migration (**b**) a gap in the reflector can be distinguished. To properly image the terminations, data must be recorded beyond the zone of interest by a distance called the aperture. In addition, the trace spacing must be small enough to record the diffraction dips without aliasing.

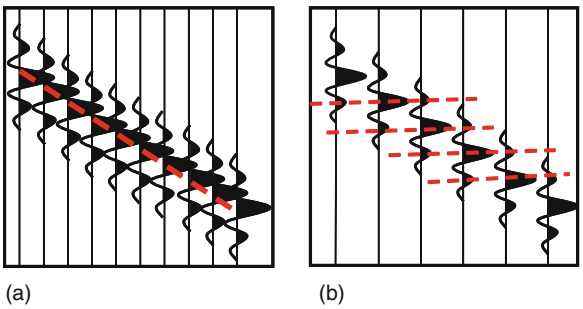


FIGURE 6 Bin size and spatial aliasing. The distance from trace to trace or bin to bin must be small enough to sample dipping events without aliasing. In (**a**), the sampling is sufficient to identify the event dip, and migration will be able to reposition the event to its proper location. If the trace spacing is coarse so that the aliasing criterion of two samples per apparent wavelength is not met, then the dip is ambiguous. As shown in (**b**), the apparent dip can be different from the true dip. In this case, the dipping event cannot be properly imaged with migration.

The lateral resolution (Vermeer, 1999) is limited by the ability to record and then position diffractions from terminations. For a given frequency and bin size, lateral resolution is determined by the largest dip angle, of diffraction tails that can be recovered without aliasing. Typically 40 degrees is a practical limit on recovery.

$$\text{Lateral resolution} = \frac{\text{vertical resolution}}{\text{SIN (maximum dip angle)}} \tag{3}$$

Vertical or temporal resolution is determined by the width of the seismic bandwidth above the noise for the target reflectors and by the interval velocities at the target (Figure 7). For good-quality data, the vertical resolution can be estimated using the Raleigh criterion (Kallweit and Wood, 1982).

$$\text{Vertical resolution} = \frac{\text{interval velocity}}{(4 \times \text{bandwidth})} \tag{4}$$

Line direction can also be important for imaging. It is relatively inexpensive to obtain fine sampling along the receiver lines. For example, station spacing on a standard marine cable is 12.5 m. It is much more expensive to sample finely in the cross-line

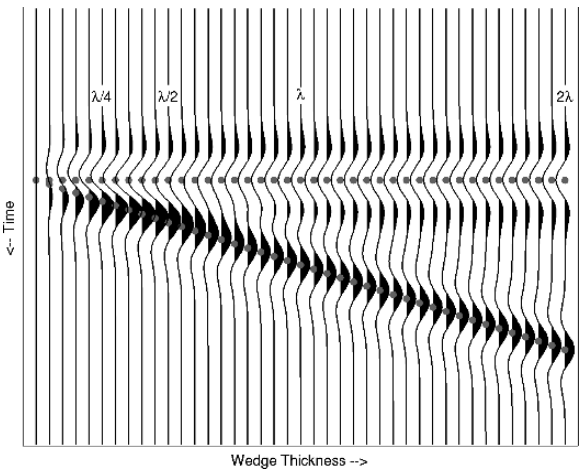


FIGURE 7 Vertical resolution and bandwidth. The vertical resolution is the smallest resolvable temporal distance between two reflectors. This is set by the bandwidth of the seismic wavelet and the amount of noise in the data. For very good data quality, we expect to resolve 1/4 of the dominant wavelength. We can often detect, but not measure, smaller events because interference or tuning will cause an amplitude increase in the reflection.

TABLE 5 Imaging parameters

Parameter	Examples	Comments
Aperture	Target depth	The increase in survey size compared to imaged area depends on dip and target depth.
Record length	2-way time to target plus 2s	Must capture diffractions from target.
Bin size	25 m	Must be large enough to record dipping energy without aliasing at the maximum frequency.
Line direction	Direction of maximum dip	Finer sampling is easier to achieve in the line direction.

direction. For this reason, it is generally preferred to orient the receiver lines along the direction of maximum dip to utilize the finer sampling. There are cases, however, where strike orientation is better for logistical issues or to simplify the effects of complex geometry on reflections and multiples.

Once the basic parameters are determined, the designer will choose a geometry or pattern for the sources and receivers. Offshore, the geometry may be dictated by the available acquisition vessel; onshore there is much greater flexibility (Cordsen, et al., 2000). The number of receiver lines active per shot and the receiver line spacing must be determined, and the shot pattern selected. The geometry can range from in-line shooting with the source line parallel to the receiver line to orthogonal shooting with the source line perpendicular to the receiver lines. Other common patterns include brick, zigzag, and slant configurations. For a given fold, orthogonal patterns can be constructed with less source effort (expense) than in line patterns but there tend to be large variations in offset distributions from bin to bin, which can create amplitude artifacts in the processed data. More uniform offset distributions can be constructed with inline patterns. Any pattern may have an acquisition footprint, which must be evaluated considering the survey objectives and the strength, depth, and structure of the target reflections. Selected survey parameters for migration imaging fidelity are shown in Table 5.

Summary

Trade-offs exist in data acquisition between increasing sampling and reducing costs. It is important to consider the effect of the sampling on imaging and on signal-to-noise optimization. Amplitude fidelity and footprints should also be evaluated. The

optimum parameters are difficult to estimate and experience is invaluable. Modeling based on prior information is useful and is especially important for complex structures with large velocity changes. However, our information is uncertain; the survey to determine the subsurface geology has not yet been acquired and this uncertainty should be taken into account.

References

Anstey, N.A. 1986, "What ever happened to ground roll" TLE, March 1986, 40–45.

Baeten, G. and Ziolkowski, A., 1990, The Vibroseis Source, Elsevier.

Cordsen, A., Galbraith, M., and Peirce, J., 2000, Planning Land 3-D Seismic Surveys, B. A. Hardage, ed, Geophysical Developments Series No. 9, Society of Exploration Geophysicists.

Caldwell, J. and Dragoset, W, 2000, "A brief overview of seismic air-gun arrays," The Leading Edge, Vol. 19, No. 8, 898–902.

Dragoset, W., 2000, "Introduction to air guns and air-gun arrays", The Leading Edge, Vol. 19, No. 8, 892–897.

Dragoset, W., 2005, "Expert answers: Digital geophones", The Leading Edge, Vol. 24, No. 2, 131–135.

Houck, R. T., 2004, "Predicting the economic impact of acquisition artifacts and noise" TLE, October 2004, 1024–1031.

Kallweit, R. S. and Wood, L. C., 1982, "The limits of resolution of zero-phase wavelets" Geophysics, Vol. 47, 1035–1046.

Liner, C. L. and Underwood, W. D., 1999, "3-D seismic survey design for linear v(z) media," Geophysics, Vol. 64, 486–493.

Ongkiehong, L. and H. J. Askin, 1988,"Towards the universal seismic acquisition technique" First Break, Vol 6, No. 2, 46–63.

Pritchett, W. C. 1990, Acquiring Better Seismic Data, Chapman and Hall.

Vermeer, G. J. O., 1999, "Factors affecting spatial resolution" Geophysics, Vol. 64, 942–953.

86

Signal Models in Seismic Processing

A.J. Berkhout

Delft University of Technology, Delft, The Netherlands

1 • Introduction

The data of a seismic survey is always a discrete spatial sampling of the wavefield. Therefore, seismic data traces can be conveniently arranged in a so-called data matrix. After removing waves that have travelled along the surface, the data matrix contains signals that can be expressed in terms of wavefield operators describing propagation and reflection in the subsurface. Wavefield operators may refer to different wave types, i.e., longitudinal and shear. Multiple scattering is addressed by adding feedback paths to the signal model for single scattering. The different steps in seismic processing can be easily explained with the aid of the presented signal models.

After the introduction of seismic digital signal processing in the 1960s, seismic processing in the 1970s was characterized by the time series approach. Processes such as statistical deconvolution, common-midpoint (CMP) stacking and time migration treated seismic data as time objects; both input and output were in seismic time. In the early 1980s, however, wave theory gained a principal role (Claerbout, 1976; Berkhout, 1982). The seismic community realized that seismic data should be considered as sampled *wavefields*, measured with proper aliasing protection at the data acquisition surface (Ongkiehong and Askin, 1988). Moreover, the transformation of surface measurements into

wavefields below the acquisition surface – wavefield extrapolation – evolved into one of the key operations in seismic signal processing. In this chapter, I will show seismic signal models that are based on wave theory and, therefore, are useful for formulating and understanding *multi-channel* seismic signal processing algorithms.

2 • The Data Matrix

It has been shown (Berkhout, 1982, 1993) that the data of a seismic survey (2D or 3D) can be conveniently arranged with the aid of the so-called *data matrix* $P(z_a, z_b)$ – a matrix of frequency domain signals – where the (i,j) element represents one frequency

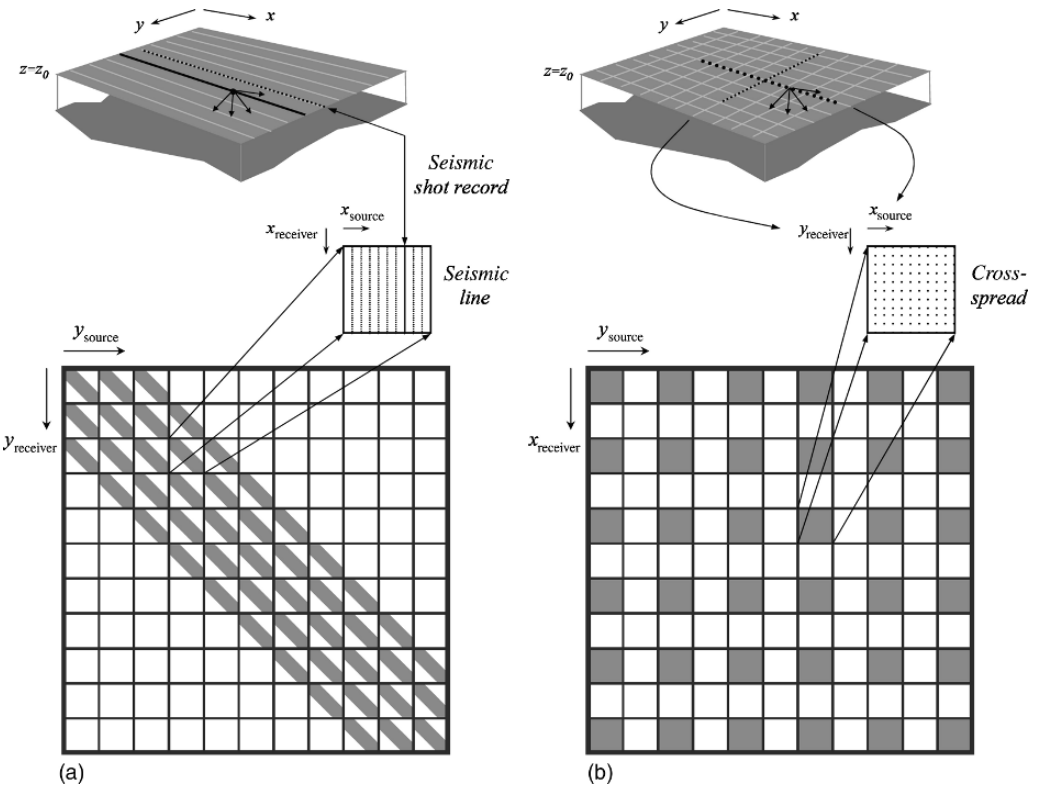


FIGURE 1 The data matrix for 3D seismic measurements. In multi-streamer marine data (**a**), one column represents one shot record (in this example five streamers are shown). In multi-cross-spread land data (**b**), one submatrix represents the data of a single cross-spread (in this example 36 cross-spreads are shown).

component of the signal recorded by a receiver at lateral location i from a source at lateral location j, where the receiver and source depths are z_a and z_b, respectively (Figure 1). This matrix can be used directly for the formulation of wave theory-based numerical algorithms in seismic signal processing such as multiple removal and prestack migration. Note that in most theoretical considerations, the data matrix is assumed to be filled completely with regularly sampled measurements. Figure 1 shows that this is not the case in practice. During a seismic survey, every surface location is not occupied by both a source and a receiver, so some of the elements of the matrix will be empty. For example, if only zero-offset data were collected, then the data matrix would be diagonal. Note also that in the frequency domain, each element of the data matrix represents one frequency component of a single trace, i.e., one complex number. It is understood that all matrices and operators discussed will be functions of frequency unless otherwise indicated.

3 • Source, Transfer and Receiver Matrices

The wave equation is linear and time-invariant, so that the signal recorded by a receiver at the ith location from a source at the jth location may be written as the convolution of the source wavelet $s(t;j)$ with a medium impulse response $h(t;i,j)$ and a receiver impulse response $d(t;i)$. In frequency domain, the data matrix may then be written as the product

$$\mathbf{P}(z_0, z_0) = \mathbf{D}^-(z_0)\ \mathbf{X}(z_0, z_0)\ \mathbf{S}^+(z_0). \tag{1}$$

In Equation (1), the jth column of the *source matrix* $\mathbf{S}^+(z_o)$ defines the downgoing source wavefield at surface z_o when the source array is at position j. The (i,j) element of the *transfer matrix* $\mathbf{X}(z_0, z_0)$ defines the upgoing wavefield at location i produced by a unit source at location j (source and receiver both at the surface z_0), and the ith row of the *receiver matrix* $\mathbf{D}^-(z_o)$ gives the sensitivity of the ith seismic trace to the upgoing wavefield at surface z_o when the receiver array is at location i. The (i,j) element of the data matrix $\mathbf{P}(z_0, z_0)$ is thus the seismic trace when the receiver array is at location i and the source array is at location j. In the special situation of single-point sources and single-point receivers (no arrays), $\mathbf{S}^+$ and $\mathbf{D}^-$ are diagonal matrices. We use the superscript $(+)$ and $(-)$ to remind the reader that we refer to operators on the downgoing $(+)$ and upgoing $(-)$ wavefields, respectively.

3.1 Signal Model for Single Scattering

While the full data matrix contains the raw measurements with surface-related multiples, interbed multiples and surface waves, the first goal of seismic processing is commonly to construct a new data matrix that contains only signals which go down to the target, reflect once and return directly to the sensor, i.e., a single scattering approximation to the full data matrix. This data matrix for single scattering, $\Delta\mathbf{P}(z_o, z_o)$, can be expressed in terms of propagation and reflection operators according to

$$\Delta\mathbf{P}(z_o, z_o) = \mathbf{D}^{-}(z_o)\,\Delta\mathbf{X}(z_o, z_o)\,\mathbf{S}^{+}(z_o) \tag{2}$$

with

$$\Delta\mathbf{X}(z_o, z_o) = \sum_m \mathbf{W}^{-}(z_o, z_m)\mathbf{R}(z_m, z_m)\,\mathbf{W}^{+}(z_m, z_o) \tag{3}$$

for each Fourier (or Laplace) component (Figure 2). Note that the source matrix $\mathbf{S}^{+}(z_0)$ and receiver matrix $\mathbf{D}^{-}(z_0)$ in Equation (2) are the same as encountered in Equation (1).

In Equation (3), the columns of $\mathbf{W}^{+}(z, z_o)$ are the (pseudo) primary[1] propagation operators of the downgoing wavefield, from each point at surface z_o to every point at depth level z. Similarly, the columns of $\mathbf{W}^{-}(z_o, z)$ are the propagation operators of the upgoing wavefield from each point at depth z to every point at the surface z_0. Due to the reciprocity principle, $\mathbf{W}^{+}(z, z_o) = \mathbf{W}^{-}(z_o, z)^{\mathrm{T}}$. As each column of $\mathbf{W}^{+}$ and $\mathbf{W}^{-}$ defines a unit point-source response, the elements of a "single-arrival" version may be written as

$$W^{+}_{\ell j}(z_m, z_o) = a^{+}_{\ell j} e^{-j\omega\tau_{\ell j}} H(\omega),$$
$$W^{-}_{ik}(z_o, z_m) = a^{-}_{ik} e^{-j\omega\tau_{ik}} H(\omega),$$

where $H(\omega)$ defines the spectrum of a band-limited temporal differentiator (optionally including fine-layering effects), $\tau_{\ell j}$ being the traveltime from source point $(\mathbf{x}_j, z_o)$ to reflection point $(\mathbf{x}_\ell, z_m)$ and τ_{ik} being the traveltime from reflection point $(\mathbf{x}_k, z_m)$ to detector point $(\mathbf{x}_i, z_o)$. Here, $a^{+}_{\ell j}$ and a^{-}_{ik} are related amplitude attenuation factors that include geometrical spreading and transmission losses. Note that in the multi-arrival version, a and τ are frequency dependent.

In Equation (3), the reflection operator for the downgoing wavefield at depth z_m is $\mathbf{R}(z_m, z_m)$ (Figure 2). If we take a diagonal

[1]The pseudo-primary contains the dispersion effect of fine layering.

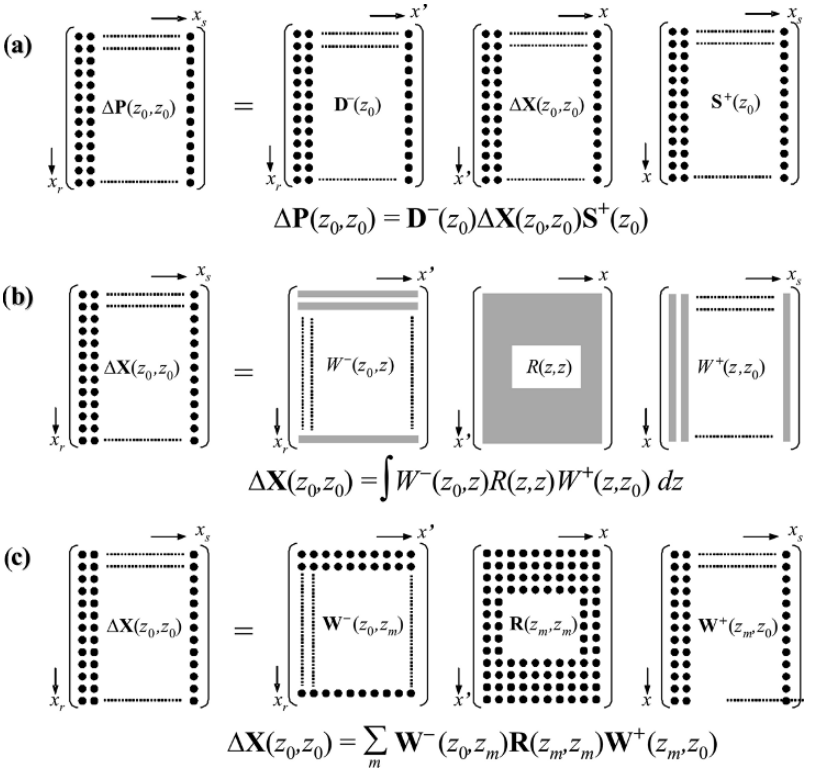


FIGURE 2 Pictorial illustration of the 3D primary reflection model for the measurements (**a**) and the transfer function (**b** and **c**), the latter being based on the continuous (b) as well as the discrete (c) formulation of wave propagation.

matrix for **R**, then the reflection process is angle *independent*. Angle-*dependent* reflectivity can be represented by a non-diagonal matrix **R**. For example, the plane wave reflection coefficient from a horizontal interface would be applied to an incoming wavefield as a multiplication in frequency-wavenumber domain. Our operator formalism represents signals in frequency-space domain, where the multiplication in wavenumber may be expressed as a convolution in space, by picking the appropriate form for **R** (Berkhout, 1982; de Bruin et al., 1990; de Bruin, 1992).

Equation (3) is a discrete implementation of continuous Kirchhoff-type integrals, describing propagation down, reflection and propagation up (**WRW**). Note that for continuous 3D primary reflections, the evaluation of each element of $\Delta\mathbf{X}(z_o, z_o)$ involves

two sets of double integrals per depth level. The first set of double integrals gives the reflected wavefield at depth z,

$$\delta X(\mathbf{x}, z; \mathbf{x}_s, z_o; \omega) = \iint R(\mathbf{x}, z; \mathbf{x}', z; \omega)\, W^+(\mathbf{x}', z; \mathbf{x}_s, z_o; \omega)\, \mathrm{d}^2\mathbf{x}', \quad (4)$$

for each source point $(\mathbf{x}_s, z_o)$ related to the elements of $\mathbf{S}^+(z_o)$, and the second set of double integrals propagates the reflected wavefield at depth z back to the receiver at the surface,

$$\Delta X(\mathbf{x}_r, z_o; \mathbf{x}_s, z_o; \omega) = \int \mathrm{d}z \iint W^-(\mathbf{x}_r, z_o; \mathbf{x}, z; \omega)\, \delta X(\mathbf{x}, z; \mathbf{x}_s, z_o; \omega)\, \mathrm{d}^2\mathbf{x}, \quad (5)$$

for each receiver point $(\mathbf{x}_r, z_o)$ related to the elements of $\mathbf{D}^-(z_o)$. The full signal is then composed by integrating over every depth slice of reflectors.

In Equations (4) and (5), any lateral position (x, y) has been indicated by the vector $\mathbf{x}$. Hence, the location of a grid point at lateral position k of depth level z_m is defined by $(\mathbf{x}_k, z_m)$. Note that in two dimensions, the vector $\mathbf{x}$ needs to be replaced by the scalar x.

Figure 3 summarizes the primary signal model, showing two levels of abstraction: the operator level and the impulse response level.

3.2 Signal Model Including Multiple Scattering

So far, the signal model has included only primary reflections, leading to primary transfer operator $\Delta\mathbf{X}$ and primary data matrix $\Delta\mathbf{P}$. Now, we will extend the signal model to multiple scattering by including a feedback path at each depth level, leading to transfer operator $\mathbf{X}$ and data matrix $\mathbf{P}$:

$$\mathbf{X}^{\cup}(z_{k+1}, z_{k+1}) = \mathbf{X}(z_{k+1}, z_{k+1}) + \mathbf{R}^{\cup}(z_{k+1}, z_{k+1}) \quad (6)$$

$$\overline{\mathbf{X}}(z_k, z_k) = \mathbf{W}(z_k, z_{k+1})\, \mathbf{X}^{\cup}(z_{k+1}, z_{k+1})\, \mathbf{W}(z_{k+1}, z_k) \quad (7)$$

$$\mathbf{X}(z_k, z_k) = \overline{\mathbf{X}}(z_k, z_k) + \overline{\mathbf{X}}(z_k, z_k)\, \mathbf{R}^{\cap}(z_k, z_k)\, \mathbf{X}(z_k, z_k) \quad (8)$$

for $k = M-1, M-2, \ldots, 0$ and with $\mathbf{X}^{\cup}(z_M, z_M) = \mathbf{R}^{\cup}(z_M, z_M)$, so that there is no upgoing signal from below depth z_M. Note that we have extended the notation in order to differentiate between down-to-up reflection ($\mathbf{R}^{\cup}$) and up-to-down reflection ($\mathbf{R}^{\cap}$).

In terms of wave theory, Figure 4 shows that each operator represents a transformation of the seismic wavefield *inside* geological layers (down and up propagation) and *at* geological boundaries (up and down reflection). The topology of the signal model in Figure 4 shows the interrelationships between these complex transformations. It explains the primary and the multiple reflections in seismic data at the highest conceptual

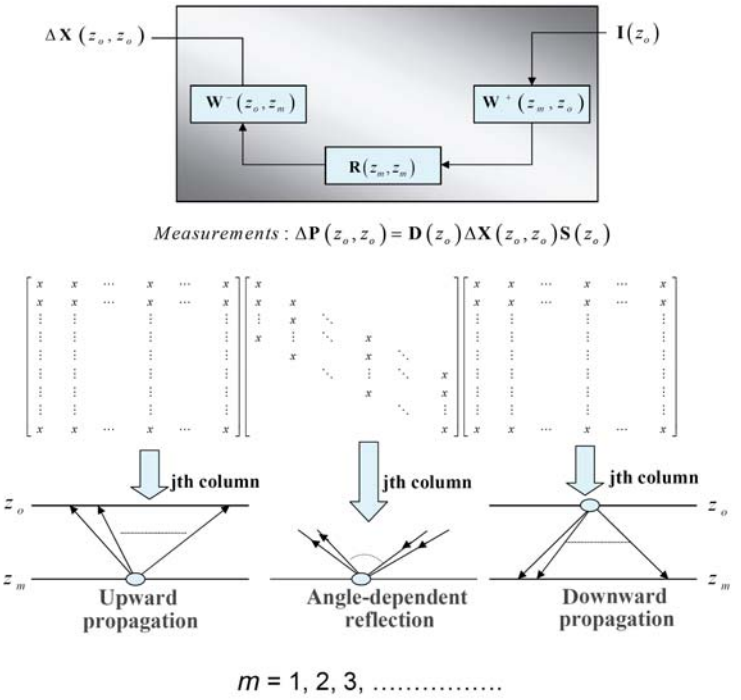
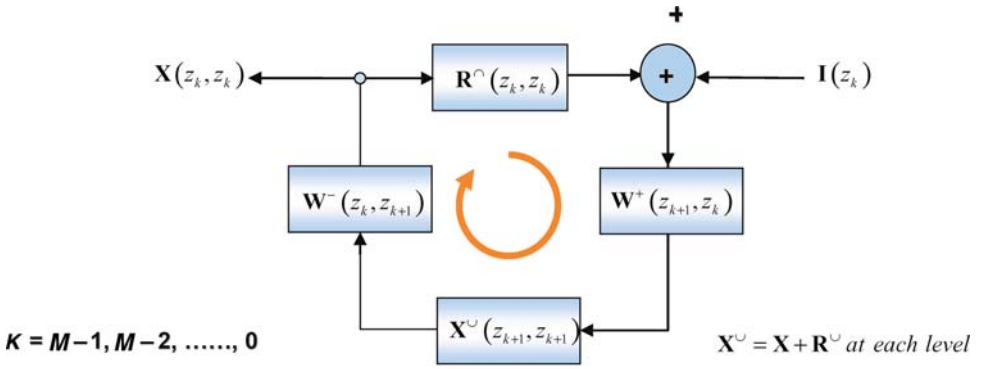


FIGURE 3 In the system formulation of seismic wave theory, wavefield operators **W** and **R** represent spatial impulse responses within one layer and at one boundary, respectively. Here, wavefield propagation (down and up) is shown in layer (z_o, z_m) and wavefield reflection (down to up) is shown at boundary z_m. Note that propagation operators $\mathbf{W}^+$ and $\mathbf{W}^-$ may refer to different wave types. Note also that the traveltimes in $\mathbf{W}^+$ and $\mathbf{W}^-$ contain the velocity information.

level. Moving one step down in the model, we have seen that each operator is formulated by a 4D integral kernel for each frequency component (two detector coordinates and two source coordinates); this kernel can be represented by a *matrix* in the discrete world. We have also seen that if we move one more step down, one column of each matrix represents a spatial impulse response, the elements of each column representing local amplitude and phase information of the involved impulse response. Hence, multiplication of operators involves multiplication of matrices, which again involves spatial convolution of impulse responses. This means that propagation (down and up) as well as reflection (up and down) are represented by multi-dimensional convolution. In inhomogeneous media, these convolutions are space variant.



$$\textit{Reflection measurements}: \mathbf{P}(z_o, z_o) = \mathbf{D}(z_o)\mathbf{X}(z_o, z_o)\mathbf{S}(z_o)$$

FIGURE 4 The signal model for seismic primary and multiple reflections (**P**) in terms of multi-dimensional impulse response (**X**), showing the underlying wavefield operators (**W** and **R**). For $k = 0$, the model adds the surface-related multiples to the subsurface response.

4 • Accuracy and Tractability in Seismic Signal Models

Mathematical modelling is finding a mathematical representation of a system which is both accurate (otherwise the model is meaningless) and tractable (otherwise the model is useless). In seismic signal processing, "accurate" means reproducing the full set of effects consistent with solutions to the linear viscoelastic, anisotropic wave equation in a heterogeneous medium. "Tractable" means that the model can be translated into a set of practical processing algorithms to reveal important geological information about the subsurface.

It is informative to consider the trade-off between accuracy and tractability when specifying the signal models at four levels of abstraction. Listed by increasing level of detail, these are as follows:

1. The physical system of operators, showing the wavefield transformations as they occur in the subsurface.
2. The mathematical system of matrices, replacing the involved operators by convolution matrices.
3. The impulse response samples, specifying the amplitude and phase value of each matrix element.
4. The optimized algorithm, taking into account economic shortcuts in the implementation.

At the highest level, i.e. level 1, reflection Equations (6, 7, 8) are perfectly accurate. Even the single scattering version of Equation (3) retains the ability to accurately handle single scattering in fully heterogeneous, attenuative and elastic media. This can be well understood by bearing in mind that in the seismic case, **W** and **R** are full elastic operators that include wave type conversions:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{pp} \, \mathbf{W}_{ps} \\ \mathbf{W}_{sp} \, \mathbf{W}_{ss} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} \mathbf{R}_{pp} \, \mathbf{R}_{ps} \\ \mathbf{R}_{sp} \, \mathbf{R}_{ss} \end{bmatrix}. \tag{9}$$

At the next level, i.e. level 2, linearity and stationarity are assumed. With these medium assumptions, the generalized operators can be specified by matrix operators, the columns of each matrix representing spatial impulse responses for one common frequency component (no harmonics).

Moving to the third level of detail, the individual samples of the impulse responses need be specified. For simple geological situations, these elements can be determined by solving the seismic wave equation (see, e.g., Aki and Richards, 1980), and analytical expressions for these elements provide useful insight. For complex geological situations, however, these analytical solutions will not represent the real world for two reasons:

1. An adequate description of complex geological layers and reflectors is not available.
2. An adequate multi-scale wave theory for dispersive heterogeneous media is not yet available.

Trying to formulate a tractable wave theory-based description at this third level of abstraction currently requires making significant approximations about the medium. This means that the current theoretical assumptions may force a significant bias on the solution. This is not always realized.

However, it is not necessary to enter this third level of detail via the traditional wave theoretical assumptions, and therein lies the power of this model. The elements of the matrix operators should not be estimated by iterative forward modelling in an assumed subsurface model, but these elements should be estimated by an inversion scheme that makes no specific assumptions about the complex interaction between elastic waves and geological media, apart from reciprocity. Processing algorithms based on conceptual formulations are called uncommitted or data-driven, in that they estimate the propagation and reflection operators directly from the data matrices after taking the acquisition and near-surface conditions into account. These data matrices embody all the details of the true impulse responses without having to commit to approximations about these responses.

It is the author's opinion that conceptual formulations at different levels of abstraction are *the* way that experts of different disciplines should communicate with each other in cross-disciplinary project teams.

4.1 Seismic Signal Models and Seismic Signal Processing

The conceptual formulations in Figures 3 and 4 yield a systems view of the seismic signal. If the wavefield operators in the signal model are represented by matrices with the predefined structure (Figures 2 and 3), then they can also be used to implement seismic signal processing algorithms. For instance, multiplication of a seismic data set with the matrix operator **W** defines a forward wavefield extrapolation process in a heterogeneous subsurface, the extrapolation operator for lateral position $\mathbf{x}_k$ being represented by the kth row of **W**. Similarly, multiplication of a seismic data set with matrix operator **R** defines an angle-dependent and frequency-dependent reflection process at any type of geological boundary, the reflection operator for lateral position $\mathbf{x}_k$ being represented by the kth row of **R**. According to Equations (2) and (3), unbiased estimation of **R** requires uncommitted knowledge of **W** (velocity problem). Examples of accurate and tractable processing algorithms derived using the same uncommitted inversion formalism include surface-related multiple removal (Berkhout, 1982; Verschuur, 1991; Berkhout and Verschuur, 1997), reflectivity estimation (Berkhout, 1982; De Bruin, 1992), controlled illumination (Berkhout, 1992; Rietveld and Berkhout, 1994; Rietveld, 1995) and common focus point (CFP) technology (Berkhout, 1997; Thorbecke, 1997).

Looking at the signal models in Figures 2–4, today's seismic processing arena could be summarized as follows:

1. In the preprocessing step, the imperfect way seismic data are acquired should be compensated for. In terms of the signal model, this means that the preprocessing step should transform **S** and **D** as much as possible to unity matrices, meaning the transformation of **P** into **X** (Figure 4). In practice, this is done by equalization, interpolation and deconvolution.
2. Next, multiple scattering must be removed. In terms of the signal model, this means that the feedback path at each depth level should be taken away, meaning that **X** is transformed into $\Delta\mathbf{X}$ (Figure 3).
3. Now, we are ready to separate the propagation effects of the overburden from the reflection properties of the geological boundaries. In terms of the signal model, this means that we have to separate **W** and **R** at each depth level (Figure 3). This is done in seismic migration.

4. Finally, the operators **W** and **R** can be used to determine the geological properties of the subsurface, and particularly the engineering properties of the (potential) reservoir. This is done in seismic reservoir characterization by parametric inversion.

Today, imaging step 3 requires a velocity model. This model is translated into operator **W** by forward modelling, and then **W** is removed from the primary data (migration), resulting in **R**. In practice, the combination velocity estimation and migration is done in an iterative manner (MVA). In the future, one may expect that iteration will not only occur at the velocity level but also at the **W** level ("uncommitted seismic migration"). This means that in uncommitted migration, the velocity model is not translated into **W** by forward modelling, but **W** is translated into the underlying velocity model by tomographic inversion.

Acknowledgement The author is grateful to Dr. Max Deffenbaugh for his helpful comments to improve the clarity of the manuscript for the non-seismic readers.

References

Aki and Richards, P.G., 1980, Quantitative Seismology, W.H. Freeman and Co.

Berkhout, A.J., 1982, Seismic Migration, Elsevier Amsterdam, pp. 151–195.

Berkhout, A.J., 1992, Areal shot record technology, J. Seis. Expl., 1, 251–264.

Berkhout, A.J., 1993, A unified approach to acoustical reflection imaging, Part I: The forward model, J. Acoust. Soc. Am., 93, No. 4, 2005–2017.

Berkhout, A.J., 1997, Pushing the limits of seismic imaging, Geophysics, 62, No. 3, 937–969.

Berkhout, A.J. and Verschuur, D.J., 1997, Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations, Geophysics, 62, No. 5, 1586–1595.

Claerbout, J.F., 1976, Fundamentals of Geophysical Data Processing, McGraw-Hill.

De Bruin, C.G.M., 1992, Linear AVO inversion by prestack depth migration: Ph.D. thesis, Delft University of Technology, The Netherlands.

De Bruin, C.G.M., Wapenaar, C.P.A., and Berkhout, A.J., 1990, Angle dependent reflectivity by means of prestack migration, Geophysics, 55, No. 9, 1223–1234.

Ongkiehong, L. and Askin, H.J., 1988, Towards universal seismic acquisition techniques, First Break, 6, No. 2, 46–63.

Rietveld, W.E.A., 1995, Controlled illumination in prestack seismic migration: Ph.D. thesis, Delft University of Technology, The Netherlands.

Rietveld, W.E.A. and Berkhout, A.J., 1994, Prestack depth migration by means of controlled illumination, Geophysics, 59, No. 5, 801–809.
Thorbecke, J.W., 1997, Common focus point technology: Ph.D. thesis, Delft University of Technology, The Netherlands.
Verschuur, D.J., 1991, Surface-related multiple elimination: Ph.D. thesis, Delft University of Technology, The Netherlands.

87

Deconvolution and Optimal Filtering in Seismology

Ramesh Neelamani

ExxonMobil Upstream Research Company, Houston, TX, USA

1 • Introduction

Deconvolution is an important, well-studied problem that is commonly encountered in seismology [1–4]. During hydrocarbon exploration, seismic receivers measure a noisy version of the earth's response that is blurred by a source wavelet (for example, from marine air guns, or land dynamite charges). Deconvolution becomes necessary to deduce the earth's response from the blurred and noisy receiver measurements. In earthquake seismology, the receivers measure seismic waves generated by earthquakes. In this case, deconvolution is employed to isolate the earthquake source time function, which characterizes the underground faulting process, from propagation effects [5–7].

Deconvolution can be formally stated as follows. The measured samples $y(m)$, with m the sample number, consist of the desired signal samples $e(m)$ first blurred by convolution (denoted by $\circledast$) with a source signature $s(m)$ and then corrupted by additive noise $n(m)$ (Figure 1)

$$y(m) = s(m) \circledast e(m) + n(m), \tag{1}$$

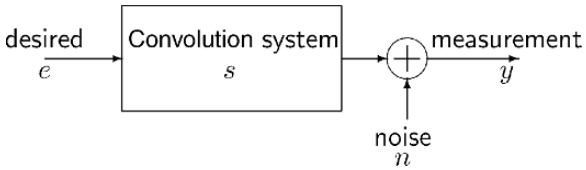


FIGURE 1 Convolutional model. The measurement y is a noisy version of the desired signal e blurred by convolution with the system impulse response s.

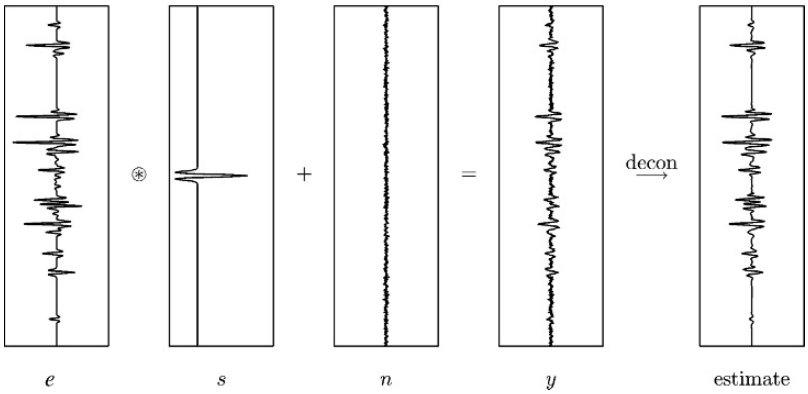


FIGURE 2 Deconvolution example. The estimate is obtained by employing ForWaRD [8], a deterministic deconvolution technique, on the blurred and noisy receiver measurement y.

Note that (1) can be rewritten in frequency domain as

$$\mathcal{Y}(f) = \mathcal{S}(f)\,\mathcal{E}(f) + \mathcal{N}(f) \tag{2}$$

with $\mathcal{Y}$, $\mathcal{S}$, $\mathcal{E}$, and $\mathcal{N}$ the respective Fourier transforms of y, s, e, and n. Deconvolution seeks to estimate e from the measurement y (Figure 2).

Deconvolution problems can be broadly classified into two types based on whether the source signature s is known. The next two sections address these two types of deconvolution. See Table 1 for a list of various deconvolution techniques.

2 • Deconvolution with Known Source Signatures

In some deconvolution problems, the source signature s in (1) is known. Such problems are often referred to as *deterministic deconvolution*. For example, in marine seismic exploration and vibroseis-based land exploration, the air gun and vibroseis source

TABLE 1 Various deconvolution techniques

Deconvolution type	Signal model	Source model	Noise model	Technique	Properties
Deterministic	Gaussian	Minimum phase	Gaussian	Causal Wiener–Levinson filter [9]	MSE-optimal filter
	Gaussian	Non-minimum phase	Gaussian	Non-causal Wiener–Levinson filter [9,18]	MSE-optimal filter
	Smooth (Sobolev functions)	Any	Gaussian	Non-causal Wiener or Tikhonov filter [18]	Asymptotically MSE-optimal
	1D piecewise smooth (Besov functions)	Frequency response $\propto f^{-\alpha}$	Gaussian	Wavelet-based [16] or ForWaRD [8]	Asymptotically MSE-optimal
	2D piecewise smooth (C^2) with smooth (C^2) boundaries	Frequency response $\propto f^{-\alpha}$	Gaussian	Curvelet-based [15]	Asymptotically MSE-optimal
	1D/2D piecewise smooth	Any	Gaussian	ForWaRD [8]	Robust performance; asymptotic MSE analysis possible
Statistical (Blind)	Gaussian	Minimum phase	Gaussian	Second-order statistics-based [19]	Robust for minimum phase
	Non-Gaussian	Any	Any	Higher-order statistics-based [20–22]	Good for sparse reflectivity
	Well-understood cepstrum	Well-understood cepstrum	Any	Homomorphic [23,24]	Useful to estimate short, band-limited sources
	Any	Surface consistent	Any	Surface consistent [25–27]	Robust; exploits multiplicity of measurements

signatures s are often assumed to be known. Another example of deterministic deconvolution is *wavelet shaping*, where the goal is to replace the existing known s with another desired $\widetilde{s}$ that typically has a larger bandwidth. That is, wavelet shaping aims to effectively deconvolve $\frac{\mathcal{S}(f)}{\widetilde{\mathcal{S}}(f)}$ from

$$\mathcal{Y}(f) = \left(\frac{\mathcal{S}(f)}{\widetilde{\mathcal{S}}(f)}\right)\left(\widetilde{\mathcal{S}}(f)\mathcal{E}(f)\right) + \mathcal{N}(f). \tag{3}$$

If the s in (1) is simply inverted, then the noise n corrupting the measurements gets amplified at all frequencies where $\mathcal{S}(f)$ is small.

$$\frac{\mathcal{Y}(f)}{\mathcal{S}(f)} = \mathcal{E}(f) + \frac{\mathcal{N}(f)}{\mathcal{S}(f)}. \tag{4}$$

The challenge lies in inverting s without noise amplification.

2.1 Linear Filtering Approaches

Linear filtering approaches aim to provide an estimate $\widehat{e}$ of e by applying a linear filter to y. A popular linear filtering technique filters y with a length-k filter a designed using the Wiener–Levinson least squares approach [9]; the estimate $\widehat{e} = y \circledast a$. To obtain a zero-delay version of e, a must be such that

$$\widehat{\mathbf{e}} = \begin{bmatrix} y(0) & 0 & \dots & 0 \\ y(1) & y(0) & \ddots & \vdots \\ y(2) & y(1) & \ddots & 0 \\ \vdots & & & y(0) \\ y(M) & & & \vdots \\ 0 & y(M) & & \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & y(M) \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ \vdots \\ a(k) \end{bmatrix} = \mathbf{Y}\,\mathbf{a} \approx \mathbf{e}. \tag{5}$$

In (5) and henceforth, bold-face capital letters (say, $\mathbf{Y}$) denote convolution matrices and bold-face small letters (say, $\mathbf{y}$) denote column vectors that are constructed using a sampled signal (say, $y(m)$). (If s is not minimum phase, then for improved results, the a must be designed with a time-delayed version of e as the desired output). The least squares optimum a is obtained by minimizing the expected value (with $\mathbb{E}$ the expectation operator) of the error $\mathbb{E}\left((\mathbf{Ya}-\mathbf{e})^T(\mathbf{Ya}-\mathbf{e})\right)$ with respect to a; this yields the *normal* equations

$$\mathbb{E}\left(\mathbf{Y}^T\mathbf{Y}\right)\mathbf{a} = \mathbb{E}\left(\mathbf{Y}^T\mathbf{e}\right) \tag{6}$$

$$\left(\mathbf{S}^T\mathbf{S} + \epsilon\mathbf{I}\right)\mathbf{a} = \mathbf{S}^T\delta, \tag{7}$$

with δ denoting the unit vector $[1\ 0 \dots 0]^T$, $\mathbf{I}$ the identity matrix, and $\frac{1}{\epsilon}$ a measure of the signal-to-noise-ratio (SNR) (ϵ is termed

the *prewhitening* factor). The above derivation implicitly assumes that the e and n are white random processes, but these assumptions can be relaxed. Typically, the Levinson recursion algorithm is employed to efficiently solve (7) [9].

The optimality of the Wiener–Levinson estimator a is that it provides the best mean-squared-error (MSE) estimate of e among all zero-delay length-k filters. The MSE decreases as the length of a increases.

If the length of the Wiener–Levinson filter is not constrained, and if causality is not imposed, then the performance of the finite length Wiener–Levinson estimator can be improved by using the Wiener deconvolution filter. Its frequency response is given by

$$\frac{1}{\mathcal{S}(f)} \frac{|\mathcal{S}(f)|^2}{|\mathcal{S}(f)|^2 + \frac{\mathcal{P}_{nn}(f)}{\mathcal{P}_{ee}(f)}}, \tag{8}$$

with $\mathcal{P}_{ee}(f)$ and $\mathcal{P}_{nn}(f)$ denoting the frequency spectra of the e and n respectively. (The Wiener–Levinson filter's frequency response tends to (8) when its length k is increased, and the delay assumption is dropped.) If the e and n are both stationary Gaussian random processes, then the MSE performance of (8) cannot be improved.

In practice, deconvolution is typically performed on a trace-by-trace basis. Multichannel Wiener–Levinson deconvolution filters [10] can act on several traces simultaneously to yield improved estimates by exploiting the coherency of typical reflectors across traces. However, such multichannel filtering techniques have (surprisingly) not been widely adopted for deterministic deconvolution in seismic signal processing.

2.2 Non-linear Approaches

Typical seismic reflectivity signals e are spiky in time and coherent across traces. Such signals are not well-modeled by Gaussian random processes [11,12]. For non-Gaussian e, the linear Wiener–Levinson estimator (for any delay and length) is no longer optimal. Better estimation can be obtained by techniques that employ non-linear estimation in novel multiscale domains such as wavelet [13] and curvelet domains [14]. Such techniques leverage the sparsity of wavelet and curvelet representations to attenuate the noise that gets amplified (see (4)) during inversion of s [8, 15–17].

3 • Deconvolution with Unknown Source Signatures

In many seismic deconvolution applications, the source response s in (1) is unknown (or known only partially) and needs to be estimated along with the desired signal e. This problem is often

referred to as *blind deconvolution* in signal processing and as *statistical deconvolution* in seismology. For example, when a dynamite source is used for land exploration, the source signature s is unknown. In marine exploration, the undesirable *ghost reflections* and *water column reverberations* (also called *multiples*) [28] can be approximately modeled by the source signature s in (1). In both cases, only partial information about s is available. *Ghost reflections* are caused by a part of the air gun energy reflecting off the sea surface. They manifest in seismic measurements as a time-delayed, polarity-reversed repetition of any given reflection. Using Z-transforms their effect can be (approximately) modeled as multiplying the source wavelet by $1 + r_w z^\alpha$, with r_w the water surface reflection coefficient (typically -1) and α the two-way time delay between the source and the water surface [9, p. 251]. *Water column reverberations* are caused by the air gun source energy getting trapped in the water column and bouncing multiple times between the water surface and the sea floor before and/or after reflecting from the subsurface target [28]. They appear as multiple scaled, time-delayed repetitions of reflections and can be approximately modeled by multiplying the source wavelet with $\left(1 + \sum_{k=1}^{\infty} (r_w r_b z^\alpha)^k\right)^2$, where r_b is the sea floor reflection coefficient and α the two-way time delay between the water surface and the sea bottom [9, p. 251].

The statistical deconvolution problem is challenging even in the absence of noise because there exist multiple combinations of s and e that can give rise to the observation y in (1). The challenge lies in overcoming this ambiguity by exploiting prior knowledge about the structure of s and e. The next four subsections review various statistical deconvolution approaches and their assumptions. To understand the performances of these techniques, we refer the reader to [3,29,30].

3.1 Deconvolution Using Second-Order Statistics

Deconvolution using second-order statistics is based on two key assumptions.

1. Earth's reflectivity e is a stationary, white random process.
2. The unknown source s is minimum phase (that is, the energy of s is maximally concentrated at the start).

3.1.1 Source Signature Estimation

When the above two assumptions hold, the unknown s can be estimated from the spectrum—a second-order statistic—of the measurements y. Since e is white, we have (assuming no n) $|\mathcal{S}(f)|^2 \propto |\mathcal{Y}(f)|^2$. The minimum phase assumption ensures that the phase of $\mathcal{S}(f)$ is related to its magnitude. Hence a source signature estimate $\widehat{s}$ is given by the inverse Fourier transform of [9,31]

$$\widehat{\mathcal{S}}(f) = |\mathcal{Y}(f)| \exp\left(i 2\pi \theta(f)\right),$$

with $\theta(f)$

$$= -2\sum_{m=1}^{\infty} \sin 2\pi fm \int_0^{\pi} \cos um \log|\mathcal{Y}(u)|du. \tag{9}$$

If s is assumed to be zero-phase, then $\widehat{s}$ is given by the inverse Fourier transform of $|\mathcal{Y}(f)|$ (that is, $\theta(f) = 0$ in (9)). For non-white e (with known coloring), an additional prewhitening step is also required [32] before estimating $\widehat{s}$. The e is obtained by using the estimated s with techniques such as Wiener–Levinson estimation outlined in Section 2.

In practice, the noise n in y can adversely impact wavelet estimation (particularly, the phase of $\widehat{s}$) [33,34]. This effect can be alleviated by attenuating the noise in the receiver spectrum $|\mathcal{Y}(f)|$ by averaging across an ensemble of traces or by subtracting an estimated noise level from the $|\mathcal{Y}(f)|$ in (9) [33].

In reality, the earth's reflectivity, as sensed by the surface receivers, is not stationary but varies with depth; this effect is typically captured by an attenuation factor termed Q [3]. (Instead, the source s can also be viewed as being time-varying.) To offset this non-stationarity, often different time windows of y are individually subjected to wavelet estimation and deconvolution [3,29,33].

3.1.2 Predictive Deconvolution

In ghost reflection and water reverberation attenuation, typically, a good estimate of the time delay α between the primary and the multiple reflections is available. *Predictive deconvolution* uses the delay estimate to predict and deconvolve the multiple reflections from the measurements [19,35]. The technique adopts the Wiener–Levinson least squares approach to design a length-k *prediction filter* a that uses past samples of y to predict values of y at α time-samples into the future,

$$a(m) \circledast y(m) \approx y(m+\alpha). \tag{10}$$

Determining a involves solving a set of normal equations similar to (7). Since e is assumed to be white, and therefore unpredictable, the *prediction error* $y(m+\alpha) - a(m) \circledast y(m)$ provides a good estimate of the reflectivity e. The deconvolution is also referred to as *gapped deconvolution* when $\alpha > 1$ and as *spiking deconvolution* when $\alpha = 1$. When multiple measurements y are available to estimate e, multichannel predictive deconvolution [10] can provide significant improvements over its single channel counterpart [36,37].

In practice, predictive deconvolution provides improved resolution when the s is minimum phase but its performance deteriorates as the validity of the phase assumption deteriorates

[38]. The technique (in particular, spiking deconvolution) is also sensitive to noise n in y; this drawback can be alleviated by choosing an appropriate prewhitening factor (ϵ in (7)) [29]. The efficacy of predictive deconvolution also hinges on the stationarity of the primary-to-multiple time separation. Adapting the a with time/depth does not typically provide any significant improvements on data where the source-to-receiver distance (termed offset) is small or on stacked data, which approximates zero-offset data [29]. However, at large source to receiver offsets, adapting the α with time [39] or preprocessing the data to keep the required α constant [40] yields better predictive deconvolution results.

3.2 Deconvolution Using Higher-Order Statistics

Deconvolution using higher-order statistics does not make any prior assumptions about the phase of s; such phase assumptions are often inaccurate when s is long [22]. Instead, such methods assume that the e is a sparse series of spikes; from a statistical perspective, such methods assume that the e is non-Gaussian.

Such methods decompose y into e and s such that a higher-order statistic-based cost function that measures the "Gaussianity" (also called entropy) of the resulting e is minimized. That is,

$$\text{Estimate of } s = \min_{\tilde{s}} \text{ Entropy}\left(\tilde{s}^{-1} \circledast y\right). \tag{11}$$

For example, minimum entropy deconvolution (MED) [20,21] uses the negative of the *varimax* cost function as a measure of the entropy; that is,

$$\text{MED Estimate of } s = \min_{\tilde{s}} -\text{Varimax}\left(\tilde{s}^{-1} \circledast y\right), \tag{12}$$

$$\text{with Varimax}(e) = \frac{\sum_m (e(m))^4}{\left(\sum_m (e(m))^2\right)^2}. \tag{13}$$

See [22,41] for theoretical analysis of other such techniques and for a variety of other cost functions.

A higher-order statistics perspective can also be used to understand Bayesian approaches [42]. Bayesian techniques estimate e and s by maximizing a probability-based cost function that incorporates prior statistical information such as the non-Gaussianity and coherency of e across traces [43–45].

In practice, when the underlying e is sparse, such higher-order techniques perform well irrespective of the phase of the s [29]. However, for non-sparse e, the actual probability distribution of e strongly influences the final results [46]. Further, compared to second-order techniques, higher-order techniques are more adversely affected by the presence of noise n in (1) [41].

3.3 Homomorphic Deconvolution

Homomorphic deconvolution can potentially deconvolve non-minimum phase sources s even when the underlying e is not sparse. It assumes that the s and e map to different regions in the *cepstral* domain [24]. A signal can be transformed from the time domain to the complex-valued cepstral domain by first taking the Fourier transform, then taking the log operation [23]; that is, the cepstral transform (denoted by $\mathcal{C}$) of s is given by

$$\mathcal{C}(s) = \log(\mathcal{S}(f)) = \log\left(|\mathcal{S}(f)|\, e^{j\angle\mathcal{S}(f)}\right) = \log(|\mathcal{S}(f)|) + \angle\mathcal{S}(f), \quad (14)$$

with $\angle\mathcal{S}(f)$ denoting the phase of the complex-valued $\mathcal{S}(f)$. The cepstral domain is attractive for deconvolution because convolutions are transformed into additions. In particular, if we assume that the noise $\mathcal{N}(f)$ in (2) is negligible, then

$$\mathcal{C}(y) \approx \mathcal{C}(s) + \mathcal{C}(e). \quad (15)$$

The essence of homomorphic deconvolution lies in understanding and exploiting the characteristics of $\mathcal{C}(s)$ and $\mathcal{C}(e)$. For short band-limited sources $\mathcal{C}(s)$ tends to concentrate near the origin, and for water reverberations $\mathcal{C}(s)$ tends to be a periodic spike train. In contrast, $\mathcal{C}(e)$ is typically distributed over a large region [24].

The deconvolution estimate is obtained by using the following steps. First, transform y to the cepstral domain as in (15). Second, exploit the characteristics of $\mathcal{C}(s)$ and $\mathcal{C}(e)$ to filter $\mathcal{C}(y)$ and extract $\mathcal{C}(e)$. Finally, transform the extracted $\mathcal{C}(e)$ back into the time domain to obtain an estimate of e.

In practice, homomorphic deconvolution is typically employed only to estimate short, band-limited sources s, but not to obtain e or to deconvolve multiples [24]. For good results, phase unwrapping and editing are also necessary because n in the y (see (1)) perturbs the phase of the complex-valued cepstrum of y [47].

3.4 Surface-Consistent Deconvolution

Surface-consistent deconvolution attempts to estimate the s that is consistent with the surface geometry. To describe this further, we set up the following notations.

- $y_{i,j}$ denotes the receiver measurements when the source is located at i and the receiver is located at j.
- $s_{s,i}$ denotes the response due to the near-surface effects at the source location i.
- $s_{r,j}$ denotes the response due to the near-surface effects at the receiver location j.

- $s_{o,i-j}$ denotes the response due to offset (separation) between the source and the receiver.
- $e_{i+j/2}$ denotes the earth's response. The midpoint index $i+j/2$ between the source and the receiver locations, i and j respectively, is assumed to characterize the earth's response.
- $n_{i,j}$ denotes the noise corrupting $y_{i,j}$.

The underlying assumption is that the effective source for each measured trace can be decomposed into source, receiver, and offset effects; that is,

$$y_{i,j}(m) = \left(s_{s,i}(m) \circledast s_{r,j}(m) \circledast s_{o,i-j}(m)\right) \circledast e_{\frac{i+j}{2}}(m) + n_{i,j}(m). \qquad (16)$$

Above, $\left(s_{s,i}(m) \circledast s_{r,j}(m) \circledast s_{o,i-j}(m)\right)$ is the effective source probing the earth.

Given a collection of receiver measurements $y_{i,j}$ for various i and j, surface-consistent deconvolution first transforms the $y_{i,j}$ into the cepstral domain. Assuming that the noise $n_{i,j}$ is negligible, we have [25–27].

$$\mathcal{C}\left(y_{i,j}\right) \approx \mathcal{C}\left(s_{s,j}\right) + \mathcal{C}\left(s_{r,j}\right) + \mathcal{C}\left(s_{o,i-j}\right) + \mathcal{C}\left(e_{\frac{i+j}{2}}\right). \qquad (17)$$

Second, the earth's response and the three components comprising the source are estimated by solving the above equation as a linear inverse problem.

In practice, surface-consistent deconvolution can provide robust estimates even when the reflectivity e is not white [27], assuming that the surface-consistency conditions hold and that sufficient number of measurements are available [48]. However, when the number of observations are insufficient, the cepstral domain linear inverse problem based on (17) becomes severely ill-posed [27]. In such cases, the deconvolution needs to be performed carefully. Otherwise, the estimates could be severely corrupted by the noise $n_{i,j}$ in $y_{i,j}$ [48,49].

4 • Conclusions

This chapter illustrated the role of deconvolution in seismology and outlined various deconvolution techniques. Broadly, the deconvolution techniques were classified into two types: one where the source is known, termed deterministic deconvolution, and the other where the source is unknown, termed statistical deconvolution.

In deterministic deconvolution, the challenge is to control the noise amplification that results when the inverse of the

source signature is applied to the measured signal. Conventionally the desired signal is estimated by filtering the observations. Such techniques implicitly attenuate the appropriate frequency components of the noise that get amplified during inversion. Some recent techniques aim to attenuate the amplified noise in the wavelet or curvelet domain. These domains are better suited than the frequency domain to exploit the structure of typical reflectivity signals, and can thus provide improved estimates.

The statistical deconvolution problem is difficult even in the absence of noise because it contains multiple solutions. Since the problem is severely ill-posed, all statistical deconvolution techniques make some assumptions about the underlying source and reflectivity. Hence, a technique should be chosen only after understanding the assumptions satisfied by the deconvolution problem at hand. The most popular assumption, made by second-order statistics-based techniques, is that the unknown source is minimum-phase and that the reflectivity is white. These techniques are fairly robust to the presence of noise in the observation. In contrast, higher-order, statistics-based techniques assume that the desired signal is a sparse series of spikes. Such techniques are more sensitive to noise. Homomorphic deconvolution assumes that the source and earth reflectivity map to different regions in the cepstral domain. Homomorphic deconvolution works particularly well when the source is short and band-limited. Surface-consistent deconvolution assumes that the effective source for each trace in a collection can be decomposed into various components that are consistent with the surface geometry. It provides excellent estimates when the surface-consistency conditions hold, but its estimates are adversely affected by noise particularly when the number of traces in the collection (and thereby, the redundancy) is insufficient.

This chapter described many but not all deconvolution problems and solutions in seismology. For example, the problem of merging seismic datasets such as ocean bottom cable hydrophone and geophone data can be formulated as a multi-input single-output deconvolution (MISO-D) problem. Such problems are topics of active research; see [50,51] for a description of the problem and some optimal solutions.

New data types and novel acquisition scenarios are continuing to expand the demands on deconvolution. Hence, this signal processing technique will continue to play an important role in seismology and attract new research ideas.

Acknowledgments I would like to thank Warren Ross for his insightful comments.

References

1. G. M. Webster, ed., *Deconvolution*. Soc. Expl. Geophys., Geophysics reprint series; No. 1, 1978.
2. E. A. Robinson and O. M. Osman, eds., *Deconvolution 2*. Soc. Expl. Geophys., Geophysics reprint series; No. 17, 1996.
3. Öz Yilmaz, *Seismic Data Analysis: Vols. I and II*. Soc. Expl. Geophys., second ed., 2001.
4. J. M. Mendel, *Optimal seismic deconvolution: An estimation-based approach*. NY: Academic Press Inc., 1983.
5. T. Lay and T. C. Wallace, *Modern Global Seismology*. NY: Academic Press, 1995.
6. P. Shearer, *Introduction to Seismology*. Cambridge, UK: Cambridge University Press, 1999.
7. M. Bertero, D. Bindi, P. Boccacci, M. Cattaneo, C. Eva, and V. Lanza, "Application of the projected Landweber method to the estimation of the source time function in seismology," *Inverse Problems*, vol. 13, pp. 465–486, 1997.
8. R. Neelamani, H. Choi, and R. G. Baraniuk, "ForWaRD: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Trans. Signal Processing*, vol. 52, pp. 418–433, Feb. 2004.
9. E. A. Robinson and S. Treitel, *Geophysical Signal Analysis*. Englewood Cliffs, NJ, USA: Prentice Hall, 1980.
10. S. Treitel, "Principles of digital multichannel filtering," in *Deconvolution* (G. M. Webster, ed.), pp. 334–360, Soc. Expl. Geophys., 1978. Reprinted from Geophysics vol. 35, no. 5, pp. 785–811.
11. J. Agard and G. Grau, "Etude statistique de sismogrammes," *Geophys. Prosp.*, vol. 09, no. 04, pp. 503–525, 1961.
12. A. T. Walden and J. W. J. Hosken, "The nature of the non-Gaussianity of primary reflection coefficients and its significance for deconvolution," *Geophys. Prosp.*, vol. 34, no. 07, pp. 1038–1066, 1986.
13. S. Mallat, *A Wavelet Tour of Signal Processing*. NY: Academic Press, 1998.
14. E. Candès and D. Donoho, "Curvelets and curvilinear integrals," *J. Approx. Theory.*, vol. 113, pp. 59–90, 2000.
15. E. J. Candés and D. L. Donoho, "Recovering edges in ill-posed inverse problems optimality of curvelet frames," *Ann. Statist.*, vol. 30, pp. 784–842, 2000.
16. D. L. Donoho, "Nonlinear solution of linear inverse problems by Wavelet-Vaguelette Decomposition," *Appl. Comput. Harmon. Anal.*, vol. 2, pp. 101–126, 1995.
17. G. Hennenfent, F. Herrmann, and R. Neelamani, "Seismic deconvolution revisited with curvelet frames," in *EAGE 67th Conference & Exhibition Proceedings*, June 2005.
18. K. R. Castleman, *Digital Image Processing*. NJ: Prentice Hall, 1996.
19. E. A. Robinson, "Predictive decomposition of seismic traces," in *Deconvolution 2* (E. A. Robinson and O. M. Osman, eds.), pp. 60–71, Soc. Expl. Geophys., 1996.

20. R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 26, pp. 21–35, 1978.
21. M. D. Sacchi and T. J. Ulrych, "Nonminimum-phase wavelet estimation using higher order statistics," *The Leading Edge*, vol. 19, no. 01, pp. 80–83, 2000.
22. D. L. Donoho, "On minimum entropy deconvolution," in *Applied Time-Series Analysis II* (D. F. Findley, ed.), pp. 556–608, NY: Academic Press, 1981.
23. A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
24. B. Buttkus, "Homomorphic filtering: Theory and practice," in *Deconvolution 2* (E. A. Robinson and O. M. Osman, eds.), pp. 209–245, Soc. Expl. Geophys., 1996.
25. M. T. Taner and F. Koehler, "Surface consistent corrections," *Geophysics*, vol. 46, no. 01, pp. 17–22, 1981.
26. S. A. Levin, "Surface-consistent deconvolution," in *Deconvolution 2* (E. A. Robinson and O. M. Osman, eds.), pp. 366–376, Soc. Expl. Geophys., 1996.
27. G. Cambois and P. L. Stoffa, "Surface-consistent deconvolution in the log/Fourier domain," in *Deconvolution 2* (E. A. Robinson and O. M. Osman, eds.), pp. 377–394, Soc. Expl. Geophys., 1996.
28. M. M. Backus, "Water reverberations - Their nature and elimination," in *Deconvolution* (G. M. Webster, ed.), pp. 136–164, Soc. Expl. Geophys., 1978. Reprinted from Geophysics vol. 24, no. 2, pp. 233–261.
29. A. Jurkevics and R. Wiggins, "A critique of seismic deconvolution methods," *Geophysics*, vol. 49, no. 12, pp. 2109–2116, 1984.
30. L. R. Lines and T. J. Ulrych, "The old and the new in seismic deconvolution and wavelet estimation," *Geophys. Prosp.*, vol. 25, no. 03, pp. 512–540, 1977.
31. R. E. White and P. N. S. O'Brien, "Estimation of the primary seismic pulse," *Geophys. Prosp.*, vol. 22, no. 04, pp. 627–651, 1974.
32. M. Saggaf and E. A. Robinson, "A unified framework for the deconvolution of traces of nonwhite reflectivity," *Geophysics*, vol. 65, no. 5, pp. 1660–1676, 2000.
33. B. J. Newman, "Deconvolution of noisy seismic data – A method for prestack wavelet extraction," *Geophysics*, vol. 51, no. 01, pp. 34–44, 1986.
34. A. J. Berkhout, "Least-squares inverse filtering and wavelet deconvolution," *Geophysics*, vol. 42, no. 07, pp. 1369–1383, 1977.
35. K. L. Peacock and S. Treitel, "Predictive deconvolution – Theory and practice," in *Deconvolution* (G. M. Webster, ed.), pp. 221–235, Soc. Expl. Geophys., 1978. Reprinted from Geophysics vol. 34, no. 2, pp. 155–169.
36. M. T. Taner, R. F. O'Doherty, and E. Baysal, "Multiple suppression by multi-channel predictive deconvolution," in *53rd Mtg.*, pp. 22–23, Eur. Assn. of Expl. Geophys., 1991.

37. M. T. Taner, R. F. O'Doherty, and F. Koehler, "Long period multiple suppression by predictive deconvolution in the x-t domain," *Geophys. Prosp.*, vol. 43, no. 04, pp. 433–468, 1995.
38. T. J. Ulrych and T. Matsuoka, "The output of predictive deconvolution," *Geophysics*, vol. 56, no. 03, pp. 371–377, 1991.
39. S. C. Hornbostel, "Multiple suppression with time-varying gaps and maximized polarity reversals," in *68th Ann. Internat. Mtg*, pp. 1142–1145, Soc. Expl. Geophys., 1998.
40. M. Schoenberger and L. M. Houston, "Stationarity transformation of multiples to improve the performance of predictive deconvolution," *Geophysics*, vol. 63, no. 02s, pp. 723–737, 1998.
41. A. T. Walden, "Non-Gaussian reflectivity entropy and deconvolution," *Geophysics*, vol. 50, no. 12, pp. 2862–2888, 1985.
42. G. B. Giannakis and J. M. Mendel, "Entropy interpretation of maximum-likelihood deconvolution," *Geophysics*, vol. 52, no. 12, pp. 1621–1630, 1987.
43. M. Lavielle, "2-D Bayesian deconvolution," in *Deconvolution 2* (E. A. Robinson and O. M. Osman, eds.), pp. 584–594, Soc. Expl. Geophys., 1996.
44. K. Kaaresen and T. Taxt, "Multichannel blind deconvolution of seismic signals," *Geophysics*, vol. 63, no. 06, pp. 2093–2107, 1998.
45. P. Sheng, B. White, B. Nair, and S. Kerford, "Bayesian deconvolution of gamma-ray logs," *Geophysics*, vol. 52, no. 11, pp. 1535–1546, 1987.
46. R. Wiggins, "Entropy guided deconvolution," in *Deconvolution 2* (E. A. Robinson and O. M. Osman, eds.), pp. 265–271, Soc. Expl. Geophys., 1996.
47. P. L. Stoffa, P. Buhl, and G. M. Bryan, "The application of homomorphic deconvolution to shallow-water marine seismology – Part I: Models," *Geophysics*, vol. 39, no. 04, pp. 401–416, 1974.
48. S. A. Levin, "Lessons in surface-consistent deconvolution," in *58th Ann. Internat. Mtg*, pp. 718–722, Soc. Expl. Geophys., 1988.
49. F. Kirchheimer and R. Ferber, "Robust surface-consistent deconvolution with noise suppression," in *71st Ann. Internat. Mtg*, pp. 1831–1834, Soc. Expl. Geophys., 2001.
50. M. Deffenbaugh and R. Neelamani, "OBC multiple suppression with the Texas Two-step," in *75th Ann. Internat. Mtg*, SPMUL 2.8, Soc. Expl. Geophys., 2005.
51. R. Neelamani, M. Deffenbaugh, and R. G. Baraniuk, "Texas two-step: A framework for optimal multi-input single-output deconvolution," IEEE Transactions on Image Processing, vol. 16, no. 11, pp. 2752–2765, Nov. 2007.

88

Seismic Migration Imaging

Michael Fehler

Massachusetts Institute of Technology, Cambridge, USA

1 • Introduction

The goal of seismic migration imaging is to gain information about the geometry of subsurface structures and thus gain insight into the geology and spatial distribution of rock-types within the imaged region. The geometrical information obtained is in the form of a reflectivity map that shows the locations in the subsurface where seismic waves are reflected. Reflections occur where rock properties change. These changes may be in rock type, the volume and geometry of pores within the rock, or the fluid type within the pores (e.g., oil vs. water). Recent advances in migration imaging have provided more information than just the geometry of subsurface structures. For petroleum exploration, information about the contrast in rock properties across a boundary is also important, and approximate information about these contrasts can be obtained from seismic reflection data using various methods.

Imaging using seismic reflection data continues to be an area where significant advances are being made [1,2]. The advances have been driven by improved data acquisition methods that provide more data of better quality. Newer data sets typically have higher frequency content and are collected with source and receiver geometries that are more advantageous for imaging. The

availability of faster computers, better algorithms, and better data have lead to increasingly reliable migration imaging.

2 • The Migration Problem

An intuitive way to visualize migration is to think of the wavefield propagating into the earth from a source located at the surface. As waves arrive at velocity or density discontinuities they are reflected and/or scattered. The waves that are reflected back to the earth's surface are recorded by receivers. Migration is the process of backprojecting the reflected waves recorded at the surface into the earth to find when and where they were generated.

Migration involves the use of an *imaging condition*, which specifies how we decide where the reflected waves were generated, and thus where reflectors are located in the subsurface. A common imaging condition is a convolutional one, which says that the reflected waves were generated at the time and location within the earth where a wavefield propagated forward in time from the source would intersect a wavefield backprojected (propagated backward in time) from the receivers [3].

Until very recently, most seismic migration has been done using the assumption that the earth is an acoustic medium where shear waves do not propagate. This assumption has worked well because (1) the seismic sources used for reflection imaging are sources of strong compressional waves and generate only weak shear waves, (2) migration imaging has been applied in regions that are generally laterally homogeneous (e.g., velocity varies weakly with horizontal position), and (3) only nearly vertical reflections from discontinuities are recorded. There has been recent interest in imaging using shear waves but, given the challenges of imaging with only compressional waves, the use of shear waves is still a research topic.

Consider the seismic waves recorded by a receiver located on the earth's surface in a region where there is only one subsurface interface. The source is at the surface. A single reflected arrival will be recorded at a time equal to the combined traveltime of a ray from the source to the point of specular reflection on the interface and then to the receiver located at a distance d from the source. Considering only the traveltime information, the reflection could have occurred anywhere along an ellipse (in 2D) having foci located at the positions of the source and receiver and having major axis length, a, equal to the traveltime multiplied by the medium velocity divided by two and minor axis length, b, where $b^2 = a^2 - (d/2)^2$. By recording data using a spatial distribution of sources and receivers, a large number of ellipses can

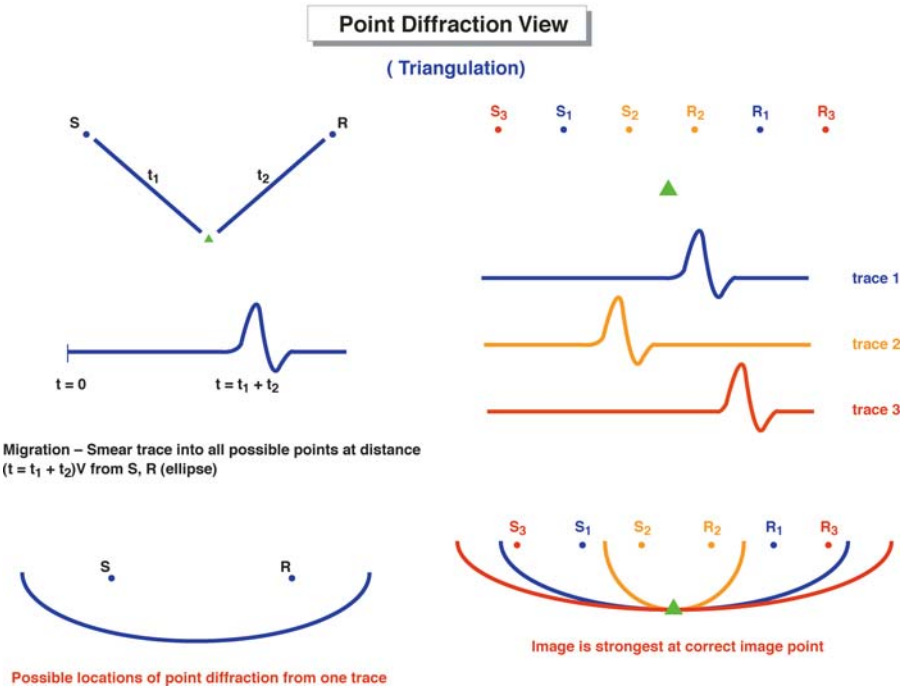


FIGURE 1 Graphical illustration of migration. On the *left side*, the wavefield generated by a source is detected by one receiver. The simplified model contains one point scatterer that produces a single pulse on the received wavefield. The distribution of possible scattering points within the subsurface defines an ellipse in 2D. Using data from a single station and a single source, the location of the scatterer on the ellipse cannot be determined. Using data from multiple shots recorded at multiple receivers, as shown on the *right*, multiple ellipses of possible scattering locations can be inferred. The point where the ellipses intersect is taken as the location of the scattering point.

be drawn and the location of the interface determined as the trend where multiple ellipses intersect. Figure 1 illustrates this process. On the upper left hand side, a seismic source is placed at S and the resulting wavefield is recorded at R. If there is a single point-scatterer within the earth that is located at the position of the triangle, the reflection waveform will consist of one pulse that occurs at the time $t_1 + t_2$, where t_1 is the traveltime from the source to the scatterer and t_2 is the traveltime from the scatterer to the receiver. (We neglect the portion of the seismogram that is caused by the pulse that travels from the source directly to the receiver. This is accomplished by a process called muting.) Without any other information, the scatterer could be located

anywhere on an ellipse whose foci are at the source and receiver and whose major axes has length $(t_1 + t_2)v$ where v is the medium velocity. By recording the scattered pulse at a number of receivers and from a number of sources, as shown in the right side of Figure 1, a series of ellipses can be drawn. The location of the scatterer is given by the point where the ellipses intersect.

3 • Time Migration

Prior to the availability of computers, migration was implemented as a graphical procedure. Numerous mechanical devices were created to perform migration for simple models of subsurface velocity structure, such as a linear increase in velocity with depth. Papers describing some of these methods can be found in [4].

As computers became available, migration was implemented for the so-called zero-offset data, which simulated data collected when the source and receiver are coincident at the surface, using either a constant velocity model or one with a linear velocity change with depth. This procedure, known as *poststack time migration*, is computationally very fast and is still used in areas of relatively simple geology or to make an initial assessment of a dataset. Zero-offset data can be obtained by stacking, which involves time-shifting or stretching and then summing multiple traces of recorded data to get a single trace that mimics what would have been recorded if the source and receiver had been at the same position. There are numerous approaches for stacking data. Most often, traces are stacked together which have common subsurface reflection points, called a *common depth point* or *common midpoint* stack (see Chapter 85: Seismic Data Acquisition).

The numerical implementation of poststack time migration consists of smearing each portion of the stacked zero-offset data onto a circle (for a homogeneous medium) centered at the position of the source/receiver. The circle radius is defined by the product of the traveltime of the point on the trace that is smeared and the assumed medium velocity. Constructive interference leads to the locations of strong reflectors having high amplitudes after the smearing process is complete. The use of a simplified velocity model in time migration leads to poor image quality for regions where geology is complex.

Stacking of seismic data from individual shots into a zero-offset trace for each shot location greatly reduces the data volume, increases signal-to-noise ratio compared to that of a single receiver

at the source position, and speeds migration processing. Stacking also reduces image quality because much information is lost from the individual waveforms during the stacking procedure. To avoid the stacking, a procedure known as *prestack time migration* is used. This approach uses the simplified earth structure for the migration but raw trace data are used instead of stacked data. The traveltimes between sources, image points, and receivers are calculated for the simplified velocity model and the trace data are projected onto ellipses for homogeneous media or deformed ellipses for inhomogeneous media. Prestack time migration generally provides better images of the earth but the image quality is still limited by the simplifying assumption about the velocity structure.

4 • Depth Migration

When geology becomes more complex so that time migration cannot provide clear images of the earth's subsurface, a procedure called *depth migration* is employed. Depth migration uses a 3D representation of the earth's velocity structure to backpropagate the receiver wavefield into the earth and find the locations where the reflections occurred. The first computational approach to depth migration imaging is called ray-based Kirchhoff migration [5]. Ray-based Kirchhoff migration backprojects the wavefield using a backprojection operator, or migration operator, that is constructed using ray theory [6,7,8,9], which is a high-frequency approximation to the wave equation. We can think of this operator as telling us how to smear a given reflection recorded at a given receiver into the heterogeneous medium; for example, instead of smearing onto an ellipse like we did in a homogeneous medium, we smear along a deformed ellipse whose shape is determined from ray tracing through the medium. Ray-based methods, based on a high-frequency approximation, work well when the dominant wavelength of the imaging wavefield is smaller than the typical scale of heterogeneity in the medium. Early implementations of ray-based Kirchhoff migration used only the fastest arrival between an image point and a source or position on the surface. More recent implementations include the use of multiple-valued traveltime tables, ray amplitudes, and ray phases that can be calculated from various ray-tracing implementations.

The high-frequency approximation inherent in ray tracing can be eliminated by using wave-equation imaging [10,11], based on implementations of solutions of the wave equation, one-way

wave equation, and approximations to the Lippmann–Schwinger equation [12,13]. Wave-equation methods take account of wave phenomena such as focusing, defocusing, and diffraction that are important in many geological environments where imaging is used for petroleum exploration. Most wave-equation imaging is based on the one-way or paraxial wave equation and thus reverberations are ignored. In addition, the amplitude of the back-projected wavefield may not be reliable when there is significant backscattering.

Migration imaging requires some initial estimation of the earth's velocity structure to use for backprojecting the seismic trace data. This information is usually determined from the data using a process called velocity model building. One approach to velocity model building is to perform traveltime tomography. Velocity models are often improved by iterative application of migration followed by velocity analysis to improve the velocity model, repeating until the migrated image comes into focus.

Figure 2 shows the velocity model for a 2D model that was developed as being representative of a region in Nigeria. The model has considerable complexity due to the steeply dipping

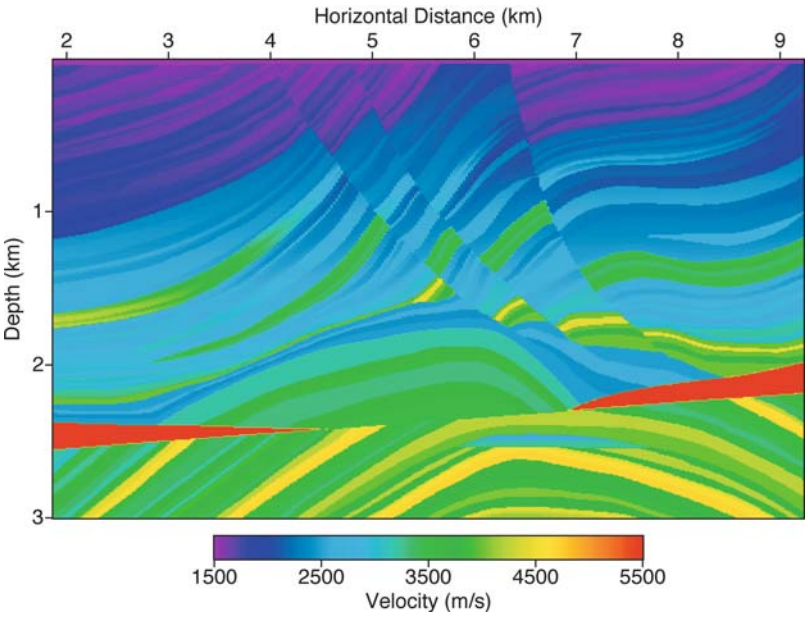


FIGURE 2 Velocity model for a complex structure. The model is considered to be acoustic so on shear waves are generated within the model.

strata and the high-velocity region near the bottom, below which petroleum reserves were considered to reside. A seismic data acquisition experiment was conducted by doing finite difference calculations over this model to produce trace data for an array of source and receiver locations. These data were subsequently migrated. To aid the migration, traveltimes were calculated from points on the surface into a grid of points in the subsurface using a wavefront-construction ray-tracing code [14]. The Se traveltimes were highly variable due to the complexity of the velocity model. Figure 3 shows a depth-migrated image obtained by migrating the trace data using a wave-equation-based approach like the one described in [12]. This dataset provides a good test of a migration code; many approaches have yielded poor-quality images of the structure. The wave-equation-based approach provides a good image.

While seismic migration imaging is used mostly for petroleum exploration, there have also been applications of various types of migration imaging in basic studies of earth structure. Such studies have been made to investigate deep earth structure and large-scale lithospheric structure using waveforms from teleseisms as sources [15,16].

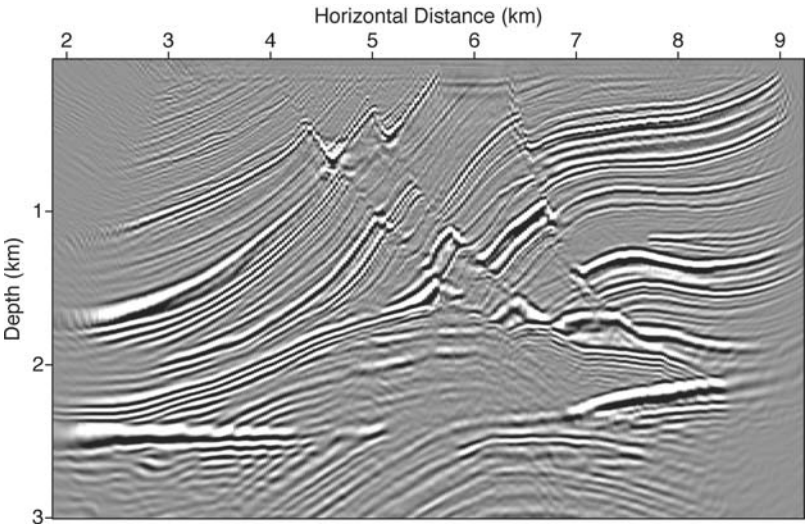


FIGURE 3 Image obtained from migrating finite-difference–generated seismograms calculated for the velocity model shown in Figure 2. Migration was performed using a wave-equation–based approach like the one described in [12].

References

1. Gray, S., J. Etgen, J. Delinger., and D. Whitmore., Seismic migration problems and solutions, *Geophysics* **66**, 1622–1640 (2001).
2. Fehler, M. and L. Huang., Modern imaging using seismic reflection data, *Annu. Rev. Earth Planet. Sci.* **30**, 259–284 DOI: 10.1146/annurev.earth.30.091201.140909 (2002).
3. Claerbout, J. F., Toward a unified theory of reflector mapping, *Geophysics* **36**, 467–481 (1971).
4. Gardner, G., *Migration of Seismic Data, Geophysics Reprint Series Number* 4, Society of Exploration Geophysicsits, Tulsa, Ok (1985).
5. Schneider, W., Integral formulation for migration in two and three dimensions, *Geophysics* **43**, 49–76 (1978).
6. Cerveny, V., *Seismic Ray Theory*, Cambridge University Press, Cambridge, UK (2001).
7. Vinje, V., Iversen, E., and Gjoystdal H., Traveltime and amplitude estimation using wavefront construction. *Geophysics* **58**, 1157–1166 (1993).
8. Vidale, J., Finite-difference calculations of travel times. *Bull. Seismol. Soc. Am.* **78**, 2062–2076 (1988).
9. Um, J. and Thurber C., A fast algorithm for two-point seismic ray tracing. *Bull. Seismol. Soc. Am.* **77**, 972–986 (1987).
10. Stolt, R., Migration by Fourier transform, *Geophysics* **43**, 23–48 (1978).
11. Gazdag, J., Wave-equation migration with the phase-shift method, *Geophysics* **43**, 1342–1351 (1978).
12. Huang, L., M. Fehler, P. Roberts., and C. C. Burch., Extended local Rytov Fourier migration method. *Geophysics* **64**, 1535–1545 (1999).
13. Huang, L., M. Fehler., and R. Wu, Extended local Born Fourier migration method. *Geophysics* **64**, 1524–1534 (1999).
14. Fehler, M., S. Hildebrand., D. Alde., and L. Huang., Wavefront Construction Kirchhoff Migration with Ray-Amplitude Corrections (expanded abstract), European Association of Geoscientists and Engineers in Florence, Italy, (2002).
15. Flanagan, M. P. and P. M. Shearer., Topography on the 410-km seismic velocity discontinuity near subduction zones from stacking of sS, sP, and precursors, *J. Geophys. Res.* **103**, 21165–21182, 1998.
16. Aprea, C., S. Hildebrand., M. Fehler., L. Steck., W. S. Baldridge., C. Thurber., and W. Lutter., 3-D Kirchhoff migration: imaging of the Jemez volcanic field using teleseismic data, *J. Geophys. Res.* **107**, 2247 (2002).

89

Geophysical Parameter Estimation

Max Deffenbaugh

ExxonMobil Research and Engineering Company, Annandale, NJ, USA

1 • Introduction

In parameter estimation, properties of interest are quantified based on a set of related observations. Examples of geophysical parameter estimation include determining the porosity of an oil reservoir from seismic reflection amplitudes, determining the degree of velocity heterogeneity in the mantle from the duration of the seismic coda, or classifying the source of a particular seismic event based on various signal attributes. Formally, all these problems involve determining *parameters* (θ) of a conditional probability distribution based on *observations* (y) drawn from the distribution $y \sim f(y|\theta)$.

Good solutions to practical parameter estimation problems involve four steps: (1) selecting a model, (2) assessing the available information, (3) selecting an estimator, and (4) validating estimator performance. This chapter describes techniques which seismologists and geophysicists have found useful for each of these steps. In practice, the four steps are best applied iteratively, starting with the simplest possible models and estimators to get a "feel" for the problem, refining the model and/or the estimator as needed, and then stopping when the required estimate accuracy is either achieved or shown to be unattainable.

2 • Model Selection

The most difficult step in parameter estimation is choosing a suitable model. The model is a conditional probability distribution of the observations given the parameters, $f(y|\theta)$. There are several trade-offs to consider. (See Burnham and Anderson, 2003 for more discussion.)

1. *Physics vs. empiricism*. A model can be derived from the physics which relates observations and parameters, or it can be based on empirical relationships extracted from data. Physics-based models are preferred when all the relevant physics is well-understood and leads to a tractable model. Empirical models are preferred when the physics is too complicated but observations are available for instances covering the full range of possible parameter values. Empirical models simply mimic an observed correlation without providing physical insight, so it is difficult to say whether an empirical model will be accurate under new conditions.
2. *Accuracy vs. tractability*. A good model both adequately captures the relationship between the observations and the parameters *and* is computationally tractable. A model which is complicated numerically or analytically hinders all aspects of estimator design and testing. The benefits it affords in accuracy should be weighed against the inconvenience it creates.
3. *Overfitting vs. underfitting*. The goal of modeling is not to fit a certain set of observations, but rather to fit the probability distribution from which those observations were drawn. A model with too few parameters may not adequately capture the underlying probability distribution (underfitting), but a model with too many parameters may canonize the variability within one observation set as a fixed feature of the model (overfitting). An extreme case of overfitting would be a probability distribution made of equal delta functions at each sample value – a perfect fit to the observations, but almost certainly not the true distribution that produced the observations.

2.1 Probability Distributions

Many models in seismology are based on the probability distributions described in this section. The parameters of the distributions (e.g., mean or variance) are often complicated functions of the geophysical parameters of interest. For example, one sample in a seismic data trace may have a mean value which is a function of the seismic wave velocity and density at a point in the subsurface and a variance which depends on both environmental noises (wind, waves, traffic) and source-related noises neglected in the function for the mean value (interference from other wave types,

scattering from heterogeneities not included in the velocity and density model). Some of the noises in seismic data are described in Bormann (2002), Wilson et al. (2002), and Olhovich (1964).

2.1.1 Central Limit Distributions

Random variables measured in nature are often the combination of many independent events drawn from the same probability distribution. The central limit theorem tells us that when the individual events combine additively and their common distribution has finite variance, their sum has a *Gaussian* distribution. Background noise often consists of the sum of many independent contributions of similar size and is therefore often modeled using a Gaussian distribution in both whole earth seismology (Peterson, 1993) and exploration seismology (Sax, 1968). The Gaussian distribution is a good starting point for modeling symmetric distributions of data, such as reflection coefficients in sequences of sedimentary rock (Todoeschuck et al., 1990).

When many independent events combine multiplicatively, each increasing or decreasing the product by some percentage, a *lognormal* distribution results. The logarithm of the product is equal to the sum of the logarithms of the individual events, so the logarithm of the product has a Gaussian distribution according to the central limit theorem; this means the product has a lognormal distribution. A sand grain is shaped by numerous small collisions, each reducing its size by a random percentage. Thus the size distribution of sand grains tends to be lognormal. (Kondolf and Adhikari, 2000). The size of an oilfield is the result of numerous sedimentary and geochemical processes, the variability of which might be regarded as increasing or decreasing the field size by a percentage compared with neighboring fields. Thus the size of oilfields within a geologic basin is often modeled by a lognormal distribution. (Greenman and Fryer, 1996) The lognormal distribution is a good starting point for modeling properties which are always positive and vary by at least an order of magnitude.

When many independent heavy-tailed events combine additively, a *Levy-stable* distribution results. Heavy-tailed means that for large values of y, $f(y) \sim |y|^{-(\alpha+1)}$ with α between 0 and 2, so that these distributions have infinite variance. If $\alpha < 1$, the distribution also has an undefined expected value, but can still be a legitimate probability distribution (i.e., $f(x)$ is positive and integrates to 1) so long as $\alpha > 0$. The Levy-stable distribution is heavy tailed with the same value of α as the contributing effects. An analytical form is available for the characteristic function of the distribution, but not for the distribution itself for most values of α. As $\alpha \longrightarrow 2$, the Levy-stable distribution approaches the Gaussian distribution, and for $\alpha = 1$, the Levy-stable distribution is the

Cauchy distribution. The Levy-stable distribution has been used to model seismic reflection amplitudes (Painter et al., 1995; Painter, 2003). The stress at a point within a random distribution of point defects can be approximated using a Cauchy distribution (Zolotarev and Strunin, 1971). In seismogenic regions, perturbations to the local stress field cause relative rotations between the focal mechanisms of earthquakes. These relative rotations often have a Cauchy distribution (Kagan, 1992). When estimating parameters of a heavy tailed distribution, a quantile-based method or other robust technique should be used instead of moments, since enough finite moments may not exist.

2.1.2 Power-Law Distributions

Many processes in nature have no characteristic scale; their statistical behavior remains the same at all scales. A probability distribution is scale-invariant under truncation if, when we exclude all values less than b, we find that the probability of obtaining a value equal to c^*b is the same regardless of b. The *Pareto* distribution has this important property and therefore occurs frequently in nature in describing self-similar or fractal processes. Many of the important physical processes in seismology tend to produce power-law trends. Fracturing processes tend to have power-law properties, which lends a Pareto distribution to earthquake magnitude (Scholz, 1997) as well as fracture-generated reservoir permeability (Laubach et al., 2000; Marrett et al., 1999). Statistics of floods, earthquakes, and landslides are famously power law (Turcotte, 1997), and they tend to impart a power-law size distribution to the sediment deposits they form (Beattie and Dade, 1996).

2.1.3 Likelihood of Events

The likelihood of random events is often described by the probability distribution of the time between events. The *exponential* distribution is a special case, where the time until the next event is always independent of the time since the last event. In other words, the next event becomes neither more nor less likely as time passes since the last event. This unique "memoryless" property of the exponential distribution is discussed in reference to the likelihood of the next earthquake (Sornette and Knopoff, 1997). The *Laplace* distribution is a double-sided exponential distribution.

When the time between events has an exponential distribution with scale parameter (expected time between events) s, the probability of n events occurring in time s has a *Poisson* distribution with parameter $a = s$. Poisson distributions can describe the number of earthquake shocks in a fixed period of time (Madhava-Rao and Kaila, 1986) or the net effect of many small contributions

such as the size of accumulations of carbonate sedimentary bodies (Wilkinson et al., 1999).

Sometimes the next event becomes more or less likely as time increases since the last event. The *Weibull* distribution can model both uniformly increasing ($a > 1$) or uniformly decreasing ($a < 1$) likelihood with time. This distribution, which includes the exponential as a special case ($a = 1$), is often used to model the time between earthquakes (Scholz, 1997) or volcanic eruptions (Ho, 1991) when the possibility of non-random timing is investigated. Such non-random timing is often associated with the failure of materials (Cox and Oakes, 1984). The Weibull distribution is sometimes used as a parameterized family of distributions, as it includes the exponential ($a = 1$) and Rayleigh ($a = 2$) as special cases and can approximate the Gaussian ($a = 3.4$). The *Logistic* distribution has a hazard function (likelihood of the next event) that first increases then later decreases with time since the last event. Conner et al. (2003) used this distribution to model the time between volcano eruptions where competing physical processes caused the gas bubble pressure to initially increase and later decrease with time since the last eruption.

Instead of describing times between events, we may simply assign a probability that an event occurs in a particular trial. The *Bernoulli* distribution describes the likelihood of a binary event, that is, one which occurs in a given trial with probability p. The *binomial* distribution gives the probability that n such events will occur in (fixed) N Bernoulli trials. The Binomial distribution has been used to test whether seismic waves from large earthquakes trigger additional earthquakes (called "dynamic triggering"). The number of days with elevated earthquake activity in a period following a large earthquake was compared to a baseline using the binomial distribution to see whether the change was statistically significant. (Pankow, et al., 2004). The *negative binomial* distribution gives the probability that the (fixed) Nth event will occur on the $N+n$th trial, i.e., after n failures. Sometimes the negative binomial distribution is called the *Pascal* distribution when N is strictly a natural number and the *Polya* distribution when N is allowed to be a real number greater than zero. The negative binomial distribution is a better model than the Poisson distribution for events which tend to cluster in time, such as shallow earth quakes (Rao and Kaila, 1986). The *geometric* distribution is the special case of the negative binomial distribution where $n = 1$. Since the probability of success for each event is independent of the other events, the geometric distribution is memoryless and is the discrete analog of the continuous exponential distribution.

2.1.4 Distributions of Limited Domain

Most natural properties have physical upper and lower limits on their value. If their likely values are far from the limits, we may model these properties with an infinite tail distribution. For example, the reflection coefficient is necessarily between −1 and 1, but in sequences of sedimentary rocks is usually between −0.2 and 0.2, so we model it with infinite tail distributions and accept as modeling error the fact that our distribution admits non-physical values with some negligible probability. Other properties, like angles or percentages, may have substantial likelihood near the edges of their domain of support. In these cases, a probability distribution with a bounded domain may be required.

The *uniform* distribution describes random variables which assume with equal probability a value in a bounded range. The uniform distribution is often used to describe phases or angles. It also is the probability distribution of digitization (round-off) noise in analog-to-digital converters, where a continuous valued analog voltage (e.g., from a geophone) is represented by one of 2^{16} or 2^{24} discrete integers.

The *triangle* distribution is specified by a maximum and minimum value as well as a most likely value in that range. The probability increases linearly from zero at the extreme values to the most likely value, forming a triangle. It is often used in elicitation, when experts are asked to specify a maximum, minimum, and most likely value for some property based on their experience.

2.1.5 Parameterized Families of Distributions

The *beta* distribution has support on the interval [0,1] and is a very versatile distribution for fitting data that is limited to that interval. The *generalized beta* distribution adds a location and scale parameter to fit distributions on any bounded domain, which makes it useful for approximating even infinite tail distributions (Gupta and Nadarajah, 2004; Karian and Dudewicz, 2000).

The *generalized lambda* distribution is another bounded-domain distribution which can provide good approximations to most common distributions (Karian and Dudewicz, 2000). Its quantile function is easily specified, so it can be conveniently fit to data by matching four quantiles of the distribution to the same quantiles of the sampled data. A good choice of metrics for fitting data are the median, the difference between the 90th and the 10th percentile, left–right tail weight ratio (ratio of distance from median to 10th percentile to distance from 90th percentile to the median), and the tail weight factor (ratio of the interquartile

range to the interdecile range). Sometimes it is necessary to try other percentile metrics if those do not produce quick convergence to a good fit. Since the percentile function is specified, it is convenient to generate samples numerically from the distribution. Uniformly distributed samples on the interval [0,1] are generated, and the sample value corresponding to that quantile is found using the quantile function (Dudewicz, 1999).

The *generalized gamma* distribution is useful for fitting many infinite tail distributions. Special cases of this distribution include the gamma, chi-square, exponential, and Gaussian distributions. There is a convergence problem with estimating the parameters of the common form of the distribution, but a better parameterization is available (Lawless, 2002).

The *maximum entropy* distribution can be viewed as a parameterized family of distributions. It is the distribution with the largest entropy subject to constraints on the first N moments (e.g., for $N = 2$, it equals the Gaussian distribution). Solving this constrained maximization using the method of Lagrange multipliers produces the distribution form in Table 1 (Jaynes, 1957). The method of moments (Section 2.2.2) may then be applied to fit the parameters $\lambda_1, \lambda_2, \ldots, \lambda_N$ of the maximum entropy distribution to a set of observations. This involves some numerical computation, but the distribution it produces is a member of the exponential family (Section 2.1.7), which makes it a convenient choice.

2.1.6 Functions of Random Variables

Certain functions of random variables have distributions with a convenient analytical form.

Sums

A sum of random variables has a mean value equal to the sum of the individual means. If the random variables are *independent*, their sum has a variance equal to the sum of the individual variances and a distribution equal to the convolution of their respective distributions. If enough random variables are summed together, the distribution of their sum approaches one of the central limit distributions (see Section 2.1.1).

Products

Taking a logarithm turns the product of positive-valued random variables into a sum of logarithms of the variables, after which rules for sums of random variables can be applied. For example, if enough random variables with finite variance are multiplied

together, the logarithm of their product has a Gaussian distribution which means their product has a lognormal distribution (see Section 2.1.1).

Functions of Gaussian Random Variables

Many useful functions of Gaussian random variables have named distributions. In Table 2, some of these functions are given, where z_i are independent Gaussian random variables with means μ_i and variances σ_i^2. The corresponding distribution is given, along with constraints on the parameters of z_i and expressions for the parameters of the named distribution as it appears in Table 1.

The square root of the sum of the squared values of N independent zero-mean same-variance Gaussian random variables has a *chi* distribution with N degrees of freedom. Special cases of the chi distribution are the *Rayleigh* ($N = 2$) and *Maxwell* ($N = 3$) distributions. The envelope of a narrowband signal with independent Gaussian real and imaginary parts has a Rayleigh distribution, and the Euclidian distance in three dimensions has a Maxwell distribution when the locations along each axis have a Gaussian distribution. The sum of the squared values of N independent zero-mean, unit-variance Gaussian random variables has a *chi-square* distribution with N degrees of freedom. The *gamma* distribution is a generalization of the chi-square distribution where the number of degrees of freedom is not an integer.

Certain measures of signal coherence used in array signal processing have distributions from Table 2. The F-statistic has the central or non-central F-distribution depending on whether a signal is present at the array (Douze and Laster, 1979; Blandford, 1974). The semblance has a central or non-central beta distribution (Kimball and Scheibner, 1998).

Unfortunately, the non-central distributions all have an awkward analytical form, being expressed as infinite sums which converge slowly. They are not included in Table 1.

Functions of Uniform Random Variables

Certain functions of uniform random variables come up frequently in nature. If x has a uniform distribution between $-\pi/2$ and $\pi/2$, then $y = \tan(x)$ has a Cauchy distribution. When a narrowband signal is formed by the interference of many independent scattering events, each with uniformly distributed phase from 0 to 2π, the real and imaginary parts of its envelope are independent and Gaussian, and the magnitude of the envelope is Rayleigh. Fehler et al. (2000) note that the envelopes of seismic codas have a Rayleigh distribution.

Functions of uniform random variables are important in random number generation. Most computer languages

implement a uniform random number generator, and various algorithms are available to map the uniformly distributed values into the distribution of interest. More information on generating random numbers of any distribution is found in Gentle (2003).

2.1.7 The Exponential Family

A probability distribution belongs to the N-parameter *exponential family* of distributions if it can be written in the exponential family form given in Table 1. For example, the Gaussian distribution with mean t and variance s^2 is a member of the two-parameter exponential family, with $A_1 = t/s^2$, $B_1 = y$, $A_2 = -1/2s^2$, $B_2 = y^2$, $C = 1$, $D = (2\pi s^2)^{-1/2} \exp(-t^2/2s^2)$. The two-parameter Weibull distribution is a member of the exponential family when s is unknown, but not when a is unknown, as there is no way to write the Weibull distribution in the exponential family form if a is an unknown parameter. The Cauchy distribution and other heavy-tail distributions are not members of the exponential family (for additional examples, see Garthwaite et al. 2002). Column 5 of Table 1 indicates which distributions are members of the exponential family and for which unknown parameters or combinations of unknown parameters. There is a generalization of the exponential family to multivariate distributions (Kotz et al., 2000).

There are numerous benefits to modeling with an exponential family distribution. The maximum likelihood (ML) estimator (Section 4.1.1) is usually easy to formulate using the log-likelihood function, and it achieves the Cramér-Rao bound (Section 3.2.1). This means the ML estimate is a sufficient statistic (Section 3.3). Also, a *conjugate prior distribution* exists for all the parameters of an exponential family distribution. If $g(\theta)$ is a conjugate prior distribution for a model $f(y|\theta)$, then the posterior distribution $h(\theta|y)$ is the same type of distribution as $g(\theta)$, only with different parameters. This simplifies the mathematical analysis of the problem. Conjugate prior distributions are identified for some of the distributions in column 5 of Table 1.

2.1.8 References on Distributions

References with more extensive information include the encyclopedic five-volume set by Johnson, Kotz, and others who divide their material up by univariate discrete distributions (Johnson et al., 2005), continuous univariate distributions (in two volumes) (Johnson et al., 1994; Johnson et al., 1995), multivariate discrete distributions (Johnson et al., 1997), and multivariate continuous distributions (Kotz et al., 2000). These volumes also give some historical background on the distributions, special properties of the distributions, and information on how to estimate their parameters. Shorter works include

Evans et al. (2000) which also gives information about how to numerically generate random variables of each distribution and Patel et al. (1976) which emphasizes properties of the distributions, like their sufficient statistics and membership in the exponential family. Online searches can provide a great supply of information on particular distributions. Useful sites include www.mathworld.wolfram.com, www.wikipedia.com, and www.itl.nist.gov/div898/handbook/index.htm.

Table 1 lists some important probability distributions and their properties. The first column gives the distribution name and its domain of support. The second column gives the functional form for each distribution: $f(y)$ indicates a probability density function for a continuous-valued random variable y, while $f(n)$ indicated a probability mass function for a discrete-valued random variable n. When applicable, the location parameter is t, and the scale parameter is $s (s > 0)$. Some distributions have additional positive-valued parameters a and b. Some of the discrete distributions contain the parameter p, which is the probability of a Bernoulli event, and N, which is a natural number.

The third column of Table 1 gives moments or moment generating functions for many of the distributions. The moment generating function for a random variable y is

$$M_Y(\mathbf{x}) = E[e^{\mathbf{x}\bullet\mathbf{y}}]$$

where the "dot" indicates multiplication if y is a scalar or the vector dot product if y is a vector. The kth moment of $f(y)$ can be found easily from $M_Y(x)$ as

$$\mu_k = E\left[Y^k\right] = \left.\frac{d^k M_Y(x)}{dx^k}\right|_{x=0}$$

The fourth column of Table 1 gives the quantile function $Q(p)$ for some distributions where p is between 0 and 1. The quantile function is simply the inverse function to the cumulative distribution function (Gilchrist, 2000). In addition to their application in estimating distributions, quantile functions are useful for generating random variables numerically; the value for p is generated as a uniformly distributed random variable between 0 and 1, and the value of the quantile function at p gives an observation from the distribution of interest. More information on generating random variables is found in Gentle (2003).

The fifth column of Table 1 identifies whether the distribution is a member of the exponential family and for which parameters. For example, EF: $a, b, (a, b)$ indicates that the distribution is a member of the exponential family when parameter a is

unknown, when parameter b is unknown, and when both a and b are unknown. The fifth column also identifies the conjugate prior distribution for the parameters when that distribution has a name. For example, CP: p∼beta indicates that the conjugate prior distribution for parameter p is the beta distribution.

The special functions beta $B(a,b)$ and gamma $\Gamma(a)$ are used in Table 1. $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. The gamma function can be found at arguments of $n/2$ where n is a natural number by using $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$, and the recursion $\Gamma(n+1) = n\Gamma(n)$. Note that $\Gamma(n) = (n-1)!$ when n is a positive integer. Values of the gamma function for additional arguments are given in Abramowitz and Stegun (1972) and are available as built in functions in most mathematical software packages such as MATLAB and Mathematica.

2.2 Estimating Distributions

The model distribution is often estimated from a collection of observations. Methods are described as "parametric" or "non-parametric". In parametric methods, the family of distributions (e.g., Gaussian) is selected ahead of time, and the observations are used to estimate the parameters of that family (e.g., mean and standard deviation). Among the parametric methods, the ML method and the method of moments are described here. In non-parametric methods, few assumptions are made about the structure of the distribution. Here we describe the histogram method. Finally, elicitation methods are mentioned which obtain distributions based on expert opinion rather than a set of observations.

2.2.1 Maximum Likelihood

The ML method is the most popular way to estimate parameters of a distribution from a collection of independent samples drawn from that distribution. It is described in Section 4.1.1.

2.2.2 Method of Moments

In the method of moments, model parameters are chosen so that the first K moments of the model distribution equal the corresponding sample moments of the observations (Casella and Berger, 2002). The kth moment of a distribution is

$$\mu_k = \int y^k f(y)\,\mathrm{d}y$$

The kth sample moment of a set of N observations y_i is

$$\bar{\mu}_k = \frac{1}{N}\sum_{i=1}^{N} y_i^k$$

The method of moments estimator is often computationally faster but less accurate than ML. Sometimes the method of moments is used to get initial estimates of parameter values, which are then refined using ML. Expressions for moments of many distributions are given in column 3 of Table 1.

In a variation on the method of moments, the median and median absolute deviation or quantile values can be used instead of moments. This makes the estimate more robust to outliers and is necessary when attempting to estimate parameters of heavy tail distributions, like Cauchy. The quantile function is the inverse function to the cumulative distribution function. It is given in column 4 of Table 1 for some of the distributions.

2.2.3 Histogram/ Kernel Methods

In the histogram method, the range of N observations is divided into bins. The probability density function at y is estimated as the fraction of the observations (n_i/N) which falls in the same bin as y, divided by the width of the bin (h_i).

$$\hat{f}(y) = \frac{n_i}{Nh_i}$$

This method requires many observations to converge and produces discontinuous estimates of the density function. However, it imposes few assumptions on the data, so it should be consulted first to get a sense of what a distribution looks like. It should also be compared to the final estimate of the model distribution to make sure that the assumptions which went into the final estimate have not made the model distribution inconsistent with the observations.

The histogram estimate can be smoothed by applying a kernel function g,

$$\hat{f}(y) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_i}g\left(\frac{y - Y_i}{h_i}\right),$$

where Y_i is the ith out of N observations. The Gaussian probability density function is a common choice for the kernel function. The scale h_i can be the distance from Y_i to its kth nearest neighbor to make the kernel wider where the observations are sparse, improving estimates in the undersampled tails of the distribution.

2.2.4 Elicitation

In the absence of actual observations, a probability distribution may be obtained by asking experts to specify it for you (Cook, 1991; Ayyub, 2000; O'Hagen, 1998). Unfortunately, there are many biases which can unconsciously enter such a specification, and people – even experts – tend to give inaccurate

answers. Elicitation theory deals with how to formulate questions and process responses to obtain the desired information while reducing such biases. The design of an effective elicitation process generally includes formulation of questions in the way most likely to produce accurate responses (e.g., people specify percentile values more accurately than moments of a distribution) and feedback based on consistency checks (e.g., the specified median must be larger than the specified 25th percentile.)

Elicitation has been used in exploration seismology to establish prior distributions for rock properties or to assess the likelihoods of various geological scenarios (Curtis and Wood, 2004), and in earthquake seismology for hazard analysis (NRC, 1988).

2.3 Comparing Models

There are quantitative methods for selecting the best model among a set of options. Some of these methods involve minimizing the Kullback–Leibler (K–L) information, which is the information lost when probability distribution g is used to approximate the true probability distribution f,

$$I(f,g) = \int f(y) \ln\left(\frac{f(y)}{g(y)}\right) \mathrm{d}y$$

The K–L information is simply the expected value of $\ln(f/g)$ (Kullback and Leibler, 1951). The best model is the one which loses the least information, that is, the one which minimizes the K–L information. Unfortunately, in model selection problems we usually do not know the true model, f.

Akaike (1974) found an estimator for the relative K–L information based on the maximized log-likelihood function for a model distribution and the number of parameters K in the model. A refinement is applied to Akaike's estimator to make it more accurate for small numbers of observations, which gives the Akaike information criterion (corrected) (AICC) (Burnham and Anderson, 2003).

$$\mathrm{AICC} = -2\ln\left(L\left(\hat{\theta}\right)\right) + 2K + \frac{2K(K+1)}{n-K-1}$$

The value of AICC is calculated for each candidate model, then the lowest such value is subtracted from all the values. Differences < 2 indicate that the model has substantial support from the data, 4–7 considerably less support, and >10 essentially no support.

These differences can be turned into weights, called the Akaike weights,

$$w_i = \frac{\exp\{-(\mathrm{AICC}_i - \mathrm{AICC}_{\min})\}}{\sum_{j=1}^{N} \exp\{-(\mathrm{AICC}_i - \mathrm{AICC}_{\min})\}}$$

where w_i can be interpreted as the probability that model i is the K–L best model for the data, conditioned on the data and the full set of a priori models being considered. Using these weights, it is possible to construct an estimate which is not conditional on the particular model selected, called multimodel inference (Burnham and Anderson, 2004).

Applications of the Akaike information criterion in seismology include determining the number of wavefronts present when computing moveout velocity spectra in preparation for stacking seismic data (Biondi and Kostov, 1989), and optimizing the complexity of neural networks for various applications (van der Baan and Jutten, 2000).

3 • Information Assessment

For a given model, there is a bound on how accurately parameters can be estimated from observations. This bound should be evaluated as soon as a validated model is constructed and before creating and testing an estimator. Evaluation will prevent futile efforts on estimator design when the observations contain in adequate information about the parameters and can reveal whether an estimator fully utilizes the information in the observations.

If evaluation of the current model shows that estimates will be too inaccurate, the only way forward is to change the model. The model can be changed by adding additional observations (e.g., repeating measurements, incorporating new types of measurements) or by improving the constraints on the parameters and/or observations (e.g., removing model approximations made in relating the parameters to the observations, incorporating new physical insight into relationships between parameters and/or observations, specifying a prior distribution of the parameters).

3.1 Quantifying Information

Information is the potential to reduce uncertainty. While the word "information" is commonly used in a qualitative sense, it has a specific quantitative definition which is useful in optimizing the design of models and estimators. Fisher (1925) quantified the

information about a parameter θ which is provided by a sample drawn from the distribution $f(\mathbf{y}|\theta)$,

$$I(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\ln f(\mathbf{y}|\theta)\right)^2\right] = -E\left[\frac{\partial^2 \ln f(\mathbf{y}|\theta)}{\partial\theta^2}\right]$$

When $\boldsymbol{\theta}$ is a vector of K parameters,

$$I_{ij}(\boldsymbol{\theta}) = E\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\ln f(\mathbf{y}|\boldsymbol{\theta})\,\frac{\partial}{\partial\boldsymbol{\theta}_j}\ln f(\mathbf{y}|\boldsymbol{\theta})\right]$$

where $I_{ij}(\boldsymbol{\theta})$ is the Fisher information matrix, which is a positive semi-definite symmetric $K \times K$ matrix.

Information cannot be created by manipulating the observations. Any function of the observations (including an estimator) will have the same or less information than the original observations, but never more information. When two samples are independent, their combined Fisher information is the sum of their separate Fisher informations.

3.2 Performance Bounds

The amount of information that observations contain about parameters imposes a bound on how accurately the parameters can be estimated. Bounds are given here for the variance of unbiased estimates and the mean squared error of all estimates. When the likelihood function contains many maxima, these bounds may not be tight, so improvements under those conditions are discussed.

3.2.1 Cramér–Rao Bound and Efficiency

The Cramér–Rao bound is the reciprocal of the Fisher information, and is a lower bound on the variance of all unbiased estimates of θ formed from the observations $\mathbf{y}$.

$$E\left[\left(\hat{\theta}(\mathbf{y}) - \theta\right)^2\right] \geq -E\left[\frac{\partial^2 \ln f(\mathbf{y}|\theta)}{\partial\theta^2}\right]^{-1}$$

If a vector of parameters is to be estimated,

$$E\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}\right)^{\mathrm{T}}\right] \geq \left(E\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\ln f(\mathbf{y}|\boldsymbol{\theta})\,\frac{\partial}{\partial\boldsymbol{\theta}_j}\ln f(\mathbf{y}|\boldsymbol{\theta})\right]\right)^{-1}$$

where the inequality indicates that the matrix on the left side minus the matrix on the right side is positive semi-definite. More information about the Cramér-Rao bound is in Kay (1993), including results for complex-valued observations and parameters.

An estimator is called *efficient* if it is unbiased and its variance (or covariance matrix) matches the bound. An efficient estimator exists for θ when $f(\mathbf{y}|\theta)$ is a member of the exponential family

of distributions (and in a few additional cases (Joshi, 1976)). Some estimators are only *asymptotically efficient*, which means that in the limit where the number of observations becomes large, their bias approaches zero (i.e., they are *asymptotically unbiased*) and their variance approaches the Cramér–Rao bound. For most distributions outside the exponential family, no efficient estimator exists. That is, the Cramér–Rao bound underestimates the minimum achievable variance for an unbiased estimator. In these cases, better (larger) lower bounds can be found (Kendall and Stuart, 1979, Section 17.20).

3.2.2 Bayesian Cramér–Rao Bound

When the prior distribution of the parameters is known, a mean squared error bound is available which applies to all estimators, both biased and unbiased. This is sometimes called the "Van Trees inequality" or the "Bayesian Cramér–Rao bound" or the "a posteriori Cramér–Rao bound" (Van Trees, 1968, p. 72; Gill and Levit, 2005).

$$E\left[\left(\hat{\theta}(y)-\theta\right)^2\right] \geq -E\left[\frac{\partial^2 \ln f(y,\theta)}{\partial\theta^2}\right]^{-1}$$

For the multi-parameter case,

$$E\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{y})-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{y})-\boldsymbol{\theta}\right)^{\mathrm{T}}\right] \geq \left(E\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\ln f(\mathbf{y},\boldsymbol{\theta})\,\frac{\partial}{\partial\boldsymbol{\theta}_j}\ln f(\mathbf{y},\boldsymbol{\theta})\right]\right)^{-1}$$

Bounds can be used to identify when a particular algorithm achieves the best possible error performance using data described by the model. The Bayesian Cramér-Rao bound has been used to evaluate methods for multiple suppression in ocean bottom seismic data (Deffenbaugh and Neelamani, 2005). The bound has also been applied in optimizing combinations of algorithms in a seismic processing flow and in research planning and valuation. In these applications, the bound was used as an admissible heuristic in a branch-and-bound tree search where the objective function depends on estimator mean squared error (Deffenbaugh, 2002, 2007). Additional information about the theory and applications of Bayesian bounds is compiled in Van Trees and Bell (2007).

3.2.3 Likelihood Functions with Multiple Maxima

When there are local maxima in the likelihood function and the noise is large enough to sometimes turn the wrong local maximum into the global maximum, occasional large errors will occur. These large errors are not accounted for by the Cramér–Rao inequalities which will give an overly optimistic (loose) lower bound. However, tighter bounds are available for such cases (Xu et al., 2006).

3.3 Sufficient Statistics

The concept of sufficient statistics is a powerful tool for designing optimal estimators. A function of the observations $t(y)$ is a *sufficient statistic* for parameters θ if it contains all the information in y that is relevant to estimation of θ. Formally, the posterior distribution of θ given t is the same as the posterior distribution of θ given y (Garthwaite et al., 2002, p. 151).

$$f(\theta|t) = g(\theta|y)$$

An equivalent condition that is sometimes easier to evaluate is that t is sufficient for θ if $f(y|t)$ can be factored as, (Kendall and Stuart, 1979, Section 17.32)

$$f(y|\theta) = g(t|\theta)\,h(y)$$

where the function $h(y)$ is independent of θ. The sufficient statistic t will have the same Fisher information as y and thus the same Cramér–Rao bound. Examples of sufficient statistics for θ include any one-to-one function of another sufficient statistic for θ, an efficient estimate of θ from y, and any estimate of θ from y which is optimal under some strictly convex loss function (e.g., MMSE). If an estimate can be written entirely as a function of a variable which is not a sufficient statistic, it is not an optimal estimate under any strictly convex loss function. (Rao–Blackwell Theorem in Lehmann & Casella (1998)). For distributions $f(y|\boldsymbol{\theta})$ which are absolutely continuous in $\boldsymbol{\theta}$ for fixed y, and are members of the k-parameter exponential family, a vector of k statistics exists which is sufficient for $\boldsymbol{\theta}$. Otherwise, more than k statistics are required to be sufficient. (Joshi, 1976).

Seismic signals are typically passed through a sequence of processing steps, each of which forms an intermediate data product. It can be difficult to analyze the performance of the estimator formed by the aggregate of all the processing steps. Properties of sufficient statistics can be invoked here to gain some insight. If one of the intermediate data products is not a sufficient statistic, then all data products formed from it are also not sufficient statistics, and the final estimate cannot be optimal under any strictly convex loss function. If, however, all intermediate data products are sufficient statistics, and the final processing step is an optimal estimator, then the compound estimator is also optimal in the same sense as the final step (e.g., if the final step is MMSE optimal, than the compound estimator is MMSE optimal). This principle is exploited in the merging and deconvolution of seismic data sets by the "Texas Two-Step" algorithm, where a multi-input statistical deconvolution is accomplished by first merging the multiple input data sets into a single data

set which is a sufficient statistic and second performing an optimal single-input deconvolution on the sufficient statistic (Deffenbaugh and Neelamani, 2005; Neelamani et al., 2007).

4 • Estimator Selection

The estimators discussed here fall into three classes: (1) Estimators for deterministic parameters, where nothing is assumed about the prior probability distributions of the parameters (2) Estimators for random parameters, where the prior probability distributions are specified; and (3) Robust estimators, which are relatively insensitive to outliers.

4.1 Estimating Deterministic Parameters

If a prior probability distribution for the parameters is not known, or if it is so much wider than the posterior distribution that it adds little information compared to the measurements, then the parameters can be treated as unknown but deterministic values.

4.1.1 Maximum Likelihood Estimators

Maximum likelihood (ML) is the most widely used criterion for estimator design. The likelihood function $L(\mathbf{\theta}|\mathbf{y})$ is equal to the conditional joint probability density function of the observations $f(\mathbf{y}|\mathbf{\theta})$, but while the conditional probability density is thought of as a function of $\mathbf{y}$ for a particular value of $\mathbf{\theta}$, the likelihood function is thought of as a function of $\mathbf{\theta}$ for a particular value of the observation vector $\mathbf{y}$. Note that $L(\mathbf{\theta}|\mathbf{y})$ is not itself a probability distribution since its integral over $\mathbf{\theta}$ generally does not equal 1. The ML estimator finds the value of $\mathbf{\theta}$ that maximizes $L(\mathbf{\theta}|\mathbf{y})$,

$$\hat{\mathbf{\theta}}_{\mathrm{ML}} = \arg\max_{\mathbf{\theta}} L(\mathbf{\theta}|\mathbf{y})$$

It is often more convenient to work with the natural logarithm of the likelihood function, which is maximized at the same parameter value as the likelihood function. Often, the maximum of $L(\mathbf{\theta}|\mathbf{y})$ or $\ln L(\mathbf{\theta}|\mathbf{y})$ can be found analytically as the value of $\mathbf{\theta}$ for which $dL(\mathbf{\theta}|\mathbf{y})/d\mathbf{\theta} = \mathbf{0}$, but i is wise to check the second derivative at such analytical solutions to confirm that a maximum and not a minimum has been found (Pawitan, 2001). The maximum can also be found numerically (Everitt, 1987; Garthwaite et al., 2002).

The ML estimate has some useful mathematical properties. If $\hat{\theta}$ is the ML estimate of θ, then for any function $g(\theta)$, the ML estimate of $g(\theta)$ is $g\left(\hat{\theta}\right)$. This is called the "invariance property" (Casella and Berger, 2002, p. 320). The ML estimate is efficient if an efficient estimator exists (Kendall and Stuart, 1979,

Section 18.5) and is asymptotically efficient even if no efficient estimator exists, provided $f(y|\theta)$ meets certain regularity conditions. The necessary regularity conditions have to do with being able to interchange the order of integration over y and differentiation with respect to θ. They are violated, for example, when the range of acceptable values for y depends on θ. (Lehmann and Casella, 1998) The ML estimate can develop problems when the regularity conditions are not met (Garthwaite et al., 2002, Section 3.2.3; Good and Hardin, 2003).

Applications of ML estimation in seismology include deconvolution (Giannakis and Mendel, 1987; Ursin, 1997 and estimation of moveout velocity spectra (Biondi and Kostov, 1989)).

4.1.2 Least Squares

The least squares (LS) method is routinely used to estimate parameters of a deterministic function contaminated by additive noise: $\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{n}$. The estimate $\boldsymbol{\theta}$ is chosen to minimize the squared difference between $\mathbf{f}(\boldsymbol{\theta})$ and $\mathbf{y}$. The LS estimate, along with its most popular variants, can be written in the form:

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = \arg\min_{\boldsymbol{\theta}} \left\{ (\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}))^{\mathrm{T}} \mathbf{W}_{\mathrm{N}}^{-1} (\mathbf{y} - \mathbf{f}(\boldsymbol{\theta})) + \boldsymbol{\theta}^{\mathrm{T}} \mathbf{W}_{\mathrm{P}}^{-1} \boldsymbol{\theta} \right\}$$

Setting $\mathbf{W}_{\mathrm{N}}{}^{-1} = \mathbf{I}$, $\mathbf{W}_{\mathrm{P}}{}^{-1} = \mathbf{0}$ gives ordinary LS. In weighted LS, $\mathbf{W}_{\mathrm{N}}$ assigns relative weights to the measurements. For correlated noise, the LS estimate is found by making $\mathbf{W}_{\mathrm{N}}$ the covariance matrix of the noise. If $\mathbf{W}_{\mathrm{P}}{}^{-1}$ is non-zero, the solution is not strictly speaking a LS solution, but it is a good idea to set $\mathbf{W}_{\mathrm{P}}{}^{-1} = \varepsilon^2 \mathbf{I}$ where ε is small (sometimes called damped LS or Tikhonov regularization) to prevent wild excursions in those parameter values which are poorly constrained by the measurements (Menke, 1989).

When the noise is Gaussian, the LS estimate (with $\mathbf{W}_{\mathrm{N}}$ the noise covariance matrix, and $\mathbf{W}_{\mathrm{P}}{}^{-1} = \mathbf{0}$) is the ML estimate. When the model is linear ($\mathbf{f}(\boldsymbol{\theta}) = \mathbf{C}\boldsymbol{\theta}$ for some matrix $\mathbf{C}$), the LS estimate is unbiased and is a linear function of the observations. If the model is linear *and* the noise is Gaussian, the LS estimate is the minimum variance estimate, but for linear models with non-Gaussian noises there may be a non-linear estimator with smaller variance than the LS estimate.

The LS estimate can be found by setting the derivative of the objective function with respect to θ to zero. Symbolic matrix derivatives are very useful for making this calculation (Petersen and Pedersen, 2007; Hjorungnes and Gesbert, 2007). Let the matrix $\mathbf{C} = \mathrm{d}\mathbf{f}(\boldsymbol{\theta})/\mathrm{d}\boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta}_0$. Then,

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = \boldsymbol{\theta}_0 + \left(\mathbf{C}^{\mathrm{T}} \mathbf{W}_{\mathrm{N}}^{-1} \mathbf{C} + \mathbf{W}_{\mathrm{P}}^{-1} \right)^{-1} \mathbf{C}^{\mathrm{T}} \mathbf{W}_{\mathrm{N}}^{-1} (\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}_0))$$

An equivalent form by the matrix inversion lemma is

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = \boldsymbol{\theta}_0 + \mathbf{W}_{\mathrm{P}}\mathbf{C}^{\mathrm{T}}\left(\mathbf{C}\mathbf{W}_{\mathrm{P}}\mathbf{C}^{\mathrm{T}} + \mathbf{W}_{\mathrm{N}}\right)^{-1}\left(\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}_0)\right)$$

If the model is not linear, it may be necessary to iterate several times, re-evaluating $\mathbf{C}$ at each new parameter vector estimate until convergence is achieved. If the matrices are complex-valued, then the matrix transposes are replaced by conjugate transposes (Kay, 1993).

Recursive estimators process data one observation at a time. Instead of remembering all prior observations, they remember only a sufficient statistic for the observations and then update this sufficient statistic as each new observation becomes available. In addition to the obvious applications for providing current best estimates from a stream of data (e.g., tracking a radar target), recursive estimators can be used to rapidly process a large data set. Processing the data set one sample at a time is often computationally much faster than handing the data in one large block.

4.1.3 Least Absolute Deviation

The least absolute deviation (LAD) method (also called "least absolute error" or "least absolute value") estimates $\boldsymbol{\theta}$ from noisy $\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{n}$ by fitting $\mathbf{f}(\boldsymbol{\theta})$ to the *median* of the data. In contrast, the LS estimate fits $\mathbf{f}(\boldsymbol{\theta})$ to the *mean* (Li and Arce, 2004). The LAD solution is less sensitive than LS to occasional large noises, so it is usually applied when the noise distribution has a longer tail than the Gaussian or when outliers may be present in the data. The estimate $\boldsymbol{\theta}$ is chosen to minimize the absolute value of the difference between $\mathbf{f}(\boldsymbol{\theta})$ and $\mathbf{y}$. If $\mathbf{W}_{\mathrm{N}}$ is the covariance matrix of the noise and $\mathbf{1}^{\mathrm{T}}$ is a row vector with all elements equal to 1,

$$\hat{\boldsymbol{\theta}}_{\mathrm{LAD}} = \arg\min_{\boldsymbol{\theta}} \left\{ \mathbf{1}^{\mathrm{T}} \left| \mathbf{W}_{\mathrm{N}}^{-1/2}\left(\mathbf{y} - \mathbf{f}\left(\boldsymbol{\theta}\right)\right) \right| \right\}$$

The LAD estimate is the ML estimate when the noise is Laplacian. The LAD estimate has been applied to localization (Shearer, 1997), tomography (Scales et al., 1988), seismic signal processing (Claerbout and Muir, 1973), and many other cases where the occasional large outlier would be overweighted by LS calculations.

4.2 Estimating Random Parameters

When the a priori distributions of the parameters are known, this information can be used to improve the accuracy of the estimates. Estimators which use prior information are described here.

4.2.1 Maximum a Posteriori Estimators

The maximum a posteriori (MAP) estimate seeks the most likely parameter value given the observations,

$$\hat{\boldsymbol{\theta}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\theta}} f\left(\boldsymbol{\theta}|y\right) = \arg\max_{\boldsymbol{\theta}} f\left(y|\boldsymbol{\theta}\right) g\left(\boldsymbol{\theta}\right),$$

where we have used Bayes' rule to rewrite the objective function, and removed a term that does not depend on $\boldsymbol{\theta}$. As the prior distribution $g(\boldsymbol{\theta})$ becomes much wider (less informative) than the posterior distribution $f(\boldsymbol{\theta}|y)$, the MAP estimate approaches the ML estimate. The MAP estimate is the mode of the posterior distribution. Note that for unimodal, symmetric posterior distributions, the MAP estimate equals the conditional mean (MMSE) estimate.

The MAP estimator requires a model for the prior distribution of the parameters. Since it uses the information in this model, it is more accurate than the ML estimate. Applications in seismology include estimating seismic reflectivity (Ursin, 1997) and inversion for elastic properties (Riedel, et al., 2003).

4.2.2 Minimum Mean Squared Error

The MMSE estimate is the mean of the posterior distribution,

$$\hat{\boldsymbol{\theta}}_{\mathrm{MMSE}} = E\left[\boldsymbol{\theta}|\mathbf{y}\right] = \arg\min_{\boldsymbol{\theta}'} E\left\lfloor \left(\boldsymbol{\theta}' - \boldsymbol{\theta}\right)^{\mathrm{T}} \left(\boldsymbol{\theta}' - \boldsymbol{\theta}\right) \right\rfloor$$

For unimodal, symmetric posterior distributions, it is the same as the MAP estimate. For models of the form $\mathbf{y} = \mathbf{C}\boldsymbol{\theta} + \mathbf{n}$ where $\mathbf{n}$ is zero-mean Gaussian, and the prior distribution of $\boldsymbol{\theta}$ is also zero-mean Gaussian, the MMSE estimator is linear, $\hat{\boldsymbol{\theta}}_{\mathrm{MMSE}} = \mathbf{K}\mathbf{y}$, where the matrix $\mathbf{K}$ can be found from the results of Section 4.1.2 by letting $\mathbf{W_P}$ be the covariance matrix of the prior distribution of $\boldsymbol{\theta}$ (Kailath et al., 2000). For non-Gaussian models the linear estimator is not MMSE optimal, but it has the advantage that it can be formulated knowing only the first and second moments of $f(\mathbf{y}|\boldsymbol{\theta})$. Symbolic matrix derivatives can be helpful in finding the matrix $\mathbf{K}$ for linear estimators (Petersen and Pedersen, 2007; Hjorungnes and Gesbert, 2007).

4.2.3 Other Minimum Expected Loss Estimators

An estimator can be designed to minimize the expected value of any loss function λ which depends on the true parameter value and the estimated parameter value, not just the squared error loss seen in the MMSE estimator. A general form for such an estimator is

$$\hat{\boldsymbol{\theta}}_{\mathrm{MLF}} = \arg\min_{\hat{\boldsymbol{\theta}}} E\left[\lambda\left(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}\right)|\mathbf{y}\right]$$

The loss function can be customized to reflect the real-world objective behind a particular estimation problem. For example, it can mimic the actual financial costs associated with different sized errors, possibly assessing a greater penalty to overestimation than to underestimation, or assessing no penalty to errors within a certain range but increasing error outside that range.

4.3 Robust Estimators

Robust estimators tolerate mismatch between the model and the true probability distribution of the observations. In exchange for robustness to model errors, they generally underperform non-robust estimators when there are no model errors. The breakdown point and the efficiency of robust estimators are properties that capture this trade-off. The breakdown point is the minimum percentage of observations which must go to infinity (i.e., become outliers) to make the estimate go to infinity. For the sample mean (which is not robust), the breakdown point is 0%. For the median, the breakdown point is 50%. Efficiency is the ratio of the Cramér-Rao bound to the variance achieved by the estimator. An efficient estimator has an efficiency of 100%.

4.3.1 *L*-Estimators

L-estimators are linear combinations of the *order statistics*, the observations sorted from smallest to largest. They have the form,

$$\hat{\theta} = \sum_{i=1}^{N} a_i y_{(i)}$$

where $y_{(i)}$ is the ith largest observation and a_i are weights.

The *median* is a simple *L*-estimator, which corresponds to $a_{(N+1)/2} = 1$, and all other weights zero, for odd N. (For even N, we might take $a_{N/2} = a_{N/2+1} = 0.5$, and all other weights zero.) The median offers a 50% breakdown point and 64% efficiency for estimating the mean of a Gaussian distribution. The efficiency would be different for other distributions. Median filters, based on taking the median of a sliding window of data can reproduce sharp transitions in data while removing outliers or "speckle". The use of medians has been proposed for a variety of problems in seismology (Bednar, 1983).

The *trimmed mean* is found by discarding the R largest and R smallest samples, then averaging the rest: $a_i = 1/(N-2R)$, for i between $R+1$ and $N-R$, and zero otherwise (Bednar and Watt, 1984). This is sometimes called the alpha-trimmed mean where $\alpha = R/N$. It offers a breakdown point of α. When estimating the mean of a Gaussian distribution, the efficiency is 94% for $\alpha = 0.1$ and 85% for $\alpha = 0.25$. Applications of the trimmed mean have included stacking seismic data where occasional noisy traces

can contaminate the stacked (summed) trace (Haldorsen and Farmer, 1989). When attenuation is estimated from ratios of frequency spectra at different receiver locations, the residuals have a heavy-tailed distribution, and trimmed means have been applied (Patton, 1988).

The *interquartile range* is the difference between the values at the 75th and 25th percentiles: $a_i = 1$ for $i = 0.75N$, $a_i = -1$ for $i = 0.25N$, and $a_i = 0$ otherwise. It can be used to calculate robust estimates of scale parameters (such as the standard deviation). The distance between other percentiles could also be used. Higher order differences among the order statistics can be calculated and related to higher order moments or other parameters of a distribution. One such example is "*L*-moments" (Ulrych et al., 2000).

For very large data sets, estimators which rely on order statistics can be slow and cumbersome: slow, because generating an order statistic requires sorting all the observations, and cumbersome, because sorting data generally requires holding all or most of it in memory simultaneously. Fortunately, clever algorithms are available for rapid and low-memory recovery of order statistics (Babu and McDermott, 2002; Hurley and Modarres, 1995).

4.3.2 *M*-Estimators

The *M*-estimator is a generalization of the ML estimator For *N*-independent measurements,

$$\hat{\theta}_M = \arg\min_{\theta} \sum_{i=1}^{N} \rho\,(y_i|\theta).$$

This would be equivalent to the ML estimate if $\rho(y|\theta) = -\ln f(y|\theta)$, but $\rho(y|\theta)$ is usually chosen to make the estimate more robust (less sensitive) to outliers. If θ is a location parameter, a popular choice of $\rho(y|\theta)$ is $(y-\theta)^2$ (LS) for $|y-\theta| < c$, but $2c|y-\theta| - c^2$ (LAD) for $|y-\theta| > c$ (Huber, 1981). As with the ML estimate, the *M*-estimate is not necessarily unique, and one must be cautious of local minima for certain choices of $\rho(y|\theta)$.

In seismology, *M*-estimators have been proposed for many of the same applications as *L*-estimators, such as deconvolution (Walden, 1985), stacking (Elston, 1990), and array beam forming (Bear and Pavis, 1997).

4.3.3 Outlier Detection

Another approach to making a robust estimator is to first identify and remove the outliers from a data set and then form a non-robust estimate from the remaining observations. Outlier detection techniques are tests of the statistical hypothesis that a particular observation is not a member of a particular model

distribution. A good survey of these techniques is contained in Barnett and Lewis (1994).

5 • Estimator Validation

The final step in solving parameter estimation problems is to validate the performance of the selected estimator. Both real data and synthetic data play an important role in validating performance. Real data contains all the relevant physics, while in synthetic data, the true parameter values are known.

5.1 Selecting Data for Validation

Data points used to calibrate parameters of the model should not be reused to evaluate estimator performance. The simplest way to avoid this is to divide the available data into two sets, use one set to determine the model parameters and another set to evaluate the model. A more efficient use of limited data, called *cross-validation*, is to divide the data into K sets, withhold one set for the evaluation and use the other $K-1$ sets to compute the parameters of the model. Repeat this process K times, each time withholding a different set for the evaluation, and then average the results of the evaluations. An even better approach, called *bootstrapping*, is to randomly select (with replacement) samples of the data for the evaluation, each time using the remaining samples to compute the parameters of the model. The bootstrap method of randomly selecting samples of the data is useful whenever samples are required having the same probability distribution as the data.

5.2 Validation Questions

The following questions should be addressed in validating the performance of an estimator.

5.2.1 Error Magnitudes

Are the estimate errors small enough? If the errors are not small enough, the options are to (1) include more observations, (2) improve the model, or (3) improve the estimator (if the estimator is not already fully utilizing the information in the data). The effect of outliers should be considered here by throwing in some spurious data points, and making sure that either they are detectable by the human operator or the estimator is robust to their presence.

5.2.2 Error Distribution

Do the estimate errors have the anticipated distribution? In many cases the estimate errors should have a known mean, variance, or even distribution form. These anticipated properties should be confirmed. The form of the error distribution can be checked by making a histogram of the errors and overplotting

the expected distribution form. If the errors do not have the anticipated statistics, something is wrong with the model or the estimator.

5.2.3 Error Correlations

Are the estimate errors uncorrelated with themselves and the parameters? If there is correlation between the errors and the parameters or among the errors themselves, then there is some unused information and the errors are larger than they need to be. It is useful to cross plot the estimate errors with each of the parameters to see if there is any residual correlation. Such residual correlation indicates that the data contains more information about the parameters than is being captured by the chosen combination of model and estimator.

6 • Conclusion

Much of seismology is concerned with estimating earth properties from a limited supply of indirect measurements. A four step method is recommended. The first step, model selection, is the most difficult. The earth has structure at scales far below the resolution of seismic waves, and simulations of wave propagation must make significant approximations to provide results in reasonable time. At best, seismic waves reveal the elastic properties of the earth. Further assumptions are required to relate the sensed elastic properties to the information about process and composition that we seek for scientific or economic reasons. The random component of models in seismology is typically dominated by the residues of these approximations rather than ambient noises. Due to model complexity, seismic signal processing is dominated by ad hoc methods which are useful but not rigorously optimal and by processing flows that combine many separate algorithms to form an estimator. The second step, information assessment, derives from the model an absolute standard for estimator performance. Ad hoc methods can be compared to this standard and powerful results are available for identifying the weak links in processing flows. The third step, estimator selection, is guided by the statistics in the model, whether prior information on parameter values is available, whether spurious data points could corrupt the estimate, and ease of use considerations. The fourth and final step, estimator validation, involves more than just trying the estimator on some real data to see if it gives reasonable answers. Clever methods are available for getting the most out of limited test data, and careful statistical analysis of estimate errors can reveal model or implementation problems and detect the potential for improved estimate accuracy.

TABLE 1 Some Important Probability Distributions and Distribution Families

Name	Probability Distribution	Moments	Quantiles	Properties*
Bernoulli $n \in \{0,1\}$	$f(n) = \begin{cases} p & \text{for } n=1 \\ 1-p & \text{for } n=0 \end{cases}$	$\mu_k = p$	$Q(p') = \begin{cases} 1 & \text{for } p' > p \\ 0 & \text{for } p' < p \end{cases}$	EF: p CP: p∼beta
Beta $y \in [0,1]$	$f(y) = \frac{1}{B(a,b)} y^{a-1}(1-y)^{b-1}$	$\mu_k = \frac{B(a+k,b)}{B(a,b)}$	–	EF: a, b, (a,b)
Binomial $n \in \{0,1,\cdots,N\}$	$f(n) = \frac{N!}{n!(N-n)!}(1-p)^{N-n}p^n$	$M_n(x) = (1-p+pe^x)^N$	–	EF: p CP: p∼beta
Cauchy $y \in (-\infty,\infty)$	$f(y) = \frac{1}{s\pi}\left[1+\left(\frac{y-t}{s}\right)^2\right]^{-1}$	μ_{odd} = undefined $\mu_{even} = \infty$	$Q(p) = s\tan\left(\pi p - \frac{\pi}{2}\right)+t$	EF: No
Chi $y \in [0,\infty)$	$f(y) = \frac{1}{2^{(N/2)-1}s\Gamma\left(\frac{N}{2}\right)}\left(\frac{y}{s}\right)^{N-1}\exp\left(-\frac{1}{2}\left(\frac{y}{s}\right)^2\right)$	$\mu_k = 2^{k/2}s^k\frac{\Gamma((N+k)/2)}{\Gamma(N/2)}$	For N = 2, see Rayleigh	EF: s, N, (s,N)
Chi-square $y \in [0,\infty)$	$f(y) = \frac{1}{2\Gamma\left(\frac{N}{2}\right)}\left(\frac{y}{2}\right)^{(N/2)-1}\exp\left[-\left(\frac{y}{2}\right)\right]$	$M_Y(x) = (1-2x)^{-N/2}$	–	EF: No
Exponential $y \in [0,\infty)$	$f(y) = \frac{1}{s}\exp\left(-\frac{y}{s}\right)$	$\mu_k = s^k k!$	$Q(p) = -s\ln(1-p)$	EF: s CP: s^{-1} ∼gamma
Exponential family	$f(y\|\theta) = C(y)D(\theta)\exp\left[\sum_{i=1}^{K} A_i(\theta)B_i(y)\right]$	–	–	EF: Yes
F-distribution $y \in [0,\infty)$	$f(y) = \frac{1}{B\left(\frac{N}{2},\frac{M}{2}\right)}\frac{\left(\frac{N}{M}\right)^{N/2}y^{N/2-1}}{\left(1+\frac{N}{M}y\right)^{(N+M)/2}}$	$\mu_k = \left(\frac{M}{N}\right)^k\frac{\Gamma\left(\frac{N}{2}+k\right)\Gamma\left(\frac{M}{2}-k\right)}{\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{M}{2}\right)}$	–	EF: No
Gamma	See Generalized gamma with b = 1	$M_Y(x) = (1-sx)^{-a}$	–	EF: s
Gaussian $y \in (-\infty,\infty)$	$f(y) = \frac{1}{s\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{y-t}{s}\right)^2\right)$	$M_Y(x) = \exp\left(tx+\frac{s^2x^2}{2}\right)$	$Q(p) =$ $t+s\sqrt{2}erf^{-1}(2p-1)$	EF: t, s, (t,s) CP: t∼Gaussian CP:$1/s^2$ ∼gamma
Gaussian (Multivariate) $y_i \in (-\infty,\infty)$	$f(\mathbf{y}) = \frac{1}{\sqrt{\det(\mathbf{S})(2\pi)^N}}\exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{t})^{\mathbf{T}}\mathbf{S}^{-1}(\mathbf{y}-\mathbf{t})\right)$	$M_Y(x) = \exp\left(\mathbf{t}^{\mathbf{T}}\mathbf{x}+\frac{1}{2}\mathbf{x}^{\mathbf{T}}\mathbf{S}\mathbf{x}\right)$	–	EF: t, S, (t,S) CP: t∼Gaussian CP: S^{-1} ∼Wishart

Generalized beta $y \in [t, t+s]$	$f(y) = \frac{1}{B(a,b)s^{a+b-1}} (y-t)^{a-1} (t+s-y)^{b-1}$	See Karian & Dudewicz, 2000	–	EF: a, b, (a,b)
Generalized gamma $y \in [0, \infty)$	$f(y) = \frac{b}{\Gamma(a)s} \left(\frac{y}{s}\right)^{ab-1} \exp\left[-\left(\frac{y}{s}\right)^{b}\right]$	For b=1, see Gamma	–	EF: s
Generalized lambda $y \in [t-s, t+s]$	–	See Dudewicz, 1996	$Q(p) =$ $t+s\left(p^{a}-(1-p)^{b}\right)$	EF: No
Geometric	See Negative Binomial with N = 1		$Q(p') = \text{ceil}\left(\frac{\ln(1-p')}{\ln(1-p)}\right)$	EF: p CP: p ~ beta
Laplace $y \in (-\infty, \infty)$	$f(y) = \frac{1}{2s} \exp\left(-\frac{\lvert y-t \rvert}{s}\right)$	$M_Y(x) = \frac{\exp(tx)}{1-s^2x^2}$	$Q(p) = t - cs$ $\ln(1-\lvert 2p-1 \rvert)$ where $c = sgn\left(p-\frac{1}{2}\right)$	EF: s
Levy-Stable Distribution $y \in (-\infty, \infty)$	$f(y) = \int_{-\infty}^{\infty} \varphi(x) \exp(-ixy)\, dx$ where, $\varphi(x) = \exp[ixt - \lvert sx \rvert^{\alpha}]$, $0 < \alpha < 2$	–	–	EF: No
Logistic $y \in (-\infty, \infty)$	$f(y) = \frac{\exp(-(y-t)/s)}{s[1+\exp(-(y-t)/s)]^2}$	$M_Y(x) = \pi s x e^{tx} \csc(\pi s x)$	$Q(p) = t - s\ln\left(\frac{1}{p}-1\right)$	EF: No
Lognormal $y \in [0, \infty)$	$f(y) = \frac{1}{ys\sqrt{2\pi}} \exp\left(-\frac{(\ln y - t)^2}{2s^2}\right)$	$\mu_k = \exp\left(tk + \frac{k^2 s^2}{2}\right)$	$Q(p) =$ $\exp\left(t + s\sqrt{2} erf^{-1}(2p-1)\right)$	EF: t, s, (t,s)
Maximum Entropy	$f(y) = \frac{1}{c(\lambda_1, \cdots, \lambda_N)} \exp\left\{\sum_{n=0}^{N} \lambda_n y^n\right\}$ c is a normalization constant	See Jaynes, 1957	–	EF: Yes
Maxwell	See Chi with N=3			
Negative Binomial $n \in \{0, 1, \cdots, \infty\}$	$f(n) = \frac{\Gamma(N+n)}{n!\Gamma(N)} p^N (1-p)^n$	$M_n(x) = \left(\frac{p}{1-(1-p)\exp(x)}\right)^N$	for N=1, see Geometric	EF: p
Normal	See Gaussian			
Normal Product $y \in (-\infty, \infty)$	$f(y) = \frac{1}{\pi s} K_0\left(\frac{\lvert y \rvert}{s}\right)$ K_0 = modified Bessel func. of 2nd kind.	–	–	EF: No

(*Continued*)

TABLE 1 (*Continued*)

Name	Probability Distribution	Moments	Quantiles	Properties*
Pareto $y \in [b, \infty)$	$f(y) = \frac{ab^a}{y^{a+1}}$	$\mu_k = \frac{ab^k}{a-k}$ for k < a $\mu_k = \infty$ for $k \geqslant a$	$Q(p) = b(1-p)^{-1/a}$	EF: a CP: a∼gamma CP: b∼pareto
Poisson $n \in \{0, 1, \cdots, \infty\}$	$f(n) = \frac{a^n}{n!}\exp(-a)$	$M_n(x) = \exp(a(e^x - 1))$	–	EF: a CP: a∼gamma
Rayleigh	See Chi with N = 2		$Q(p) = s\sqrt{-2\log(1-p)}$	EF: s
Triangle $y, c \in [t, t+s]$	$f(y) = \begin{cases} \frac{2(y-t)}{s(c-t)} & \text{for } y \leqslant c \\ \frac{2(t+s-y)}{s(t+s-c)} & \text{for } y \geqslant c \end{cases}$	–	$Q(p) = \begin{cases} t + \sqrt{ps(c-t)} & \text{for } p \leq \frac{c-t}{s} \\ t+s-\sqrt{(1-p)s(t+s-c)} & \text{for } p \geq \frac{c-t}{s} \end{cases}$	EF: No
Uniform $y \in [t, t+s]$	$f(y) = \frac{1}{s}$	$\mu_k = \frac{(t+s)^{k+1} - t^{k+1}}{s(k+1)}$	$Q(p) = sp + t$	EF: No
Weibull $y \in [0, \infty)$	$f(y) = \frac{a}{s}\left(\frac{y}{s}\right)^{a-1}\exp\left(-\left(\frac{y}{s}\right)^a\right)$	$\mu_k = s^k\Gamma(1 + k/a)$	$Q(p) = s(-\log(1-p))^{1/a}$	EF: s

* EF: exponential family parameter sets. CP: conjugate prior distributions.

TABLE 2 Distributions of Functions of Gaussian random variables

Function	Distribution of Function
$z_1+z_2+\cdots+z_N$	Gaussian: $t=\sum\mu_i, s^2=\sum\sigma_i^2$
z_1z_2	Normal Product: $\mu_i=0, s=\sigma_1\sigma_2$ (Weisstein)
z_1/z_2	Cauchy: $t=\mu_i=0, s=\sigma_i$
$\sqrt{z_1^2+z_2^2}$	Rayleigh: $\mu_i=0, s=\sigma_i$
$\sqrt{z_1^2+z_2^2+z_3^2}$	Maxwell: $\mu_i=0, s=\sigma_i$
$\sqrt{z_1^2+z_2^2+\cdots+z_N^2}$	Chi (degree N): $\mu_i=0, s=\sigma_i$
$z_1^2+z_2^2+\cdots+z_N^2$	Chi-square (degree N): $\mu_i=0, \sigma_i=1$ Non-central chi-square (degree N): $\mu_i=\lambda, \sigma_i=1$
$\frac{\frac{1}{N}\left(z_1^2+z_2^2+\cdots+z_N^2\right)}{\frac{1}{M}\left(z_{N+1}^2+z_2^2+\cdots+z_{N+M}^2\right)}$	F-distribution (degrees N,M): $\mu_i=0, \sigma_i=1$ Non-central F-distribution (degrees N,M): $\mu_{1\cdots M}=t, \mu_{M+1\cdots M+N}=u, \sigma_i=1$
$\frac{z_1^2+z_2^2+\cdots+z_N^2}{z_1^2+z_2^2+\cdots+z_M^2}, M>N$	Beta: $\mu_i=0, \sigma_i=1, a=\frac{N}{2}, b=\frac{M}{2}$ Non-central beta: $\mu_i=\lambda, \sigma_i=1, a=\frac{N}{2}, b=\frac{M}{2}$

References

Abramowitz, M. and I. A. Stegun, 1972, *Handbook of Mathematical Functions*, Dover.

Akaike, H., 1974, "A new look at statistical model identification", IEEE Trans. Automatic Control, AC-19, 716–723.

Ayyub, B. A., 2000, "Methods for expert-opinion elicitation of probabilities and consequences for Corps facilities" U.S. Army Corps of Engineers, Institute for Water Resources, Alexandria, VA, IWR Report 00-R-10.

Babu, G. J. and 'McDermott, J. P., 2002, Statistical methodology for massive datasets and model selection. In *Astronomical Data Analysis II*, Starck J. -L. and F. D. Murtagh (eds), Proceedings of SPIE, vol. **4847** 228–237. (downloaded PDF)

Barnett, V. and T. Lewis, 1994, *Outliers in Statistical Data*, 3rd ed., John Wiley & Sons: Hoboken, NJ.

Bear, L. K. and G. L. Pavis, 1997, "Estimation of slowness vectors and their uncertainties using multi-wavelet seismic array processing", Bull. Seis. Soc. Am., vol. 87, no. 3, pp. 755–769.

Beattie, P. D. and W. B. Dade, 1996, "Is scaling in turbidite deposition consistent with forcing by earthquakes?", J. Sediment. Res., vol. 66, pp. 909–915.

Bednar, J. B., 1983, "Applications of median filtering to deconvolution, pulse estimation, and statistical editing of seismic data", Geophysics, vol. 48, pp. 1598–1610.

Bednar, J. B. and T. L. Watt, 1984, "Alpha-trimmed means and their relationship to median filters", IEEE Trans. ASSP, vol. ASSP-32, no. 1, pp. 145–153.

Biondi, B. L., and C. Kostov, 1989, "High-resolution velocity spectra using eigenstructure methods", Geophysics, vol. 54, no. 7, pp. 832–842.

Blandford, R. R., 1974, "An automatic event detector at the Tonto Forest Seismic Observatory", Geophysics, vol. 39, pp. 633–643.

Bormann, P., (Ed.) 2002, IASPEI New Manual of Seismic Observatory Practice, GeoForschungsZentrum: Potsdam, vol. 1. (Online links at www.seismo.com/msop/msop_intro.html)

Burnham, K. P. and D. Anderson, 2003, *Model Selection and Multi-Model Inference,* Springer: New York .

Burnham, K. P. and D. Anderson, 2004, "Multimodel inference: Understanding the AIC and BIC in model selection" available online: http://www2.fmg.uva.nl/modelselection/presentations/AWMS2004-Burnham-paper.pdf

Casella, G. and R. L. Berger, 2002, *Statistical Inference 2nd ed.*, Duxbury Press.

Claerbout, J. F. and F. Muir, 1973, "Robust modeling with erratic data", Geophysics, vol. 38, pp. 826–844.

Connor, C. B., R. S. J. Sparks, R. M. Mason, C. Bonadonna, and S. R. Young, 2003, "Exploring links between physical and probabilistic models of volcanic eruptions: The Soufriere Hills Volcano, Montserrat", Geophys. Res. Lett., vol. 30, p. 1701.

Cook, R. M., 1991, *Experts in Uncertainty*, Oxford University Press: London.

Cox, D. R. and D. Oakes, 1984, *Analysis of Survival Data*, CRC Press: Boca Raton, FL.

Curtis, A. and R. Wood, 2004, "Optimal elicitation of prior information from experts." in Curtis, A. and R. Wood (eds) *Geological Prior Information, Geological Society of London Special Publication 239*. (Available at Curtis' website: www.geos.ed.ac.uk/homes/acurtis)

Deffenbaugh, M., 2002, "Use of the Bayesian Cramér-Rao bound in seismic exploration and research planning", J. Acoust. Soc. Am., vol. 112, p. 2391.

Deffenbaugh, M., 2007, "Optimizing seismic survey design and processing flows: A fast tree search using the Van Trees inequality", Annual IEEE Systems Conference.

Deffenbaugh, M. and R. Neelamani, 2005, "OBC multiple suppression with the Texas two-step", 75th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, pp. 2123–2127.

Douze, E. J. and S. J. Laster, 1979, "Statistics of Semblance", Geophysics, vol. 44, pp. 1999–2003.

Dudewicz, E. J., 1996, "The extended generalized Lamda distribution system for fitting distributions to data with moments", Am. J. Math. Manag. Sci., vol. 16, pp. 271–332.

Dudewicz, E. J., 1999, "Fitting the GLD system by a method of percentiles", Am. J. or Math. Manag. Sci., vol. 19, no. (1&2), pp. 1–73.

Elston, S. F., 1990, "Use of robust estimators in multichannel stacking", 60th Annual International Meeting, Society of Exploration Geophysicsts, Expanded Abstracts, 1693–1696.

Evans, M., N. Hastings, J. B. Peacock, 2000, *Statistical Distributions, 3rd ed.*, John Wiley & Sons: Hoboken, NJ.

Everitt, B. S., 1987, *Introduction to Optimization Methods and Their Application in Statistics*, Champan & Hall: London.

Fehler, M., H. Sato, and L-J Huang, 2000, "Envelope broadening of outgoing waves in 2D random media: A comparison between the Markov approximation and numerical simulations", Bull. Seismo. Soc. Am., vol. 90, no. 4, p. 914–928.

Fisher, R. A., 1925, "Theory of Statistical Estimation" Proceedings of the Cambridge Philosophical Society, 22, 700–725.

Garthwaite, P., I. Jolliffe, and B. Jones, 2002, *Statistical Inference*, Oxford University Press: New York.

Gentle, J. E., 2003, *Random Number Generation and Monte Carlo Methods, 2nd ed.*, Springer: New York.

Giannakis, G. B. and J. M. Mendel, 1987, "Entropy interpretation of maximum-likelihood deconvolution", Geophysics, vol. 52, pp. 1621–1630.

Gilchrist, W., 2000, *Statistical Modeling with Quantile Functions*, CRC Press: Boca Raton, FL.

Gill, R. D. and B. Y. Levit, 2005, "Applications of the Van Trees inequality: A Bayesian Cramér-Rao bound", Bernoulli, vol. 1, pp. 59–79.

Good, P. I., and J. W. Hardin, 2003, *Common Errors in Statistics (and How to Avoid Them)*, John Wiley & Sons: Hoboken, NJ.

Greenman, J. V. and M. J. Fryer, 1996, "Hydrocarbon field size distributions: A case study in mixed integer nonlinear programming", J. Oper. Res. Soc., vol. 47, pp. 1433–1442.

Gupta, A. K. and S. Nadarajah, 2004, *Handbook of Beta Distribution and its Applications*, Marcel Dekker: New York.

Haldorsen, J. B. U. and P. A. Farmer, 1989, "Suppression of high-energy noise using an alternative stacking procedure", Geophysics, vol. 54, no. 2, pp. 181–190.

Hjorungnes, A. and D. Gesbert, 2007, "Complex-valued matrix differentiation: Techniques and key results", IEEE Trans. Signal Processing, vol. 55, no. 6, pp. 2740–2746.

Ho, C.–H., 1991, "Nonhomogeneous Poisson model for volcanic eruptions", Math. Geol., vol. 23, pp. 167–173.

Huber, P. J., 1981, Robust Statistics, John Wiley & Sons.

Hurley, C. and R. Modarres, 1995, "Low-storage quantile estimation", Comput. Stat., vol. 10, pp. 311–325.

Jaynes, E. T., 1957, "Information theory and statistical mechanics", Phys. Rev., vol. 106, pp. 620–630.

Johnson, N. L., A. W. Kemp, and S. Kotz, 2005, *Univariate Discrete Distributions, 3rd ed.*, John Wiley & Sons, Inc.: Hoboken, NJ.

Johnson, N. L., S. Kotz, and N. Balakrishnan, 1994, *Continuous Univariate Distributions*, vol. 1, 2nd ed., John Wiley & Sons: Hoboken, NJ.

Johnson, N. L., S. Kotz, and N. Balakrishnan, 1995, *Continuous Univariate Distributions*, vol. 2, 2nd ed., John Wiley & Sons: Hoboken, NJ.

Johnson, N. L., S. Kotz, and N. Balakrishnan, 1997, *Discrete Multivariate Distributions*, John Wiley & Sons: Hoboken, NJ.

Joshi, V. M., 1976, "On attainment of the Cramér-Rao lower bound", Ann. Stat., vol. 4. pp. 998–1002.

Kagan, Y. Y., 1992, "Correlations of earthquake focal mechanisms", Geophys. J. Int. vol. 110, pp. 305–320.

Kailath, T., A. H. Sayed, and B. Hassibi, 2000, *Linear Estimation*, Prentice-Hall.

Karian, Z. A. and E. J. Dudewicz, 2000, *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods*, Boca Raton: Chapman & Hall/CRC.

Kay, S. M., 1993, Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory, Prentice-Hall.

Kendall, M. and A. Stuart, 1979, *The Advanced Theory of Statistics*, 4th ed. Vol. 2., MacMillan: New York.

Kimball, C. V. and D. J. Scheibner, 1998, "Error bars for sonic slowness measurements", Geophysics, vol. 63, pp. 345–353.

Kondolf, G. M. and A. Adhikari, 2000, "Weibull vs. Lognormal distributions for fluvial gravels", J. Sediment. Res., vol. 70, pp. 456–460.

Kotz, S., N. Balakrishnan, and N. L. Johnson, 2000, *Continuous Multivariate Distributions*, vol. 1, 2nd ed., John Wiley & Sons: Hoboken, NJ.

Kullback, S. and R. A. Leibler, 1951, "On information and sufficiency", Ann. Math. Stat., vol. 22, pp. 79–86.

Laubach, S., R. Marrett, and J. Olson, 2000, "New directions in fracture characterization", The Leading Edge, vol. 19, pp. 704–711.

Lawless, J. F., 2002, *Statistical Models and Methods for Lifetime Data*, 2nd ed., John Wiley & Sons: Hoboken, NJ.

Lehmann, E. L. and G. Casella, 1998, *Theory of Point Estimation*, 2nd ed., Springer-Verlag: New York.

Li, Y. and G. R. Arce, 2004, "A maximum likelihood approach to least absolute deviation regression", EURASIP J. Appl. Signal Processing, vol. 2004, no. 12, pp. 1762–1769.

Madhava-Rao, N. and K. L. Kaila, 1986, "Application of the negative binomial to earthquake occurrences in the Alpide-Himalayan belt", Geophysical Journal of the Royal Astronomical Society, vol. 85, no. 2, pp: 283–90.

Marrett, R., O. J. Ortega, and C. M. Kelsey, 1999, "Extent of power-law scaling for natural fractures in rock", Geology, vol. 27, pp. 799–802.

Menke, W., 1989, G*eophysical Data Analysis: Discrete Inverse Theory (revised edition), International Geophysical Series, vol. 45*, Academic Press: San Diego, CA.

Neelamani, R., M. Deffenbaugh, and R. Baraniuk, 2007, "Texas Two-Step: A framework for optimal multi-input single-output deconvolution" IEEE Trans on Image Processing, vol. 16, no. 11, p. 2752–2765.

NRC, 1988, *Probabilistic Seismic Hazard Analysis, National Research Council*, National Academy Press: Washington, DC.

O'Hagan, A., 1998, "Eliciting Expert Beliefs in Substantial Practical Applications," *Statistician*, vol. 47, no. 1, pp. 21–35. (This edition contains several papers on elicitation.)

Olhovich, V. A., 1964, "The causes of noise in seismic reflection and refraction work", Geophysics, vol. 29, no. 6, pp. 1015–1030.

Painter, S., 2003, "Statistical Characterization of Spatial Variability in Sedimentary Rock", (Chapter 7) in Goff and Holliger, *Heterogeneity in the Crust and Upper Mantle*, Kluwer Academic/Plenum Publishers.

Painter, S., G. Beresford, and L. Paterson, 1995, "On the distribution of seismic reflection coefficients and seismic amplitudes", Geophysics, vol. 60, no. 4, pp. 1187–1194.

Pankow, K. L., W. J. Arabasz, J. C. Pechmann, and S. J. Nava, 2004, "Triggered seismicity in Utah from the 3 November 2002 Delali Fault earthquake", Bull. Seismo. Soc. Am., vol. 94, no. 6B, pp. S332–S347.

Patel, J. K., C. H. Kapadia, and D. B. Owen, 1976, *Handbook of Statistical Distributions*, Marcel Dekker: New York.

Patton, S. W., 1988, "Robust and least-squares estimation of acoustic attenuation from well-log data", Geophysics, vol. 53, pp. 1225–1232.

Pawitan, Y., 2001, *In all Likelihood: Statistical Modelling and Inference Using Likelihood,* Oxford University Press: New York, 2001.

Petersen, K. B. and M. S. Pedersen, 2007, The Matrix Cookbook, http://matrixcookbook.com.

Peterson, J., 1993, "Observations and modeling of seismic background noise", U.S. Geological Survey Open-File Report 93–322.

Riedel, M., S. E. Dosso, and L. Beran, 2003, "Uncertainty estimation for amplitdue variation with offset (AVO) inversion", Geophysics, vol. 68, no. 5, pp. 1485–1496.

Sax, R. L., 1968, "Stationarity of seismic noise", Geophysics, vol. 33, pp. 668–674.

Scales, J. A., A. Gersztenkorn, S. Treitel, and L. R. Lines, 1988, "Robust optimization methods in geophysical inverse theory", 58th Annual Meeting, Society of Exploration Geophysicists, Expanded Abstracts, pp. 827–830.

Scholz, C. H., 1997, "Size distributions for large and small earthquakes", Bull. Seis. Soc. Am., vol. 87, no. 4, pp. 1074–1077.

Shearer, P. M., 1997, "Improving local earthquake locations using the L1 norm and waveform cross correlation; application to the Whittier Narrows, California, aftershock sequence", J. Geophys. Res. B: Solid Earth Planets, vol. 102, no. 4, pp. 8269–8283.

Sornette, D. and L. Knopoff, 1997, "The paradox of the expected time until the next earthquake", Bull. Seis. Soc. Am., vol. 87, no. 4, pp. 789–798.

Todoeschuck, J. P., O. G. Jensen, and S. Labonte, 1990, "Gaussian scaling noise model of seismic reflection sequences: Evidence from well logs", Geophysics, vol. 55, no 4, pp. 480–484.

Turcotte, D. L., 1997, *Fractals and Chaos in Geology and Geophysics* 2nd ed. 1997, Cambridge University Press: New York.

Ulrych, T. J., D. R. Velis, A. D. Woodbury, M. D. Sacchi, 2000, "L-moments and C-moments", Stoch. Environ. Res. Risk Assess., vol. 14, pp. 50–68.

Ursin, B., 1997, "Methods for estimating the seismic reflection response", Geophysics, vol. 62, pp. 1990–1995.

Van Trees, H. L., 1968, *Detection, Estimation, and Modulation Theory, Part 1*, John Wiley and Sons: New York.

Van Trees, H. L. and K. L. Bell, 2007, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering and Tracking*, Wiley-IEEE Press.

van der Baan, M. and C. Jutten, 2000, "Neural networks in geophysical applications", Geophysics, vol. 65, no. 4, pp. 1032–1047.

Walden, A. T., 1985, "Non-Gaussian reflectivity, entropy, and deconvolution", Geophysics, vol. 50, pp. 2862–2888.

Weisstein, Eric W., "Normal product distribution" from MathWorld—A Wolfram Web Resource, http://mathworld.worlfram.com/NormalProductDistribution.html

Wilkinson, B. H., C. N. Drummond, N. W. Diedrich, and E. D. Rothman, 1999, "Poisson processes of carbonate accumulation on Paleozoic and Holocene platforms", J. Sediment. Res., vol. 69, pp. 338–350.

Wilson, D., J. Leon, R. Aster, J. Ne, J. Schlue, S. Grand, S. Semken, S. Baldridge, and W. Gao, 2002, "Broadband seismic background noise at temporary seismic stations observed on a regional scale in the southwestern United States", Bull. Seis. Soc. Am., vol. 92, pp. 3335–3341.

Xu, W., A. B. Baggeroer, and H. Schmidt, 2006, "Performance analysis for matched-field source localization: simulations and experimental results", IEEE J. Oceanic Eng., vol. OE-31, pp. 325–344.

Zolotarev, V. M. and B. M. Strunin, 1971, "Internal stress distribution for a random distribution of point defects", Sov. Phys. Solid State, vol. 13, no. 2, pp. 481–2.

90 Earthquake Localization

Michael Fehler

Massachusetts Institute of Technology, Cambridge, USA

1 • Introduction

Earthquake localization is one of the fundamental problems of seismology and continues to attract considerable attention. The location, or hypocenter, of an earthquake consists of four parameters, three for its spatial coordinates and one for its occurrence time. Earthquake location is the first and most important information that must be obtained as a prerequisite to further study of an earthquake.

2 • The Location Problem

The location problem is one of triangulation. The distance of an earthquake from a station can be estimated by measuring the arrival time of two seismic phases at the station and using the known velocity of propagation of each phase. For example, if the arrival time of the P and S phases, t_P and t_S, respectively, are known at one station, the distance of the event from that station can be estimated. The difference in arrival time between the phases is

$$t_S - t_P = (D/V_S) - (D/V_P) = D(V_P - V_S)/V_S V_P \tag{1}$$

where V_P and V_S are the P and S wave velocities, respectively, and D is the distance from the station to the event. For the case where Poisson's ratio is 0.25, and $V_P = 6\,\text{km/s}$, we obtain

$$D \sim (t_S - t_P)V_P/(\sqrt{3} - 1) \sim 8(t_S - t_p)\,\text{km} \tag{2}$$

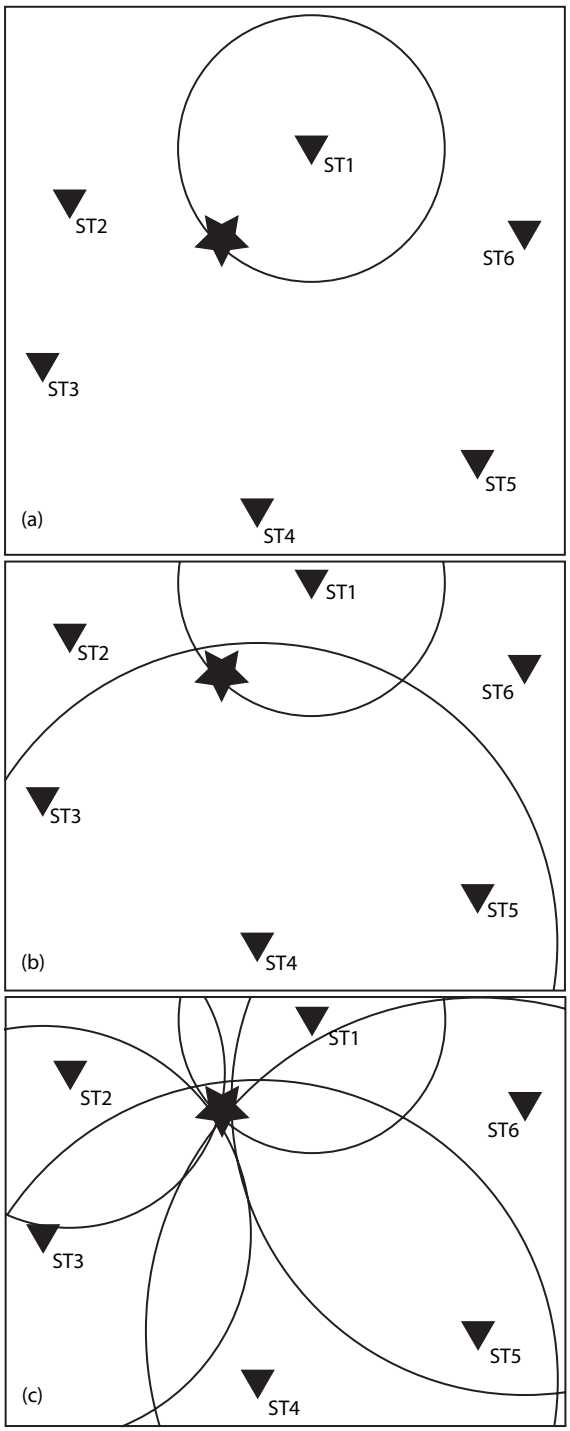


FIGURE 1 Locating an earthquake using the distance of the event from stations estimated from difference in arrival times of P and S phases. (**a**) A set of stations, called ST1–ST6, from an array that detects an earthquake. The distance between station ST1 and the event is determined using arrival time differences. The event falls on a sphere centered on the station location. The projection of the sphere onto a 2D geometry is shown. Recording station locations are marked by triangles and the true earthquake location is marked by a star. (**b**) Projections of spheres from two recording stations onto a 2D space to show the possible locations for the earthquake determined from each station. The most likely location for the earthquake is one of the points where the two spheres intersect, here shown as two points. (**c**) Projections of spheres from six stations whose distances from an earthquake are determined from arrival time differences of phases at each station. If all data were perfect and velocities were perfectly known, the spheres would intersect at the location of the earthquake. Since data and model are not perfect, spheres do not intersect at one point and the earthquake is considered to have occurred in the vicinity of the region where the spheres come closest to intersecting.

This provides a simple rule of thumb for estimating the distance to an earthquake from a single station. The estimated distance from a station defines a sphere of possible earthquake locations relative to the station as shown in Figure 1(a). Figure 1(b) shows the projections of two circles around stations illustrating possible locations for an earthquake using data from each station. The two points where the circles intersect are the most likely locations of the earthquake occurrence as inferred from the data obtained at the stations. If distances can be estimated from a number of stations, the event location can be determined by finding the intersection of spheres drawn about each station. Figure 1(c) shows projections of spheres from six stations. If all the spheres intersected in one location, that would be the earthquake location. Since there is noise in the data due to errors in estimating the arrival times of phases and uncertainty in the velocity model, including uncertainties in the velocities and lack of information about spatial heterogeneity of velocity, the spheres do not intersect at one point. The uncertainty in the earthquake location can be estimated from how closely all spheres intersect at one location. Minimizing this uncertainty is a significant goal of advanced methods for locating earthquakes.

3 • Geiger's Method

The conventional numerical approach to earthquake location is called *Geiger's method* [1], which is a straightforward inverse problem (see e.g. Lee and Stewart [2] or Menke [3]). Using the four unknowns for an earthquake location, the three spatial coordinates (x, y, z) and the origin time O, we write the arrival time of a given phase i at a given station j as

$$t_{ij} = O + D_j(x, y, z)/V_i \tag{3}$$

where $D_j(x, y, z)$ is the distance between the station and the earthquake and V_i is the propagation velocity of phase i. In practice, the velocity can vary along the ray path from earthquake to station, so the travel time between the two locations is determined using ray tracing [2].

Implementation of Geiger's method can be considered an optimization problem where the values of (x, y, z) and O are determined that give the best fit to the observed arrival times of all phases at all stations. An iterative least-squares inverse approach is often used where the location and origin time is adjusted iteratively to find the best fit between predicted and observed arrival times in the least-squares sense.

4 • Grid Search

Another popular approach for solving the optimization problem is a grid search algorithm [4]. A grid is established of possible earthquake locations within a region. Travel times between each grid point and each station are computed for all seismic phases that are likely to be observed. Then using each grid point as a trial location, the origin time is estimated independently for each phase. The grid location that gives the most consistent origin time for all observed arrival times at all stations is taken as the correct location of the earthquake.

5 • Methods to Reduce Uncertainty

Errors in arrival time estimation, misidentification of phases, uncertainty in the earth velocity structure, poor station geometry, and earthquake location procedure all contribute to uncertainty in location. The uncertainty in earthquake location can be specified by a confidence ellipsoid surrounding the most likely location.

There have been a considerable number of approaches for improving on the earthquake locations and reducing location errors. Some provide better absolute locations of events and some provide better relative locations of events within a localized region. Much effort has focused on trying to mitigate the problem of error in the earth velocity model used for calculating propagation times between earthquakes and stations. One of the first successful approaches was the use of a station time correction applied to all arrival times of a given phase at a given station. The correction would be included in Equation (3) as an extra additive term S_{ij}. Using a suite of events from the same geographic region, the value of S_{ij} would be selected for each station (j) and phase (i) to eliminate any bias in the predicted arrival times. This method is known as *joint hypocenter determination (JHD)* [5] and is a common way to improve the relative locations of events without knowing details of a velocity structure. This methodology has been adapted for relatively large regions by using a suite of well-located *ground truth* events within the *JHD* framework and determining more detailed forms of station corrections that vary spatially within a region [6]. Such well-located events may include manmade shots such as chemical or nuclear explosions, mining blasts, or earthquakes whose locations can be independently determined from geological data.

A rather new innovation for improving the determination of the relative locations of earthquakes involves the use of precise arrival time differences measured between events. These arrival time differences may be measured more reliably than the absolute arrival times in many cases because the waveforms of different events recorded at a single station often have similarities. The relative arrival times may be estimated from visual inspection of the data [7] or from cross-correlation techniques [8]. A method known as the *double difference method* involves a reformulation of the earthquake location problem in terms of the arrival time differences at stations rather than the absolute arrival times [9]. This method has gained popularity due to its ability to give significant improvements in the relative locations of earthquakes, thus revealing fault structures more reliably. This method has been analyzed in more detail by Wolfe [10] who compared it with other methods [11,12] and concludes that all the methods have the ability to eliminate the influence of unknown heterogeneous velocity structure on the location determination of closely spaced events but that the methods are not successful when event locations are widely separated. A related method for improving the relative locations of earthquakes is known as *collapsing* [13], in which earthquake locations are shifted within their uncertainty ellipsoid toward the locations of other earthquakes. The idea is that location errors caused by errors in arrival times, poor station geometries, or the location procedure will tend to decluster earthquake locations. This methodology has also been introduced within the framework of JHD [14].

Figure 2 shows the locations of a suite of events that were induced by hydraulic stimulation of a 3 km deep wellbore in Northern New Mexico [14]. The events were located using P and S arrival times measured at four stations. Figure 2(a) shows the locations determined using a single-event location method that is similar to Geiger's method. The locations appear to occur in a diffuse zone and it is difficult to identify clear spatial patterns. Figure 2(b) shows the locations determined using a JHD method that included the estimation of P and S station corrections, S_{ij}, for each of the four stations. The JHD approach yields a less diffuse zone of earthquake locations and there appear to be a few rather linear features where locations preferentially occur. Figure 2(c) shows the locations determined using the JHD collapsing approach described in [14]. The locations determined using this approach show more localization along probable lineaments, which are considered to be fractures in the hydraulically stimulated region. Figure 2(d) shows an expanded view (note

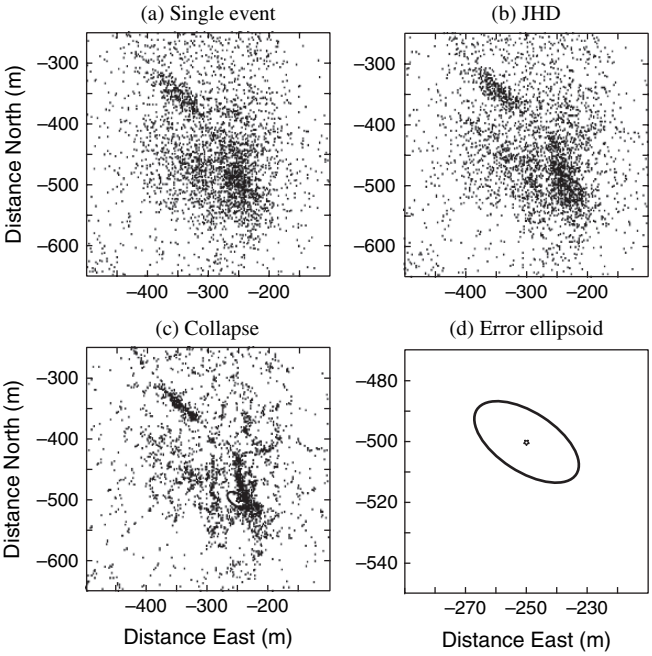


FIGURE 2 Two-dimensional projections of the locations of a set of earthquakes that were induced by hydraulic stimulation of a wellbore drilled into crystalline rock. Projections are onto a North–South and East–West plane. (**a**) Locations determined individually for events using a method that is similar to Geiger's method, (**b**) locations determined using joint hypocenter determination (JHD), (**c**) locations determined using a method that combined JHD and collapsing [14], and (**d**) projection of error ellipsoid onto the 2D plane for the event located at −500 m North and −250 m East. Note that the scale of (d) is different from the other frames. The location of the event is indicated by a star. The error ellipsoid is also shown in (c).

difference in scale from Figure 2(a)–2(c)) of the projection of the error ellipsoid onto the two-dimensional (2D) plane of the locations. The error ellipsoid defines the formal range over which the true location can be considered to occur.

The minimum number of arrival times required for an earthquake location is four since there are four unknown hypocentral parameters. Additional arrival times can be used to estimate uncertainty in the hypocentral parameters but can also be used to determine the velocity structure of the region of the earth where the earthquakes occurred. An approach for doing this was first outlined in 1979 [15] and has been improved and extensively applied since that time. Such approaches combine the earthquake

location problem with the velocity tomography problem that is described in the next chapter (Seismic Tomography). Variations have also been developed and applied, including one recent method which combines the double difference location approach with travel time tomography [16].

References

1. Geiger, L., Probability method for the determination of earthquake epicenters from arrival time only, *Bull. St. Louis Univ.* **8**, 60–71 (1912).
2. Lee, W.H.K., and S.W. Stewart, *Principles and Applications of Microearthquake Networks,* Academic Press, New York (1981).
3. Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory,* Academic Press, New York (1984).
4. Sambridge, M., and B. Kennett, A novel method of hypocenter location, *Geophys. J. Roy. Astron. Soc.* **87**, 679–697 (1986).
5. Douglas, A., Joint epicenter determination, *Nature* **215**, 47–48 (1967).
6. Myers, S., and C. Schltz, Improving sparse network seismic location with Bayesian kriging and teleseismically constrained calibration events, *Bull. Seismol. Soc. Am.* **90**, 199–211 (2000).
7. Phillips, W.S., L. House, and M. Fehler, Detailed joint structure in a geothermal reservoir from studies of induced microearthquake clusters, *J. Geophys. Res.* **102**, 11745–11763 (1997).
8. Aster, R., and C. Rowe, Automatic phase pick refinement and similar event association in large seismic datasets, *Advances in Seismic Event Location*, Kluwer Academic, The Netherlands, 231–263 (2000).
9. Waldhauser, F., and W. Ellsworth, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.* **90**, 1353–1368 (2000).
10. Wolfe, C., On the mathematics of using difference operators to relocate earthquakes, *Bull. Seismol. Soc. Am.* **92**, 2879–2892 (2002).
11. Jordan, T., and K. Sverdrup, Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific, *Bull. Seismol. Soc. Am.* **71**, 1105–1130 (1981).
12. Got, J., J. Frechet, and F. Klein, Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. Geophys. Res.* **99**, 15375–15386 (1994).
13. Jones, R., and R. C. Stewart, A method for determining significant structures in a cloud of earthquakes, *J. Geophys. Res.* **102**, 8245–8254 (1997).
14. Fehler, M., W. Phillips, L. House, R. Jones, R. Aster, and C. Rowe, A method for improving relative earthquake locations, *Bull. Seismol. Soc. Am.* **80**, 775–800 (2000).

15. Aki, K., and W.H.K. Lee, Determination of three-dimensional velocity under a seismic array using P arrival times from local earthquakes, *J. Geophys. Res.* **81**, 4381–4399 (1976).
16. Zhang, H., and C. Thurber, Double-difference tomography: The method and its application to the Hayward Fault, California, *Bull. Seismol. Soc. Am.* **93**, 1875–1889 (2003).

91

Seismic Tomography

Robert L. Nowack and Cuiping Li

Department of Earth and Atmospheric Science, Purdue University, West Lafayette, IN 47907, USA

1 • Introduction

The inversion of seismic travel-time data for radially varying media was initially investigated by Herglotz, Wiechert, and Bateman (the HWB method) in the early part of the 20th century [1]. Tomographic inversions for laterally varying media began in seismology starting in the 1970's. This included early work by Aki, Christoffersson, and Husebye who developed an inversion technique for estimating lithospheric structure beneath a seismic array from distant earthquakes (the ACH method) [2]. Also, Alekseev and others in Russia performed early inversions of refraction data for laterally varying upper mantle structure [3]. Aki and Lee [4] developed an inversion technique using travel-time data from local earthquakes. The first global tomographic images were then obtained using both body waves and surface waves [5,6]. In this chapter, we provide an overview of seismic tomography mostly applied to seismic body waves. We first describe travel-time tomography using different approaches including variable damping, iterative methods, irregular gridding, autoregressive extrapolation, nonlinear inversions, and multi-grid methods. We next provide an overview of applications of seismic tomography to cross-hole, refraction and reflection data, local earthquake data, and teleseismic data. An example is given for the inversion of data from the TomoVes experiment to image beneath the Mt. Vesuvius volcano. Finally,

a brief discussion of seismic attribute, waveform, and finite-frequency tomography for seismic body waves is given.

2 • Seismic Travel-Time Tomography

The goal of seismic travel-time tomography is to find a velocity model for a subsurface volume consistent with measured travel-times along ray paths that pass through the volume. The travel-time for high-frequency seismic waves is given by

$$T_i = \int_{\Gamma_i} \frac{1}{v(x,y,z)}\, d\vec{x} = \int_{\Gamma_i} s(x,y,z)\, d\vec{x} \tag{1}$$

where $v(x,y,z)$ is the velocity distribution, T_i is the observed travel-time, Γ_i is the ray path, and $s(x,y,z)$ is the slowness or inverse of velocity. The velocity or slowness fields can be represented by a limited set of parameters after discretization with a set of basis functions. The problem can then be written as

$$d_i = G_i(x_j) \tag{2}$$

where d_i are the observed travel-times and x_j are the model parameters. Ray tracing algorithms and model parameterization are essential to seismic tomography and a detailed overview is given in [7]. Here we mostly describe the tomographic inversion aspects.

The tomography problem can be linearized about an initial model x_0 with corresponding predicted travel-times T_0 as

$$A\,\delta x = \delta d \tag{3}$$

where $\delta d = T - T_0$ is the vector of travel-time residuals (observed minus predicted travel-times), $\delta x = x - x_0$ is the vector of model deviations from the initial model, and $A_{ij} = \partial G_i / \partial x_j$ are the components of the sensitivity matrix A evaluated for the model x_0. Often, this set of equations is operationally under-determined where $(A^T A)$ is rank deficient. In this case the observed data alone cannot completely resolve the model and the inverse solution requires some kind of regularization. For under-determined problems, one approach is to use a combined cost function which minimizes both the observed minus predicted data residuals and some norm of the model variations. For simple damped least squares (DLS), the regularization can be written as

$$\begin{bmatrix} A \\ \varepsilon I \end{bmatrix} \delta x = \begin{bmatrix} \delta d \\ 0 \end{bmatrix} \tag{4}$$

where ε is a small, positive damping parameter. Solving this combined system using least squares results in

$$\delta x = (A^T A + \varepsilon^2 I)^{-1} A^T \delta d \tag{5}$$

where T indicates the transpose. In general, this results in a compromise in finding a model that fits the data and has model deviations that are not too large.

In Eq. (4), the model parameters all have the same weight. But for most seismic tomographic experiments some weighting of the travel-times and model parameters is required. To overcome the shortcomings of the DLS method, weighting matrices W_1 for the data and W_2 for the model which are symmetric and positive definite can be used. An objective function can then be written as

$$S = (\delta d - A\,\delta x)^T W_1 (\delta d - A\,\delta x) + \varepsilon^2 \delta x^T W_2\, \delta x \tag{6a}$$

where the diagonal elements of W_1 and W_2 allow for nonuniform weighting of the data and model parameters, and the off-diagonal elements account for data correlations and model smoothing. These also result in modified inner products of the data and model spaces. Hansen [8] proposed an L-curve approach to obtain an appropriate value for ε^2. In this approach the log of the normed data misfit is plotted against the log of the normed model deviation. The trade-off parameter ε^2 at the 'knee' of this curve provides a trade-off between tightly fitting the observed data (as ε^2 goes to zero) and minimizing a norm of the model deviations (as ε^2 goes to infinity) (Figure 1). With this value of ε^2, minimizing Eq. (6a) gives

$$\delta x = (A^T W_1 A + \varepsilon^2 W_2)^{-1} A^T W_1\, \delta d \tag{6b}$$

For the above linear tomographic problem, a number of different forms of the weighting matrices W_1 and W_2 have been suggested. Using a Bayesian approach, these can be related to the data and prior model covariances as $C_D = E[\delta d\, \delta d^T] = W_1^{-1}$ and $C_{x_0} = E[\delta x\, \delta x^T] = \varepsilon^{-2} W_2^{-1}$ [9]. In this approach, the trade-off parameter is absorbed into the prior model covariance. But the model covariance must be strictly based on prior information from a Bayesian standpoint [9]. Nonetheless, Eq. (6b) is often used with weighting matrices applied operationally.

For nonuniform ray coverage and variable model cell sizes, the weighting matrices can be chosen to approximately compensate for these effects [10–12]. However, additional damping may still

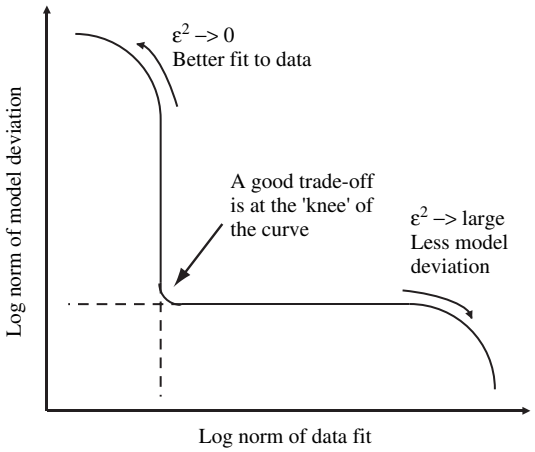


FIGURE 1 The generic form of the L-curve. The L-curve displays the compromise between minimizing the model deviation and minimizing the data misfit. The trade-off parameter can be estimated from the "knee" of the L-curve [8].

be required to avoid instabilities for cells with few rays. An example is shown in Figure 2 in which a cross-hole geometry is used to illustrate the effects of variable damping based on cell size and ray length in a given cell. Figure 2A shows the true model which is a linear gradient from 5 to 5.8 km/s. Figure 2B shows the homogeneous starting model of 5 km/s. Also shown in Figure 2B are the locations of the spline interpolated velocity nodes of the model displayed as solid circles. The rays are shown by white lines with the sources displayed as open circles and the receivers by open triangles on the left and right boundaries. This results in variable ray coverage with dense ray coverage in the upper part of the model but only two rays in the lower half of the model. Also, the cell sizes for the nodes along the boundaries of the model are effectively smaller than those for the interior nodes. The partial derivatives are computed from integrations of the perturbed travel-times along the initial rays for small perturbations of the spline interpolated nodes.

A linearized inversion for the model using simple damping is shown in Figure 2C where both W_1 and W_2 are identity matrices and ε^2 is chosen to give the best overall magnitude of model perturbations. As can be seen, the unequal ray density and different cell sizes result in differences of the inverted model from the true model. These effects can be minimized by adding variable damping terms to W_2. First a term related to the square-root of the diagonal elements of $(A^T A)$ is used to account for the

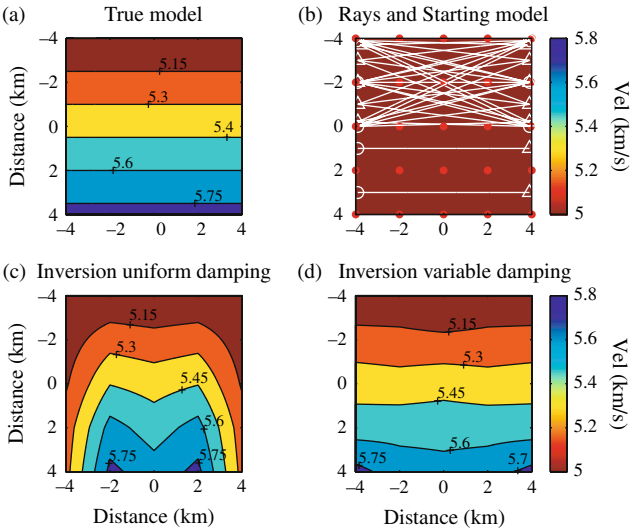


FIGURE 2 Tomographic example using unequal cell sizes and variable ray coverage. (**a**) The true model. (**b**) The starting model with ray coverage and nodes positions with the boundary nodes having smaller cells. (**c**) The inversion with uniform damping. (**d**) The inversion with variable damping.

ray lengths within different cells of the model [11]. Also, a small amount of simple damping is added to this for stability. This is then multiplied by a term which equalizes for the small and large block sizes of the model parameterization [10]. Finally, ε^2 is chosen to give the best overall magnitude of the model perturbations. Figure 2D shows the result of using this variable damping to balance the variability in ray coverage and cell size. Although this is a simple example, it illustrates the utility of a variable damping approach.

As an alternative approach, Deal and Nolet [13] suggested a nullspace shuttle technique for constraining poorly resolved parts of the model which simultaneously conserves the data fit and incorporates an a priori model into the solution. The null space component of the prior model is preserved when the data is insufficient.

2.1 Iterative Inversion for Linearized Problems

The solution of Eq. (6b) requires solving a matrix equation which can be prohibitively time consuming for large numbers of model parameters. Large-scale linear problems are usually solved using iterative methods and there are two popular iterative methods used for tomography, SIRT (simultaneous iterative reconstruction technique) and LSQR, a modified conjugate

gradient method. Dines and Lytle [14] introduced the SIRT method from medical imaging to geotomography. However, Van der Sluis and Van der Vorst [15] showed that the SIRT method does not solve the original inverse problem, Eq. (3), but instead solves a weighted least squares problem in modified data and model norms. Trampert and Leveque [16] also analyzed these properties of SIRT. In a similar fashion to using variable damping, different least squares solutions can be obtained if the metrics in the original and transformed model and data spaces are changed.

The LSQR method belongs to the family of conjugate gradient methods and was introduced by Nolet [17] to seismic tomography problems. The LSQR approach seeks the solution δx that minimizes the $||A\,\delta x - \delta d||_2$ subject to also minimizing $||\,\delta x\,||_2$. Van der Sluis and Van der Vorst [14] showed that LSQR starts the iterative process by ignoring those components associated with the smallest eigenvalues of $(A^T A)$. After more iterations, the contributions of the smaller eigenvalues eventually enter the solution. This property contributes to the intrinsic damping of the algorithm.

Another iterative approach is to use subspace methods in which the iterative process is carried out over higher dimensional subspaces of the model. An early description of this method can be found in Kennett et al. [18] (for more details see [7]).

2.2 Irregular Grid Inversion

Seismic tomography using a regular grid and uniform damping may have undesirable properties in the presence of inhomogeneous sampling of the model by the ray paths, and may result in an unnecessary over-parameterization of the model space [19,20]. The use of irregular grids may be less sensitive to regularization issues than inversions based on regular grids. Curtis and Snieder [21] suggested a genetic algorithm to do an irregular parameterization of the model space by searching a configuration of Delaunay triangles that result in the best conditioned inverse problem. Sambridge and Gudmundsson [22] proposed an irregular cell approach based on the tetrahedral and Voronoi diagrams and gave several options to do the irregular cell design. Spakman and Bijwaard [19] optimized the irregular cell parameterizations by equalizing the ray hit-counts in each cell. Sambridge and Faletic [23] recently proposed an adaptive tetrahedral cell scheme based on the maximum spatial gradients in the seismic velocity across each tetrahedral face. For their adaptive method, the study region was initially divided into a uniform tetrahedral mesh. The scheme then refined the parameterization during the inversion in response to the structural features detected.

2.3 Autoregressive Extrapolation

Autoregressive (AR) extrapolation methods have been used in signal processing to predict time series, as well as perform prediction-error filtering. Claerbout [24,25] suggested a two-stage scheme to apply autoregressive techniques to the linear least-squares problem. The first stage is to find the optimal prediction-error filter (PEF). For the second stage, the PEF is assumed to be known, and values for the missing data are found so that the power out from the prediction-error filter is minimized. This is similar to maximizing an entropy criterion [26].

In terms of an unknown model, one approach to apply the AR technique to the seismic inversion problem is to separate the problem into two steps. For model extrapolation, this requires some information about the model in order to compute the PEF. First nodes with few or no rays nearby are removed from the inversion, and only nodes with good ray coverage are inverted. Then, making use of the initial inversion result, the AR interpolation technique can be used to obtain the values for the unconstrained model nodes.

An alternate approach is to apply AR extrapolation directly to the travel-time data instead of the model. In this approach, the known data is first extrapolated to estimate the unknown data. Then, the combined known and extrapolated data are inverted for the model parameters. Using this approach, the missing data and the prediction-error filter are jointly estimated using the known data. The combined observed and extrapolated data are then used to invert for the model. This can be written as

$$F \begin{bmatrix} \delta d_a \\ \delta d_b \end{bmatrix} = r_2 \approx 0 \tag{7}$$

and

$$A\,\delta x - \begin{bmatrix} \delta d_a \\ \delta d_b \end{bmatrix} = r_1 \approx 0 \tag{8}$$

where in Eq. (7) the missing data δd_b and the PEF, F, are found from the known data δd_a by minimizing the residual r_2. In Eq. (8) the model deviations δx are then found to minimize r_1 using the known data and δd_b the extrapolated data. The extrapolated data play the role of regularization terms in Eq. (8). Equation (7) is nonlinear for the unknown F and δd_b, so this can be solved in an iterative fashion [25], and then Eq. (8) is solved for δx. Li and Nowack [27] used this approach to apply autoregressive extrapolation to the seismic tomography problem.

2.4 Nonlinear Inversion

The tomography problem can become nonlinear under large velocity perturbations if a velocity (rather than slowness) representation is used, or if the ray paths are substantially perturbed

by the velocity changes. In nonlinear inversion, the objective function to be minimized can be written as

$$S = (d - G(x))^T W_1 (d - G(x)) + \varepsilon^2 (x - x_0)^T W_2 (x - x_0) \quad (9)$$

where d is the data, $G(x)$ is the nonlinear forward problem, x is the model, x_0 is the prior model, and W_1 and W_2 are weighting matrices similar to those given earlier for the linear case. A linearization approach can be applied iteratively to solve the nonlinear problem [9]. However, the regularization is now with respect to the prior model x_0 instead of the solution at the previous iteration. For each linearization step, the linearized problem can be solved either by direct matrix methods or by iterative methods. For the error analysis, the linearized covariance and resolution operators are often computed for the converged final solution.

For strongly nonlinear problems, linearized gradient methods may fail because of a multi-minimum objective function and therefore a good starting solution is required. Alternatively, global optimization methods can be applied including MCMC (Markov Chain, Monte Carlo), simulated annealing, and genetic algorithms. A recent survey of Monte Carlo methods applied to geophysical inverse problems was given by Sambridge and Mosegaard [28]. However, Rawlinson and Sambridge [7] noted that computational expense may still limit the use of global optimization methods for seismic tomography and other inverse problems when the number of model parameters is large.

2.5 Multigrid Inversion

Traditional iterative inversion methods cannot always invert seismic data obtained from complicated earth models because of variable convergence of the high and low wavenumber components of the solution. Multigrid approaches to seismic tomography apply preliminary iterations on a coarse grid and then use this solution as an initial model on a finer grid. The use of an initial coarse grid can reduce the low wavenumber components of the model more efficiently and also allow for a variable ray coverage.

Nemeth et al. [29] summarized the following multigrid strategy for the travel-time tomography problem. First an initial iteration is performed on a smoothed model with few unknowns. After several iterations the convergence rate is likely to stall, and then the gridded model is refined by a factor of 2. The model is smoothed after refinement and new iterations are performed until convergence again stalls. This is repeated until final convergence occurs. Simple applications of a multi-grid approach were applied by Lutter and Nowack [30] for travel-times and

Nowack and Matheney [31] for mixed seismic attributes. Zhou [32] applied multi-scale tomography to cope with poor raypath coverage and solve for velocity interfaces. For global travel-time data, Chiao and Kao [33] used a wavelet approach.

3 • Examples of Travel-Time Tomography

In this section, we provide examples of travel-time tomography inversions applied to cross-borehole, refraction and reflection data, local earthquake data, and global teleseismic data. A particular example is given from the TomoVes experiment to image the shallow structure beneath the Mt. Vesuvius volcano.

3.1 Cross-Borehole, Refraction, and Reflection Tomography

In cross-borehole tomographic imaging, sources are placed in one borehole and possibly along the surface and the receivers are placed in a nearby borehole. Several examples of seismic ray tomography applied to cross-borehole data include McMechan [34], Peterson et al. [35], and Bregman et al. [36]. Menke [37] showed that cross-borehole geometries have less resolution in the crosshole direction depending on the source and receiver apertures. Lo and Inderseisen [38] provide case studies of cross-borehole tomography including monitoring steam flood injections in oil fields and imaging fault systems. Pratt and Chapman [39] and Pratt et al. [40] describe travel-time tomographic applications for anisotropic media.

Examples of 2D refraction tomography include that of White [41], Lutter et al. [42], and Zelt and Smith [43]. Lutter and Nowack [44] inverted for velocities and interface depths in a layer stripping fashion. Examples of 3D refraction tomography for crustal velocities are given by Hole [45] and Zelt and Barton [46]. Extensive reviews of refraction and wide-angle tomography are given by Nowack and Braile [47] and Rawlinson and Sambridge [7].

Examples of travel-time tomography used in reflection seismology include Bishop et al. [48] and Farra and Madariaga [49]. However, there can be an ambiguity between velocities and reflector depths as described by Williamson [50] and Stork [51] unless further a priori information is provided. The resolution is best near the surface and decreases with depth depending on the source–receiver offset range. An application of tomography to reflection seismology called stereotomography was given by Billette and Lambare [52] and uses travel-time and travel-time slope information with offset to estimate smooth velocity models. A recent assessment of this method was given by Lambare [53].

3.2 Inversion of TomoVes Experiment Data

In this example, seismic refraction data from the TomoVes experiment [54] is used to invert for shallow structure beneath the Mt. Vesuvius volcano. The imaging of the shallow structure beneath an active volcano was one of the scientific motivations of this experiment. The source and receiver geometry for the experiment is displayed in Figure 3A and can be seen to be very heterogeneous. Because of the large topographic relief in the area, a 3-D

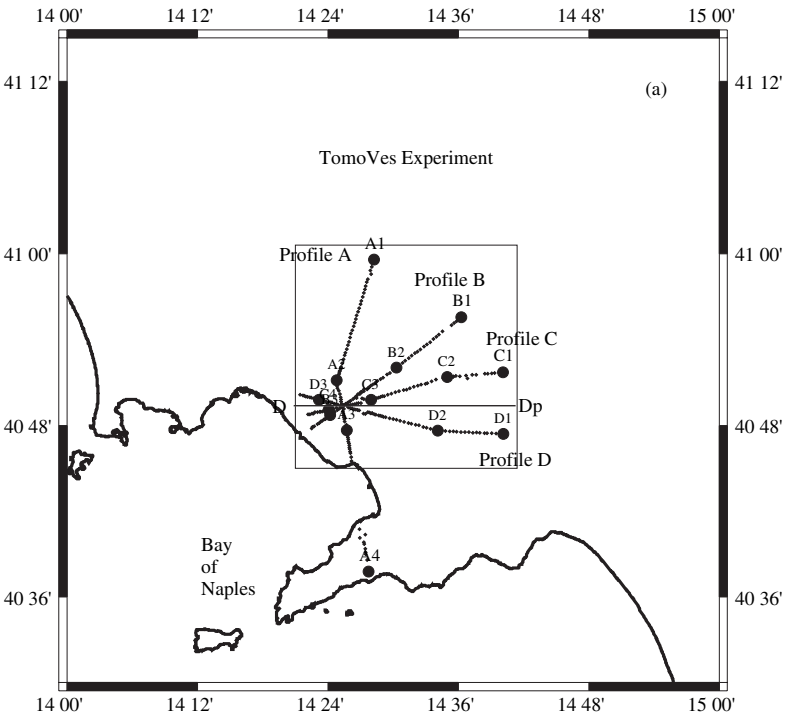


FIGURE 3 (**a**) Map of the geometry of the refraction lines from the TomoVes tomography experiment [54] going across the top of the Mt. Vesuvius volcano. The *solid dots* are the shot points and the different intersecting profiles of receivers are noted by *small dots*. The *line* D-Dp shows the location of the vertical cross-section through the inverted 3D velocity model from Nowack et al. [55,56]. (**b**) The topography and 3D node model for the Mt. Vesuvius region corresponding to the inset square in (a). The volcano is on the left and the background hills to the upper right. The *solid dots* show the tomographic nodes position used for the 3-D inversion. Rays from one shot point are also shown diving beneath the volcano. (**c**) This shows a cross-section of the tomographic velocity model from [55,56] along line D-Dp in (a).

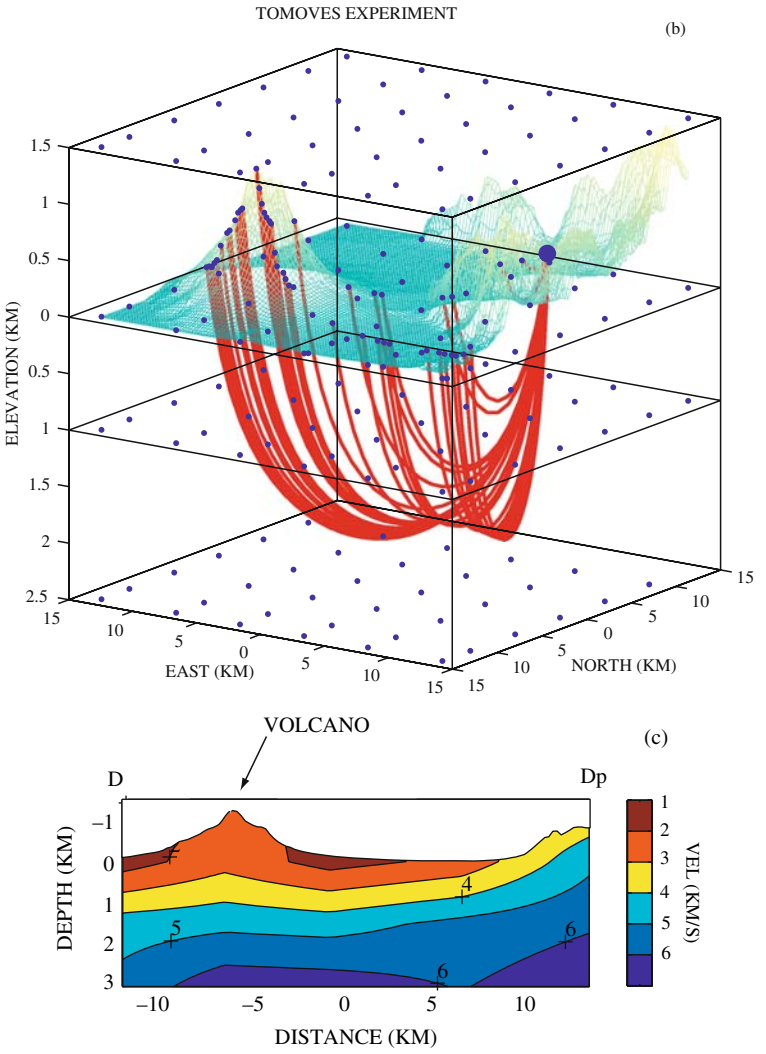


FIGURE 3 (continued)

uniform velocity grid model shown by the dots in Figure 3B was used. The velocities have been spline interpolated. The rays for one shot point are shown diving beneath the volcano. A variable damping inversion technique was applied to the first arrival travel-times and other seismic attributes of the TomoVes data by Nowack et al. [55,56]. Figure 3C illustrates the 3-D inversion result after 3 iterations shown along profile D-Dp in Figure 3A. Low surface velocities can be seen surrounding the volcano, as

well as lower velocities in the edifice of the volcano. At depths of 1.0–2.0 km, the 4 and 5 km/s velocity contour lines are seen to be shallow toward the background hills to the east and the corresponding formations are fairly continuous beneath the volcanic plain. Also, the 3 and 4 km/s contours tend to shallow directly beneath the volcano and this has been suggested to be a paleo-volcanic or sub-volcanic structure [57]. Although the derived 3D velocity model is smoothly varying, the results are similar to the more detailed velocity models along the individual 2D profiles obtained by Zollo et al. [57].

3.3 Local Earthquake Tomography

Local earthquake tomography developed as an extension of earthquake location estimation, and early examples include Aki and Lee [3] and Thurber [58]. Kissling [59] provided a detailed review of the method and Eberhart-Phillips [60] extended the method to S-waves. Hirahara [61] described using local earthquakes and teleseisms for both velocity and anisotropy estimation.

A new approach for local earthquake tomography has been the use of double-difference techniques. Zhang and Thurber [62] described an approach in which both differential arrival times as well as absolute travel-times are used for local earthquake tomography. The use of differential travel-times can improve the velocity imaging and earthquake locations by canceling some of the path effects outside the source region.

3.4 Global Tomography

Early inversions for the entire mantle were performed by Clayton and Comer [63] and Dziewonski and Anderson [64], using body waves, and by Woodhouse and Dziewonski [65], using surface waves. Since then numerous P-velocity and S-velocity global tomographic studies using both body waves and surface waves have been conducted. Recent overviews include Iyer and Hirahara [6], Kennett and Van der Hilst [66], Romanowicz [67] and Curtis and Snieder [68]. Cara [69] gives an overview of seismic anisotropy and discusses the trade-offs between anisotropy and velocity heterogeneity as well as possible anisotropy-induced artifacts on isotropic tomographic inversions.

4 • Extensions of Travel-Time Tomography

In addition to travel-times, other seismic attributes including amplitudes and pulse frequencies have been used to perform seismic tomography. Thomson [70] used body wave amplitudes to invert for velocity structure beneath the Norsar seismic array

in Norway. Neele et al. [71] performed a joint inversion of travel-times and amplitudes for upper mantle velocity structure. Using a cross-borehole geometry, Bregman et al. [72] performed an initial travel-time inversion and then used focusing-corrected amplitudes to invert for seismic attenuation. For wide-angle refraction data, Nowack and Matheney [73] inverted for both velocity and attenuation using a combined data set of travel-times, amplitudes, and instantaneous pulse frequencies. Body wave spectral ratios could also be used for the estimation of attenuation. Mitchell and Romanowicz [74] include a number of studies on the estimation and inversion of seismic attenuation. Wang [75] applied seismic amplitude inversion to reflection seismology.

The Fresnel zone depends on the seismic wavelength and the path length, and if the heterogeneity scales are on this order then wave diffraction will occur. If the heterogeneities are weak then a single scattering model, the Born approximation, can be used to model seismic wave diffraction. An early application of diffraction tomography to seismic waves was presented by Devaney [76]. Other examples include and Wu and Toksoz [77] and Pratt and Worthington [78] who used waveform tomography in the frequency-domain for cross-hole data. An overview of diffraction tomography was given by Lo and Inderwiesen [38]. Pratt et al. [79] applied waveform tomography in the frequency-domain to wide-angle seismic data, and Brenders and Pratt [80] used this approach to successfully invert for small-scale features from a synthetic, wide-angle data set. However, a starting travel-time inversion was required to provide a smooth initial velocity model for the waveform inversion.

Recently, diffraction kernels for travel-times have been investigated which allow for wavefront healing of finite frequency wavefronts. This can occur for low-frequency seismic waves having long travel paths, such as for waves traveling through the mantle. Early examples of computing Fresnel zones for finite-frequency effects include Woodward [81] and Yomogida [82]. Dahlen et al. [83] derived a fast way to compute kernels relating travel-time perturbations measured by cross-correlation of waveforms to velocity anomalies using the paraxial approximation. This was used by Montelli et al. [84] to perform finite-frequency tomography of P and PP correlation times for mantle velocity structure, and a number of hotspots fed by plumes from the lower mantle were found. They compared their results with ray tomography and concluded that finite-frequency tomography provided clearer images when using long-period seismic waves. However, this was questioned by van der Hilst and de Hoop [85] who suggested that the benefits of the more

complete theory may still not yet be realized in global tomography because of practical considerations such as damping, weighting of different data, choice of the data fit, and limited ray coverage. Although discussion of these issues has continued (Montelli et al. [86] and van der Hilst and de Hoop [87]), it is clear that further developments to improve ray coverage, incorporate finite frequency effects and include other seismic wavefield attributes will result in improved tomographic images of the Earth's interior.

References

1. K. Aki and P. G. Richards: Quantitative Seismology, (Freeman, San Francisco 1980)
2. K. Aki, A. Christoffersson and E. S. Husebye: Determination of three-dimensional seismic structure of the lithosphere, J. Geophys. Res. 82 (1977) 277–296
3. A. S. Alekseev, M. M. Lavrentiev, V. G. Romanov and M. E. Romanov: Theoretical and computational aspects in seismic tomography, Surv. Geophys. 11 (1990) 395–409
4. K. Aki and W. H. K. Lee: Determination of three-dimensional velocity anomalies under a seismic array using first P-arrival time from local earthquakes 1. A homogeneous initial model, J. Geophys. Res. 81 (1976) 4381–4399
5. G. Nolet (ed.): Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics, (Reidel, Hingham MA 1987)
6. H. M. Iyer and K. Hirahara (eds.): Seismic Tomography Theory and Practice, (Chapman and Hall, London 1993)
7. N. Rawlinson and M. Sambridge: Seismic traveltime tomography of the crust and lithosphere, In: Advances in Geophysics Vol. 46 ed. by R. Dmowska, (Elsevier, Amsterdam 2003) 81–198
8. P. C. Hansen: Rank-deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, (SIAM, Philadelphia 1998)
9. A. Tarantola: Inverse Problem Theory and Methods for Model Parameter Estimation, (SIAM, Philadelphia 2005)
10. G. Nolet: Seismic wave propagation and seismic tomography, In: Seismic Tomography ed. by G. Nolet, (Reidel, Hingham MA 1987) 1–23
11. B. Wang: Improvement of Seismic Travel-time Inversion Methods and Application to Observed Data, (Ph.D. thesis, Purdue University, West Lafayette, IN 1993)
12. J. G. Berryman: Weighted least-squares criteria for seismic travel-time tomography, IEEE Trans Geosci Remote Sens 27 (1989) 302–309
13. M. M. Deal and G. Nolet: Nullspace shuttles, Geophys. J. Int. 124 (1996) 372–380

14. K. A. Dines and R. J. Lytle: Computerized geophysical tomography, Proc. IEEE 67 (1979) 1065–1073
15. A. van der Sluis and H. A. van der Vorst: Numerical solution of large, sparse linear systems arising from tomographic problems, In: Seismic Tomography, ed. by G. Nolet, (Reidel, Hingham MA 1987) 49–83
16. J. Trampert and J. Leveque: Simultaneous iterative reconstruction technique: physical interpretation based on generalized least squares solution, J. Geophys. Res. 95 (1990) 12553–12559
17. G. Nolet: Solving or resolving inadequate and noisy tomographic systems, J. Comp. Phys. 61 (1985) 463–482
18. B L. N. Kennett, M. S. Sambridge and P. R. Williamson: Subspace methods for large scale inverse problems involving multiple parameter classes, Geophys. J. 94 (1988) 237–247
19. W. Spakman and H. Bijwaard: Optimization of cell parameterizations for tomographic, inverse problems, Pure Appl. Geophys. 158 (2001) 1401–1423
20. N. Rawlinson and M. Sambridge: Irregular interface parameterization in 3-D wide-angle seismic traveltime tomography, Geophys. J. Int. 155 (2003) 79–92
21. A. Curtis and R. Snieder: Reconditioning inverse problems using the genetic algorithm and revised parameterization, Geophysics 62 (1997)1524–1532
22. M. Sambridge and O. Gudmundsson: Tomographic systems of equations with irregular cells, J. Geophys. Res., 103 (1998) 773–781
23. M. Sambridge and R. Faletic: Adaptive whole Earth tomography, Geochemistry Geophysics Geosystems 4 (2003)
24. J. F. Claerbout: Earth Soundings Analysis: Processing Versus Inversion, (Blackwell Scientific, Oxford 1992)
25. J. F. Claerbout: Image Estimation by Example: Geophysical Soundings Image Construction: Multidimensional Autoregression, (web book, http://sepwww.stanford.edu/sep/prof/gee/toc_html 2004)
26. J. P. Burg: Maximum Entropy Spectral Analysis, (Ph.D. thesis, Stanford University, http://sepwww.stanford.edu/theses/sep06 1975)
27. C. Li and R. L. Nowack: Application of autoregressive extrapolation to seismic tomography, Bull. Seism. Soc. Am. 94 (2004) 1456–1466
28. M. Sambridge and K. Mosegaard: Monte Carlo methods in geophysical inverse problems, Rev. of Geophysics 40 (2002) doi10:1029/2000RG000089
29. T. Nemeth, E. Normark and F. H. Qin: Dynamic Smoothing in Crosswell Traveltime Tomography, Geophysics 62 (1997) 168–176
30. W. J. Lutter and R. L. Nowack: Inversion for crustal structure using reflections from the PASSCAL Ouachita experiment, J. Geophys. Res. 96 (1990) 4633–4646
31. R. L. Nowack and M. P. Matheney: Inversion of seismic attributes for velocity and attenuation structure, Geophys. J. Int. 128 (1997) 689–700

32. H. Zhou: Multi-scale traveltime tomography, Geophysics. 68 (2003) 1639–1649
33. L. Chiao and B. Kuo: Multiscale seismic tomography, Geophys. J. Int. 145 (2001) 517–527
34. G. McMechan: Seismic tomography in boreholes, Geophys. J. R. Astr. Soc. 74 (1983) 298–308
35. J. Peterson, B. Paulsson and T. McEvilly: Application of algebraic reconstruction techniques to crosshole seismic data, Geophysics 50 (1985) 1566–1580
36. N. D. Bregman, R. C. Bailey and C. H. Chapman: Crosshole seismic tomography, Geophysics 54 (1989) 200–215
37. Menke: The resolving power of cross-borehole tomography, Geophys. Res. Lett. 11 (1984) 105–108
38. T.W. Lo and P. Inderwiesen: Fundamentals of Seismic Tomography, (Society of Exploration Geophys., Tulsa OK 1994)
39. R. G. Pratt and C. H . Chapman: Traveltime tomography in anisotropic media II. Application, Geophys. J. Int. 109 (1992) 20–37
40. R. G. Pratt, W. J. McGaughey and C. H. Chapman: Anisotropic velocity tomography: A case study in a near-surface rock mass, Geophysics 58 (1993) 1748–1763
41. D. J. White: Two-dimensional seismic refraction tomography, Geophys. J. 97 (1989) 223–245
42. W. J. Lutter, R. L. Nowack and L. W. Braile: Seismic imaging of the upper crustal velocity structure using travel times from the PASSCAL Ouachita experiment, J. Geophys. Res. 95 (1990) 4621–4631
43. C. A. Zelt and R. B. Smith: Seismic travel time inversion for 2D crustal velocity structure, Geophys. J. Int. 108 (1992) 16–34
44. W. J. Lutter and R. L. Nowack: Inversion for crustal structure using reflections from the PASSCAL Ouachita experiment, J. Geophys. Res. 95 (1990) 4633–4646
45. J. A. Hole: Nonlinear high-resolution three-dimensional travel-time tomography, J. Geophys. Res. 97 (1992) 6553–6562
46. C. A. Zelt and P. J. Barton: Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin, J. Geophys. Res. 103 (1998) 7187–7210
47. R. L. Nowack and L. W. Braile: Refraction and wide-angle reflection tomography: theory and results, In: Seismic Tomography: Theory and Practice ed. by H. M. Iyer and K. Hirahara, (Chapman and Hall, London 1993)
48. T. N. Bishop, K. P. Bube, R. T. Cutler, R. T. Langan, P. L. Love, J. R. Resnick, R. T. Shuey, D. A. Spindler and H. W. Wyld: Tomographic determination of velocity and depth in laterally varying media, Geophysics 50 (1985) 903–923
49. V. Farra and R. Madariaga: Non-linear reflection tomography, Geophys. J. 95 (1988) 135–147
50. P. R. Williamson: Tomographic inversion in reflection seismology, Geophys. J. Int. 100 (1990) 255–274

51. C. Stork: Singular value decomposition of the velocity-reflector depth tradeoff, part 2: High-resolution analysis of a generic model, Geophysics 57 (1992) 933–943
52. F. Billette and G. Lambare: Velocity macro-model estimation by stereotomography, Geophys. J. Int. 135 (1998) 671–680
53. G. Lambare: Stereotomography: where we are, and where should we go, SEG extended abstracts for the 74th annual meeting (2004) 2367–2370
54. P. Gasparini and the TomoVes Working Group: Looking inside Mt Vesuvius, Trans. Am. Geophys. Un. (EAS) 79 (1998) 229–231
55. R. L. Nowack, C. Li and J. Virieux: Inversion of seismic attributes from the 1996 3-D tomography experiment at Mt. Vesuvius, Italy, Trans. Am. Geophys. Un. (EOS) 81 (2000) 903
56. R. L. Nowack, C. Li and J. Virieux: Inversion of seismic attributes from the 1996 3-D tomography experiment at Mt. Vesuvius, Italy, In: The TomoVes Project, Looking Inside Mt. Vesuvius eds. A. Zollo, A. Bobbio, P. Gasparini, R. Casale and M. Yeroyanni, (CD Book, CUEN, Naples, Italy 2003)
57. A. Zollo, L. D'Auria, R. De Matteis, A. Herrero, J. Virieux and P. Gasparini: Bayesian estimation of 2-D P-velocity models from the active seismic arrival time data: imaging of the shallow structure of Mt. Vesuvius (Southern Italy), Geophys. J. Int. 151 (2002) 566–582
58. C. H. Thurber: Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, J. Geophys. Res. 88 (1983) 8225–8236.
59. E. Kissling: Geotomography with local earthquake data, Rev. of Geophys. 26 (1988) 659–698
60. D. Eberhart-Phillips: Three-dimensional P- and S-velocity structure in the Coalinga region, California, J. Geophys Res. 95 (1990) 15, 343–14363
61. K. Hirahara: Tomography using both local earthquakes and teleseisms: velocity and anisotropy – theory, In Seismic Tomography: Theory and Practice Ed. by H. M. Iyer and K. Hirahara (Chapman and Hall, London 1993)
62. H. Zhang and C. H. Thurber: Double-difference tomography: The method and its application to the Hayward Fault, California, Bull. Seism. Soc. Am. 93 (2003) 1875–1889
63. R. W. Clayton and R. P. Comer: A tomographic analysis of mantle heterogeneities from body wave travel time, Trans. Am. Geophys. Un. (EOS) 64 (1983) 776
64. A. M. Dziewonski and D. L. Anderson: Mapping the lower mantle: determination of lateral heterogeneity in P velocity up to degree and order 6, J. Geophys. Res. 89 (1984) 5929–5952
65. J. H. Woodhouse and A. M. Dziewonski: Mapping of the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms, J. Geophys. Res. 89 (1984) 5953–5986
66. B. L. N. Kennett and R. Van der Hilst: Seismic structure of the mantle: from subduction zone to craton, In The Earth's Mantle Ed. by I. Jackson (Cambridge Univ. Press, 1998)

67. B. Romanowicz: Global mantle tomography: progress status in the past 10 years, Ann. Rev. Earth Planet. Sci. 31 (2003) 303–328
68. A. Curtis and R. Snieder: Probing the Earth's interior with seismic tomography, In: International Handbook of Earthquake and Engineering Seismology (Academic Press, San Diego 2002)
69. M. Cara: Seismic Anisotropy, In: International Handbook of Earthquake and Engineering Seismology (Academic Press, San Diego 2002) 875–885
70. C. J. Thomson: Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR, Geophys. J. R. Astr. Soc. 74 (1983) 525–558
71. F. Neele, J. VanDecar and R. Snieder: The use of P wave amplitude data in a joint inversion with travel times for upper mantle velocity structure, J. Geophys. Res. 98 (1993) 12,033–12,482
72. N Bregman, C. H. Chapman and R. C. Bailey: Travel time and amplitude analysis in seismic tomography, J. Geophys. Res. 94 (1989) 7577–7587
73. R. L. Nowack and M. P. Matheney: Inversion of seismic attributes for velocity and attenuation structure, Geophys. J. Int. 128 (1997) 689–700
74. B. J. Mitchell and B. Romanowicz (Eds.): Q of the Earth: Global, Regional and Laboratory Studies, (Birhhauser Boston 1999)
75. Y. Wang: Seismic Amplitude Inversion in Reflection Tomography, (Pergamon, Amsterdam 2003)
76. A. J. Devaney: Geophysical diffraction tomography, IEEE Trans. Geos. GE-22 (1984) 3–13
77. R. S. Wu and M. N. Toksoz: Diffraction tomography and multi-source holography applied to seismic imaging, Geophysics 52 (1987) 11–25
78. R. G. Pratt and M. H. Worthington: Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method, Geophys. Prosp. 38 (1990) 287–310
79. R. G. Pratt, Z. M. Song, P. Williamson and M. Warner: Two-dimensional velocity models from wide-angle seismic data by wavefield inversion, Geophys. J. Int. 124 (1996) 323–340
80. A. Brenders and R. G. Pratt: Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. Geophys. J. Int. 168 (2007) 133–151.
81. M. J. Woodward: Wave equation tomography, Geophysics 57 (1992) 15–26
82. K. Yomogida: Fresnel zone inversion for lateral heterogeneities in the Earth, Pure Appl. Geophys. 138 (1992) 391–406
83. F. A. Dahlen, S. H. Hung and G. Nolet: Frechet kernels for finite frequency traveltimes – I. Theory, Geophys. J. Int. 141 (2000) 157–174
84. Montelli, G. Nolet, G. Masters, F. A. Dahlen and S. H. Hung: Global P and PP traveltime tomography: rays versus rays, Geophys. J. Int. 158 (2004) 637–654

85. van der Hilst, R. D. and M. V. de Hoop: Banana-doughnut kernels and mantle tomography, Geophys. J. Int. 163 (2005) 956–961
86. Montelli, R. G. Nolet and F. A. Dahlen: Comment on 'Banana-doughnut kernels and mantle tomography' by van der Hilst and de Hoop, Geophys. J. Int. 167 (2006) 1204–1210.
87. R. van der Hilst, and M. de Hoop: Reply to comment by R. Montelli, G. Nolet and F. A. Dahlen on 'Banana-doughnut kernels and mantle tomography', Geophys. J, Int. 167 (2006) 1211–1214.

92

Array Signal Processing

Christine Thomas

Department of Earth and Ocean Sciences, University of Liverpool, Liverpool, UK

1 • Introduction

In the late 1950s and early 1960s, regional seismic arrays came into existence. Initially, applications were in identification of nuclear tests; however, seismologists quickly realised the importance of a number of linked stations in well-defined configurations for fundamental studies of the Earth's interior. The main advantages of array seismology are the possibility to obtain directional information about detected signals and to increase the signal-to-noise ratio (SNR) through directional filtering. Directional information is important in determining the origin of a wave, distinguishing between waves of different origins and determining the seismic phase of an arrival, i.e. is it a direct wave (e.g. P wave) or a reflection off a layer inside the Earth's mantle or a wave that has travelled through the Earth's core? (e.g. Weber 1993)

2 • Slowness, Backazimuth and the Complete Slowness Vector

Incoming plane waves can be identified by their angle of incidence i and backazimuth Φ (Figure 1). The angle of incidence is measured from vertical, and backazimuth is measured clockwise from North to the direction from which the wave arrives. The angle of incidence i can be translated into the

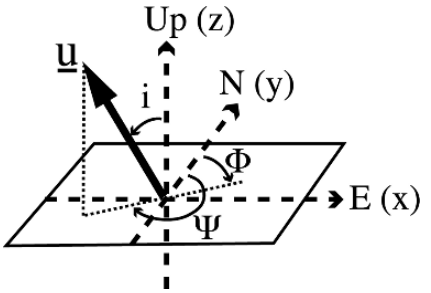


FIGURE 1 The angle of incidence i and backazimuth Φ of an incoming wave. Propagating in the direction of the slowness vector. The backazimuth is 180° opposite to the azimuth Ψ of the slowness vector and is nominally the direction from the station back to the event.

horizontal slowness, $u = (\sin i)/v_0$, where v_0 is the velocity beneath the array. The horizontal slowness of a wave does not change along the ray path in horizontally stratified media. In the following, the word slowness is used for the horizontal slowness. The complete slowness vector, which points in the direction of wave propagation, is the combination of horizontal slowness and backazimuth or azimuth ($\Psi = \Phi - 180°$):

$$\underline{u} = \frac{1}{v_0}(\sin i \sin \Psi, \sin i \cos \Psi, \cos i) \tag{1}$$

in the (East, North, Up) coordinate system. If a signal $s(t)$ is received by a station at (0,0,0), then a station at location $\underline{r} = (E, N, Z)$ receives the delayed (or advanced) signal $s(t - \underline{u} \cdot \underline{r})$.

3 • Array Methods

In the following, several methods are introduced to study seismic waves, starting from the basic beamforming methods to migration methods, together with their application, their advantages and disadvantages. A more detailed review of array methods can be found in Rost and Thomas (2002).

3.1 Beamforming

If a seismic wave arrives at the array, the arrival times at each station will differ due to the slowness and backazimuth of the incoming wave and the distances between the stations and the centre of the array. The time differences are called the delay times, which can be used to calculate the direction of the wave. In beamforming, the recorded traces are shifted with the negative delay times calculated for a specific slowness and backazimuth and then summed. After shifting and summing, spatially coherent waves that arrive with the chosen slowness and backazimuth value will stack constructively, whereas all other arrivals will diminish in amplitude.

Consider an array with M seismic stations, where the seismic trace from the ith station, $x_i(t)$, is composed of a delayed signal $s(t-d_i)$ plus noise $n_i(t)$. If the signal is a plane wave with a slowness vector $\underline{u}$, the arrival times of the signal will be delayed with respect to the centre of the array by the time difference, $d_i = \underline{r_i} \cdot \underline{u}$, where $\underline{r_i}$ is the position vector to the ith station from the centre of the array. The summed and beamformed trace is

$$b(t) = \frac{1}{M} \sum_{i=1}^{M} x_i (t + d_i) . \tag{2}$$

After shifting the traces by the negative delay times, the signal $s(t)$ is aligned, and when summing the traces, the noise will sum incoherently. Assuming the signals are perfectly correlated, the beam (delay-and-sum) trace is

$$b(t) = s(t) + \frac{1}{M} \sum_{i=1}^{M} n_i (t + d_i) . \tag{3}$$

The SNR of the array is 10 log M dB greater than the SNR of a single station if the noise is uncorrelated and of equal variance and the signal is perfectly correlated across the array (Harjes and Henger, 1973).

In using beamforming, one has to make several assumptions: Firstly, the wave has to be spatially coherent across the array; secondly, the wave has to be a plane wave and thirdly, the slowness and backazimuth of the arriving wave have to be known. In addition, to improve the SNR, the noise should be uncorrelated between traces. In many cases, these are valid assumptions; however, the spatial coherence of a wave is only given if the array aperture is sufficiently small and the subsurface is sufficiently homogeneous. The plane wave assumption is appropriate for far-field signals but not for receiver-side-produced signals such as scattering near the array. In addition, in exploration seismology, the plane wavefield assumption does not hold where there is substantial wavefront curvature along the length of the array. In this case, a more complicated expression for the delay times d_i may be required to apply to the above equation. In cases where the noise field is directional (or otherwise correlated between sensors), weightings can be applied to the traces before summing to improve the SNR (Johnson and Dudgeon, 1993).

Sometimes, the slowness and backazimuth of standard waves such as P, S, PcP, PKP, etc. can be assumed if the source location is known, but there are many instances where the slowness and/or the backazimuth of a wave may not be known. Waves may arrive

out of plane, i.e. not with the theoretical backazimuth, due to structure on their path. Local events may not have the same backazimuth and slowness from all stations. A further complication is that the structure beneath the array can influence the recorded backazimuth and slowness values. Deviations from the theoretical values can be measured for standard phases and displayed as mislocation vectors (e.g. Krüger and Weber, 1992). These in turn can then be used as an array calibration to correct the slowness and backazimuth for observed non-standard phases (Faber et al., 1986).

3.2 Determining Slowness or Backazimuth

Vespagrams (**ve**locity **sp**ectral **a**nalysis; Davies et al., 1971), also called slant stacks, are a display of the beamformed seismic trace for a range of slowness or backazimuth values. They are used to determine the slowness or backazimuth of an arriving wave (Figure 2). The vespagram can be written as

$$v_{\zeta}(t) = \frac{1}{M}\sum_{i=1}^{M} x_i(t + d_{i,\zeta}), \tag{4}$$

where ζ is either backazimuth or slowness.

When performing the vespa routine, the fixed parameter has to be known exactly due to the interdependence of slowness and backazimuth in the complete slowness vector. Deviations of the true backazimuth from the estimated value will cause a deviation of the apparent slowness (Rost and Thomas, 2002). Other disadvantages are similar to those mentioned for beamforming in that

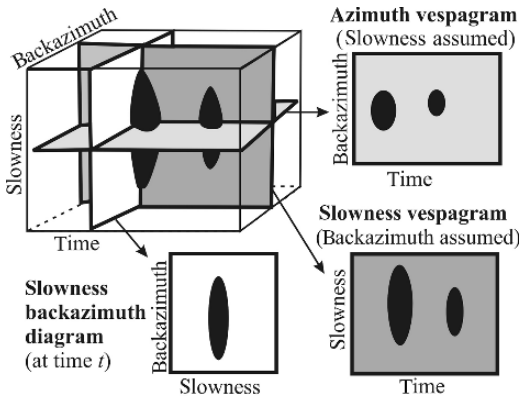


FIGURE 2 The arriving wavefield with slowness and backazimuth over time can be analysed in a vespagram (either azimuth or slowness vespagram, by assuming the other parameter) or for one specific time window simultaneously for slowness and backazimuth (slowness–backazimuth diagram).

the arriving energy needs to be a plane wave and the noise has to be uncorrelated.

3.3 Simultaneous Determination of Slowness and Azimuth

When slowness and backazimuth must both be determined (e.g. when studying deviations of waves from the theoretical slowness vector), several other processes have been developed. The most important ones are beaman (beam power analysis, King et al. (1976), also called slowness–backazimuth diagram) and frequency wavenumber (FK) analysis. In the slowness–backazimuth diagram (Figure 2), an array beam trace is formed for each slowness and backazimuth pair over a grid of values, and the energy in each beam trace is computed within a specific time window around the phase of interest. This time-domain procedure offers good time resolution but can take a lot of computer power, so FK analysis is often preferred. In FK analysis, the wavefield within one time window is transformed to frequency and wavenumber domain (Capon, 1973; Harjes and Henger, 1973). The FK spectrum provides amplitudes and phases as a function of frequency and wavenumber, $\underline{k} = (k_x, k_y)$, and may be computed as

$$X(\underline{k}, f) = \frac{1}{M}\sum_{i=1}^{M} X_i(f)\exp(j\underline{k} \bullet \underline{r}_i), \tag{5}$$

where $X_i(f)$ is the Fourier transform of the trace received at the ith station. The exponential term applies the negative delay time to the trace in frequency domain. If the received traces are composed of a signal $S(f)$ with additive noise $N_i(f)$ and with delay times based on the wavenumber vector $\underline{k}_0$ at station location $\underline{r}_i$,

$$X_i(f) = S(f)\exp(-j\underline{k}_0 \bullet \underline{r}_i) + N_i(f). \tag{6}$$

FK diagrams display the energy (integrated over frequency) at each wavenumber. If the noise at each trace has the same expected power spectral density $<|N(f)|^2>$ and is uncorrelated between traces, the expected value of the integrated energy of the beam trace at wavenumber $\underline{k}$ for an incoming plane wave at wavenumber $\underline{k}_0$ is (Kelly, 1967; Harjes and Henger, 1973; Rost and Thomas, 2002)

$$\langle E(\underline{k} - \underline{k}_0)\rangle = \int_{-\infty}^{\infty} |S(f)|^2 |A(k - k_0)|^2 + \frac{\left\langle |N(f)|^2 \right\rangle}{M} \mathrm{d}f, \tag{7}$$

where $A(k - k_0)$ is the array response function (ARF),

$$|A(\underline{k} - \underline{k}_0)|^2 = \left| \frac{1}{N}\sum_{j=1}^{N} e^{2\pi i \underline{r}_j (\underline{k} - \underline{k}_0)} \right|^2. \tag{8}$$

Equation (7) shows that the FK diagram displays the ARF of the array when the trace is dominated by a coherent wavefield (see Figure 3) and displays an approximately constant background when the trace is dominated by incoherent noise.

Wavenumber is related to slowness u and backazimuth Φ,

$$|\underline{k}| = \sqrt{(k_x^2 - k_y^2)} = 2\pi f u, \quad \Phi = \tan^{-1}\left(\frac{k_x}{k_y}\right), \tag{9}$$

where f is frequency. Based on this relationship, a grid giving slowness and backazimuth is often overplotted on an FK diagram,

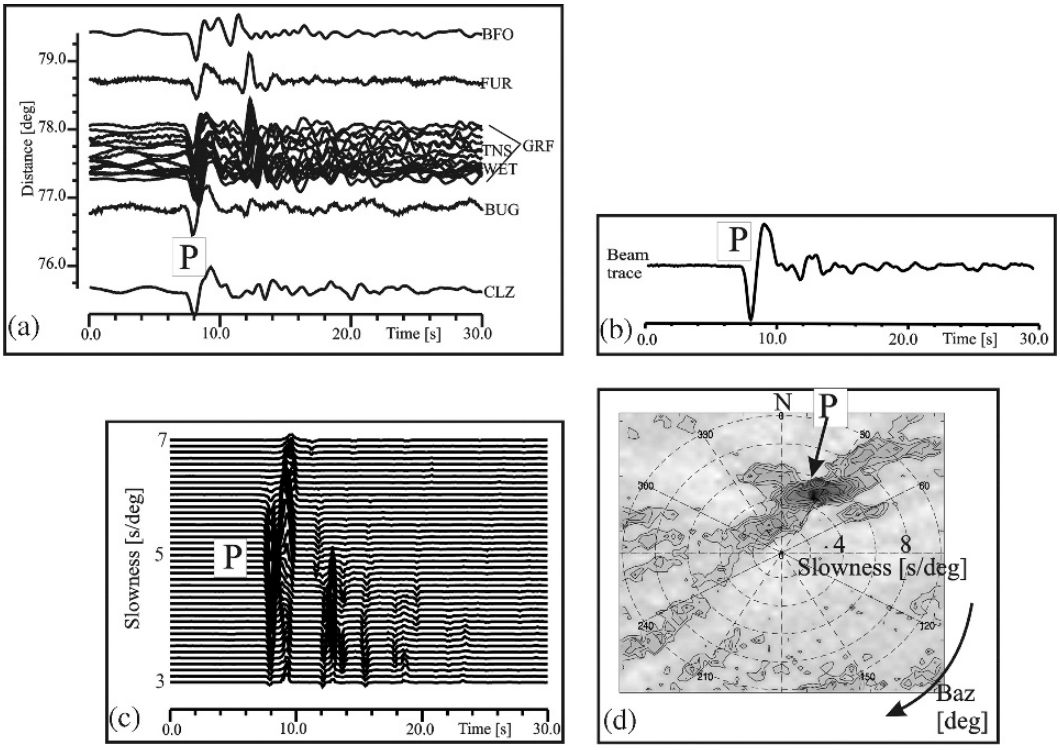


FIGURE 3 **(a)** Seismic record section aligned on P. A second wave arrival is visible with a different move-out compared to P. **(b)** Beamforming using the theoretical P values. The P-wave is amplified; the second arrival (5 s after P) has a reduced amplitude. **(c)** Vespagram for a slowness range of 3–7 s/deg. Both P and the second arrival with a lower slowness than P are amplified. **(d)** FK analysis of a window around the P-wave for GRF (Gräfenberg array) stations only. The contour maximum gives the slowness and backazimuth value of the P-wave.

where slowness is displayed in the radial direction and the backazimuth clockwise from the top (North) (Figure 3).

One limitation of the simultaneous determination of slowness and backazimuth is that it can best be performed for small time windows. If the time window is chosen too large, several phases can overlap and the results could become misleading. To overcome this limitation, the sliding window FK analysis has been introduced by Rost and Weber (2001), where short, overlapping time windows are shifted across the trace and are analysed in the same way as the FK analysis. It has to be noted again that the subsurface structure can influence the results in the same way as explained for the beamforming, and the interpretation has to be conducted with care (e.g., the event may not lie on the measured backazimuth). Figure 3 shows examples for beamforming, vespagram and FK analysis for a Kurile event recorded at the German Regional Seismic Network (GRSN).

3.4 Other stacking methods

Several improvements to slant stacking methods have been developed. These include the source stack for which several earthquakes are stacked at one receiver (e.g. Spudich and Bostwick, 1987) and a combination of source and receiver stacks, known as the double-beam method (Krüger et al., 1993; 1996; Scherbaum et al., 1997), where the SNR is improved by the number of stations times the number of events. In both the source stack and double-beam method, great care must be taken to correct for the source function and the relative location of the events (Krüger et al., 1996). The shortcomings of the source stack and double-beam methods are that usually locations of teleseismic events are not exactly known as well as the problems mentioned above for the receiver stacks.

Another development in the vespa routine is a non-linear stacking, also known as Nth root stacking (Kanasewich et al., 1973; Muirhead and Datt, 1976). This non-linear stacking of the shifted traces reduces the noise more efficiently but distorts the waveform of the arrivals (McFadden et al., 1986) so it is unusable for waveform studies. The Nth root process sums the Nth roots (e.g. $N = 2$ (square root), $3, 4, \ldots$) of the trace magnitudes, with the original signs applied,

$$v'_{\zeta,N}(t) = \frac{1}{M}\sum_{i=1}^{M} \left|x_i(t+d_{\zeta,i})\right|^{1/N} \operatorname{sgn}\left(x_i(t+d_{\zeta,i})\right). \qquad (10)$$

This stacked trace is then taken to the Nth power and its original sign is applied,

$$v_{\zeta,N}(t) = \left|v'_{\zeta,N}(t)\right|^{N} \operatorname{sgn}\left(v'_{\zeta,N}(t)\right), \qquad (11)$$

where ζ can again be slowness or azimuth. During the process, the signs of the samples are kept and therefore the polarity of the waves is preserved. In the Nth root stacking, the coherency of a wave is more important than the amplitude of the wave at the stations, and therefore, small coherent signals can be extracted more efficiently than in a linear stack.

3.5 Migration

Up to this point, the array methods assumed a plane wave signal, which is valid for many teleseismic arrivals but may be invalid for receiver-side-generated waves. These non-plane wave arrivals can be dealt with by the migration method that is applied in exploration geophysics to collapse diffractions back to their origin or to move dipping reflectors into their true position.

Many different migration techniques have been introduced in teleseismic wave analysis (e.g. Lynnes and Lay, 1989; Revenaugh, 1995; Scherbaum et al., 1997; Thomas et al., 1999; Bostock and Rondenay, 1999). In one technique (e.g. Thomas et al., 1999), a 2- or 3-dimensional grid of the subsurface is created. Travel times from the source to each grid point and from there to the receivers are calculated. The traces are shifted according to the calculated travel time and stacked. These stacked amplitudes are added to their corresponding grid cells. With enough arrivals and enough receivers, this results in a map of scattering amplitudes in the subsurface. In another type of migration (e.g. Lay and Young, 1996), scattering ellipsoids are calculated for the time of each coda arrival. These ellipsoids consist of all possible scatterer locations that could produce an arrival at the specified time. The number of instances ("hits") when each grid location is included in a scattering ellipsoid is logged, and the likelihood that a particular location contains a scatterer is assessed based on the total number of hits compared to other locations. Neal and Pavlis (1999, 2001) describe another variation on migration utilising receiver functions.

Accurate migration depends on having the correct velocity model inside the Earth, which can be a problem for teleseismic or deep Earth applications. If the velocities are assumed wrongly, the travel times will be wrong and the scattered waves will not stack constructively at the points where they originate. Another problem is that in most migration methods, only single scattering is assumed, which leads to mispositioning the origins of multiply scattered waves. It is noteworthy that for steeply arriving waves (small slowness), the spatial resolution is good, whereas the depth resolution of the scattering point is poor. In the case of more shallow arriving waves (larger slowness), the opposite is true.

4 • Array Configurations

The configuration of arrays depends on the purpose of the array (e.g. targeting a certain region within the Earth or on the surface) and sometimes on the geology beneath the array. The array response function (ARF) of Equation (9), also called the wavenumber response, describes the directional sensitivity of the array. The aperture of the array, the station spacing, the location of the stations and the frequency of the waves influence the ARF. The ARF generally has a main lobe associated with the direction of the incoming wave and side lobes. The main lobe should be sharp and distinct, the side lobes far away from the main lobe and small (Harjes and Henger, 1973). The array response for a symmetric cross array, for example, would produce a main lobe with symmetric side lobes in cross shape. A circular array on the other hand would produce a strong main lobe with circular side lobes at some radial distance from the main lobe. The ARF of the GRSN (Germany) can be seen in Figure 3d, shifted with the slowness and backazimuth values of the P-wave. Side lobes are visible around the main lobe that indicates the values for the P-wave.

In general, it is useful to build arrays that are 2-dimensional (i.e. not a line) since either the slowness or the backazimuth resolution of a line array will be poor for certain directions. For example, if seismic waves arrive perpendicular to a line array, the backazimuth resolution will be good but the slowness resolution will be poor and will result in a large elongated main lobe in the ARF. The array aperture should also increase with the bandwidth of the seismic signals that are likely to be used (larger aperture for longer periods). The ideal array would have an ARF close to a delta function with the side lobes outside the wavenumber window of interest, which would then allow measuring the slowness and backazimuth of an arriving wave perfectly. This kind of ARF can be approximated by a circular array with irregular station spacing (Harjes and Henger, 1973).

Acknowledgments I thank Max Deffenbaugh and an anonymous reviewer for their helpful comments.

References

Capon, J., Signal processing and frequency-wavenumber spectrum analysis for a large aperture seismic array, in: Methods in computational physics, 2–59, 1973.

Davies, D., E.J. Kelly and J.R. Filson, Vespa process for analysis of seismic signals, Nature Phys. Sci., 232, 8–13, 1971.

Faber, S., J. Plomerova and V. Babuska. Deep-seated lateral velocity variations beneath the GRF array inferred from mislocation patterns and P-residuals, Z. Geophys, 60, 139–148, 1986.

Harjes, H.-P. and M. Henger, Array-Seismologie, Zeitschr. f. Geophysik, 39, 865–905, 1973.

Johnson, D.H. and D.E. Dudgeon, Array Signal Processing: Concepts and Techniques, Prentice-Hall: Englewood Cliffs, NJ, 1993.

Kanasewich, E.R., C.D. Hemmings and T. Alpaslan, Nth-root stack nonlinear multichannel filter, Geophysics, 38, 327–338, 1973.

Kelly, E.J., Response of seismic signals to wide-band signals, Linc. Lab. Techn. Note, 1967.

King, D.W., E.S. Husebye and R.A.W. Haddon, Processing of seismic precursor data, Phys. Earth Planet. Int., 12, 128–134, 1976.

Krüger, F., M. Weber, F. Scherbaum and J. Schlittenhardt, Double beam analysis of anomalies in the core-mantle boundary region, Geophys. Res. Lett., 20, 1475–1478, 1993.

Krüger, F., F. Scherbaum, M. Weber and J. Schlittenhardt, Analysis of asymmetric multipathing with a generalisation of the double-beam method, Bull. Seism. Soc. Am., 86, 737–749, 1996.

Krüger, F. and M. Weber, The effect of low velocity sediments on the mislocation vectors of the GRF array, Geophys. J. Int., 108, 387–393, 1992.

Lay, T. and C.J. Young, Imaging scattering structures in the lower mantle by migration of long-period S waves, J. Geophys. Res., 101, 20023–20040, 1996.

Lynnes, C.S. and T. Lay, Inversion of P coda for isotropic scatterers at the Yucca Flat test site, Bull. Seism. Soc. Am., 79, 790–804, 1989.

McFadden, P.L., B.J. Drummond and S. Kravis, The Nth-root stack: Theory, applications, and examples, Geophysics, 51, 1879–1892, 1986.

Muirhead, K.J. and R. Datt, The N-th root process applied to seismic array data, Geophys. J. R. Astr. Soc., 47, 197–210, 1976.

Neal, S.L. and G.L. Pavlis, Imaging P- to S conversions with multichannel receiver functions, Geophys. Res. Lett., 26, 2581–2584, 1999.

Neal, S.L. and G.L. Pavlis, Imaging P- to S conversions with broad-band seismic arrays using multichannel time-domain deconvolution, Geophys. J. Int., 146, 57–67, 2001.

Revenaugh, J., The contribution of topographic scattering to teleseismic coda in Southern California. Geophys. Res. Lett., 22, 543–546, 1995.

Rost, S. and C. Thomas, Array Seismology: Methods and Applications, Rev Geophys., doi 10.1029/2000RG000100, 2002.

Rost, S. and M. Weber, A reflector at 200 km depth beneath the NW Pacific, Geophys. J. Int., 147, 12–28, 2001.

Scherbaum, F., F. Krüeger and M. Weber, Double beam imaging: Mapping lower mantle heterogeneities using combinations of source and receiver arrays, J. Geophys. Res., 102, 507–522, 1997.

Spudich, P. and T. Bostwick, Studies of the seismic coda using an earthquake cluster as a deeply buried seismograph array, J. Geophys. Res., 92, 10526–10546, 1987.

Thomas, Ch., M. Weber, Ch. Wicks and F. Scherbaum, Small scatterers in the lower mantle observed at German broadband arrays. J. Geophys. Res., 104, 15073–15088, 1999.

Weber, M.H., P- and S-wave reflections from anomalies in the lowermost mantle, Geophys. J. Int., 115, 183–210, 1993.

PART XVI
UNDERWATER AND OCEANOGRAPHIC ACOUSTICS

Edmund Sullivan[1] and Michael Taroudakis[2]

[1] EJS Consultants, Portsmouth, RI, USA
[2] University of Crete, Heraklion Crete, Greece

93 Propagation and Signal Modeling

Finn B. Jensen 1669

Introduction • The Ocean Acoustic Environment • Numerical Modeling Techniques • Time-Harmonic Solutions of Separable Problems • Time-Harmonic Solutions of Non-Separable Problems • Pulse Solutions by Fourier Synthesis • Signal Simulation Examples • Summary

94 Statistical Signal Processing

Edmund J. Sullivan 1695

Background • Detection • Estimation • Spatial Processing and Arrays

95 Inverse Problems in Underwater Acoustics

N. Ross Chapman 1723

Introduction • Linear Inverse Problems • Linearized Inversion • Nonlinear Inverse Problems • Matched Field Inversion

96 Active Sonar

Edmund J. Sullivan 1737

Introduction • The Sonar Signal • Sonar System Design • Performance Analysis • Discussion

97 Passive Sonar

Brian H. Maranda 1757

Introduction • Passive Sonar Systems • The Passive Sonar Equation • Spatial Processing and Array Gain • Detectors • Classification • Localization

98 Transducers for Sonar Systems

Kim Benjamin 1783

Transduction Materials • Conventional Designs and Modeling Approaches • Broad Bandwidth Transducer Designs • Considerations for Remotely Operated Undersea Vehicles

99 Flow Noise

Gerald C. Lauchle 1821

Introduction • Flow-Induced Noise in Hull-Mounted Sonar Systems • Transition Zone Noise Scaling • Turbulent Boundary Layer Wall Pressure Fluctuations • Transducer Size and Array Effects

93

Propagation and Signal Modeling

Finn B. Jensen

NATO Undersea Research Centre, La Spezia, Italy
email: jensen@nurc.nato.int

1 • Introduction

The use of sound in the sea is ubiquitous: Apart from the military aspect of trying to detect an adversary's mines and submarines, ship-mounted sonars measure water depth, ship speed, and the presence of fish shoals. Side-scan systems are used for mapping the bottom topography, sub-bottom profilers for getting information about the deeper layering, and other sonar systems for locating pipelines and cables on the seafloor. Sound is also used for navigating submerged vehicles, for underwater communications and for tracking marine mammals. Finally, in the realm of 'acoustical oceanography' and 'ocean acoustic tomography,' sound is used for measuring physical parameters of the ocean environment and for monitoring oceanic processes [1–6].

The range of useful frequencies in ocean acoustics covers almost six decades of frequency—from a few hertz to several hundreds of kilohertz—with wavelengths from around 1 km down to a few millimeters. Since attenuation of sound in seawater increases approximately with frequency squared, the high-frequency devices (>10 kHz) are used primarily in short-range high-resolution systems such as bottom profilers and fish sonars. The lower frequencies propagate to long ranges with little attenuation and are therefore preferred in acoustic surveillance

systems for military use as well as for acoustic sensing on a global scale, e.g., to monitor ocean warming.

2 • The Ocean Acoustic Environment

The ocean is an acoustic waveguide limited above by the sea surface and below by the seafloor. The speed of sound in the waveguide plays the same role as the index of refraction does in optics. Sound speed is normally related to density and compressibility. In the ocean, density is related to static pressure, salinity, and temperature. The sound speed in the ocean is an increasing function of temperature, salinity, and pressure, the latter being a function of depth. It is customary to express sound speed (c) as an empirical function of three independent variables: temperature (T) in degrees centigrade, salinity (S) in parts per thousand, and depth (D) in meters. A simplified expression for this dependence is

$$c = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.010T)(S - 35) + 0.016D. \quad (1)$$

In shallow water, where the depth effect on sound speed is small, the primary contributor to sound-speed variations is the temperature. Thus, for a salinity of 35 ppt, the sound speed in seawater varies between 1450 m/s at 0°C and 1545 m/s at 30°C.

To illustrate the complexity of the modeling problem associated with broadband signal simulation in the ocean, let us briefly review the acoustics of the ocean waveguide. Figure 1 is a schematic of the basic types of propagation in the ocean resulting

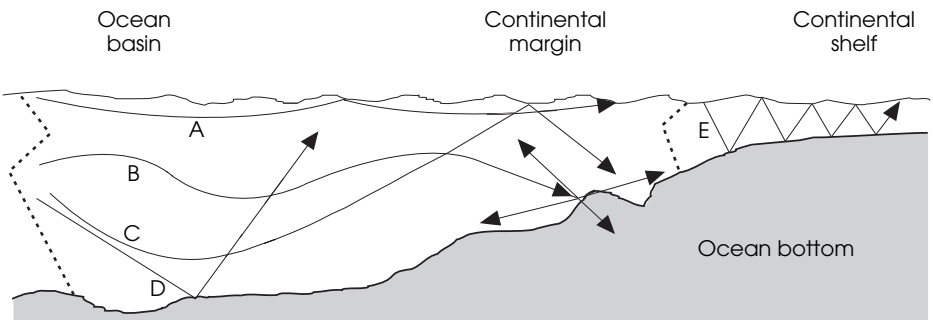


FIGURE 1 Schematic of sound propagation in the ocean.

from the sound-speed profiles (indicated by the dashed lines) discussed above. These sound paths can be understood from *Snell's law*,

$$\frac{\cos\theta}{c} = \text{const}, \tag{2}$$

which relates the ray angle, with respect to the horizontal, to the local sound speed. It is not difficult to see that the implication of this law is that sound bends locally toward regions of low sound speed (or sound is "trapped" in regions of low sound speed).

Simplest to explain are the paths about local sound-speed minima: A and B. Path A corresponds to surface duct propagation where the minimum sound speed is at the ocean surface. Path B, depicted by a ray leaving a deeper source at a shallow horizontal angle, propagates in the *deep sound channel* whose axis is at the shown sound speed minimum. For mid-latitudes, sound in the deep channel can propagate long distances without interacting with lossy boundaries; low-frequency propagation via this path has been observed to distances of thousands of kilometers. Path C, which is at slightly steeper angles than those associated with path B, is *convergence zone* propagation, a spatially periodic ($\sim$ 35–70 km) refocusing phenomenon producing zones of high intensity near the surface because of the upward refracting nature of the deep sound-speed profile. Here sound interacts with the surface but not with the bottom.

The bottom bounce path D is also a periodic phenomenon but with a shorter cycle distance and a shorter total propagation distance because of losses when sound is reflected from the ocean bottom. Finally, the right-hand-side of Figure 1 depicts propagation in a shallow water region such as a continental shelf. Here sound is channeled in a waveguide bounded above by the ocean surface and below by the ocean bottom.

The modeling of sound propagation in the ocean is complicated because the environment varies laterally—is *range dependent*—and all environmental effects on sound propagation are dependent on acoustic frequency in a rather complicated way. As a result the ray-type schematic of Figure 1 is misleading, particularly at low frequencies.

2.1 Environmental Variability

The usefulness of acoustic models as an analysis and prediction tool depends entirely on the general predictability of ocean acoustic propagation. Of course, the ocean sound-speed structure is frozen in neither time nor space. On the contrary, the ocean has its own weather system. There are currents, internal waves, and quite a bit of microstructure present, particularly in coastal

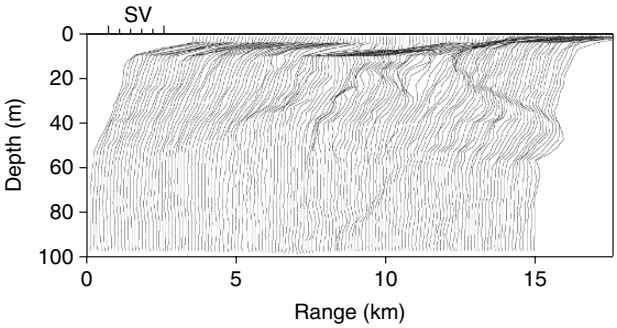


FIGURE 2 Spatial variability of sound speed in shallow-water area of the Mediterranean.

water areas. Figure 2 illustrates the sound-speed variability along a 15-km long track in the Mediterranean Sea. The data were recorded on a towed thermistor chain covering depths between 5 and 90 m. In general, this type of time-varying oceanographic structure has an effect on sound propagation, as a source of both attenuation and acoustic fluctuations.

Unfortunately, detailed measurements of water column properties as a function of range and time are generally not available. This lack of information about the real ocean environment clearly affects the reliability of acoustic propagation predictions. However, lower source frequencies are preferable in order to minimize boundary and volume scattering effects.

Turning to the upper and lower boundaries of the ocean waveguide, the sea surface is a simple horizontal boundary and a nearly perfect reflector. The seafloor, on the other hand, is a lossy boundary with varying topography. Both boundaries have small-scale roughness associated with them which causes scattering and attenuation of sound. In terms of propagation physics, the seafloor is the most complex boundary, exhibiting vastly different reflectivity characteristics in different geographical locations.

The structure of the ocean bottom in shallow water generally consists of a thin stratification of sediments overlying the continental crust. The nature of the stratification is dependent on many factors, including geological age and local geological activity. Thus, relatively recent sediments will be characterized by plane stratification parallel to the seabed, whereas older sediments and sediments close to the crustal plate boundaries may have undergone significant deformation. An example of a complicated bottom layering is given in Figure 3, which displays

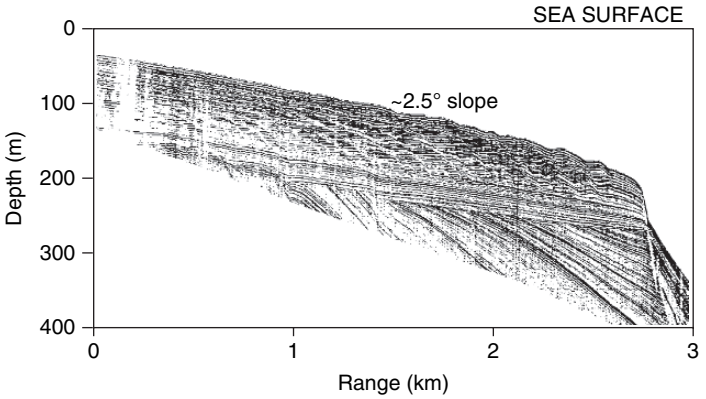


FIGURE 3 Seismic profile section from coastal-water area of the Mediterranean.

a seismic section from the Mediterranean. The track is approximately 3 km long and the water depth varies from around 40 m near the shore to 200 m at long ranges with a mean slope of approximately 2.5°. Note that while the upper stratification is almost parallel to the seafloor, deeper layers are strongly inclined.

2.2 Loss Mechanisms

Our ability to model acoustic propagation effectively in the ocean is determined by the accuracy with which acoustic loss mechanisms in the ocean environment are handled. Aside from geometrical spreading loss (spherical, cylindrical, etc.) the main loss mechanisms are volume absorption, bottom reflection loss, surface and bottom scattering loss.

Volume absorption in seawater, caused by viscosity and chemical relaxation, increases with increasing frequency. This loss mechanism is the dominant attenuation factor associated with path B in Figure 1, since this path does not interact with the boundaries. Because there is very little volume absorption at low frequencies, deep-sound-channel propagation has been observed on global scales to thousands of kilometers.

When sound interacts with the seafloor, the nature of the bottom becomes important. Reflectivity, the ratio of the amplitudes of a reflected plane wave to a plane wave incident on an interface separating two media, is an important measure of the effect of the bottom on sound propagation. Ocean bottom sediments are often modeled as fluids which means that they support only compressional waves.

The expression for reflectivity at an interface separating two homogeneous fluid media with density ρ_i and sound speed c_i, $i = 1, 2$, was first worked out by Rayleigh as

$$R(\theta_1) = \frac{(\rho_2/\rho_1)\sin\theta_1 - \sqrt{(c_1/c_2)^2 - \cos^2\theta_1}}{(\rho_2/\rho_1)\sin\theta_1 + \sqrt{(c_1/c_2)^2 - \cos^2\theta_1}}, \tag{3}$$

where θ_1 denotes the grazing angle of the incident plane wave of unit amplitude.

The reflection coefficient has unit magnitude, meaning *perfect reflection,* when the numerator and denominator of Eq. (3) are complex conjugates. This can occur only when the square root is purely imaginary, i.e., for $\cos\theta_1 > c_1/c_2$ (*total internal reflection*). The associated *critical grazing angle* below which there is perfect reflection is found to be

$$\theta_c = \arccos\left(\frac{c_1}{c_2}\right). \tag{4}$$

Note that a critical angle exists only when the sound speed of the second medium is higher than that of the first.

A closer look at Eq. (3) shows that the reflection coefficient for lossless media is real for $\theta_1 > \theta_c$, which means that there is loss ($|R| < 1$) but no phase shift associated with the reflection process. On the other hand, for $\theta_1 < \theta_c$ we have perfect reflection ($|R| = 1$) but with an angle-dependent phase shift. In the general case of lossy media (c_i complex), the reflection coefficient is complex and, consequently, there is both a loss and a phase shift associated with each reflection.

The critical-angle concept is very important for understanding the waveguide nature of ocean acoustic propagation. Figure 4 shows bottom loss curves ($\mathrm{BL} = -10\log|R|^2$) for a few simple

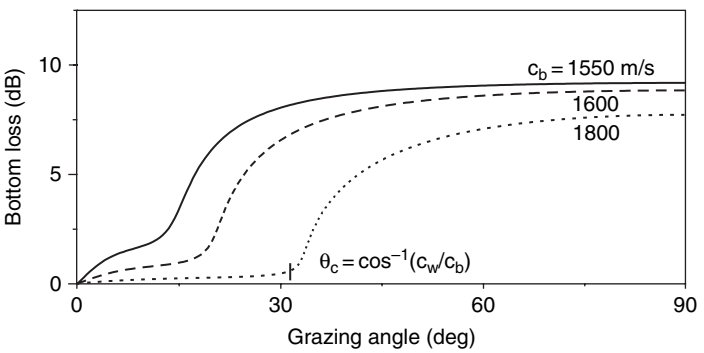


FIGURE 4 Bottom reflection loss versus grazing angle.

fluid bottoms with different compressional wave speeds, densities and attenuations. Note that for a lossy bottom we never get perfect reflection. However, there is in all cases an apparent critical angle ($\theta_c \simeq 33°$ for $c_b = 1800$ m/s in Figure 4), below which the reflection loss is much smaller than for supercritical incidence. With paths involving many bottom bounces such as in shallow-water propagation, bottom losses even as small as a few tenths of a decibel per bounce accumulate to significant total losses since the propagation path may involve many tens or even hundreds of bounces.

Real ocean bottoms are complex-layered structures of spatially varying material composition (Figure 3). A *geo-acoustic model* is defined as a model of the real seafloor with emphasis on measured, extrapolated, and predicted values of those material properties important for the modeling of sound transmission. In general, a geo-acoustic model details the true thicknesses and properties of sediment and rock layers within the seabed to a depth termed the *effective acoustic penetration depth*. Thus, at high frequencies (> 1 kHz), details of the bottom composition are required only in the upper few meters of sediment, whereas at low frequencies (< 100 Hz) information must be provided on the whole sediment column and on properties of the underlying rocks.

The information required for a complete geo-acoustic model should include the following depth-dependent material properties: the compressional wave speed c_p, the shear wave speed c_s, the compressional wave attenuation α_p, the shear wave attenuation α_s, and the density ρ. Moreover, information on the variation of all of these parameters with geographical position is required.

Finally, a rough sea surface or seafloor causes scattering of the incident sound. The result is a decay of the mean acoustic field in the water column as a function of range (scattering loss), with the scattered energy being lost to the ocean bottom through steep-angle propagation. The scattering loss increases with increasing frequency, and the propagation paths in Figure 1 mainly affected are paths A and C (surface scattering loss) and paths D and E (surface and bottom scattering loss).

A consistent mathematical model of sound propagation in the ocean must contain the physics that govern the above-mentioned propagation and loss mechanisms.

3 • Numerical Modeling Techniques

As outlined in the previous section, the ocean is an inhomogeneous, irregular waveguide limited above by a reflecting boundary (the sea surface) and below by a penetrable boundary

(the seafloor). The forward modeling problem consists in solving the acoustic wave equation with appropriate boundary and radiation conditions in order to estimate the frequency-dependent spatial and temporal properties of the sound pressure field in the ocean waveguide. This problem is generally too complex for analytical solutions, and hence we must resort to numerical methods.

The development of a numerical modeling capability in ocean acoustics has been in continuous expansion over the past 30 years [1]. Up till around 1970 the only practical technique for solving propagation problems in underwater acoustics was based on ray theory, which is a computationally efficient but approximate (infinite frequency) solution of the wave equation. The ray techniques are still in use for solving high-frequency problems in ocean acoustics.

With the advent of powerful digital computers in the early 1970s, the development of more accurate frequency-dependent solutions of the wave equation gained momentum. These wave-theory solutions (all numerically based) encompass normal mode, wavenumber integration, and parabolic equation techniques. Today, general-purpose numerical codes based on the above solution techniques are widely used for modeling acoustic propagation in complex ocean environments [6].

Even though wave-theory models provide accurate field solutions at any source frequency, the required computational effort is strongly frequency dependent. At best the calculation time increases linearly with frequency. However, for most solution techniques the calculation time increases with frequency squared, which in practice restricts the applicability of these models to low-frequency problems. In any event, with present-day computing power it is entirely feasible to apply wave-theory models to deep-water problems ($H \sim 5000\,\mathrm{m}$) for frequencies up to a few hundred hertz, and to shallow-water problems for frequencies up to a few kilohertz.

The hierarchy of computer-implemented solutions of the wave equation is shown in Figure 5. There are essentially five types of models (computer solutions to the wave equation) to describe sound propagation in the sea: wavenumber integration (WI), normal mode (NM), ray, parabolic equation (PE), and direct finite-difference (FD) or finite-element (FE) solutions of the full wave equation. All of these models permit the ocean environment to vary with depth. A model that also permits horizontal variations in the environment, i.e., sloping bottom or spatially varying oceanography, is termed *range dependent*.

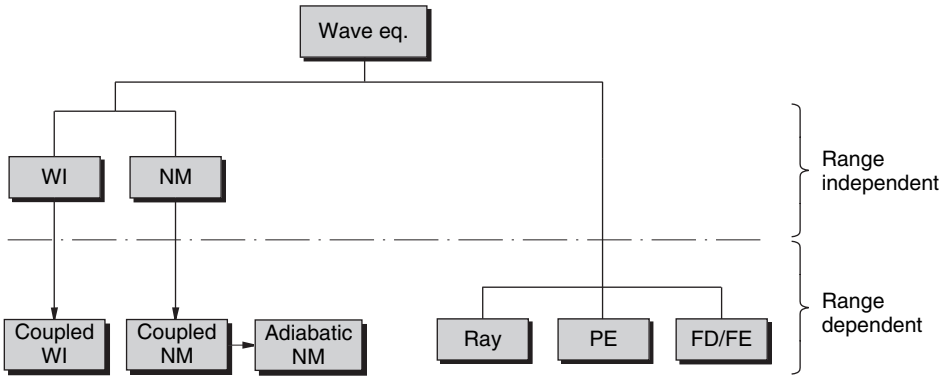


FIGURE 5 Hierarchy of numerical models in ocean acoustics. The abbreviations used are WI, wavenumber integration; NM, normal modes; PE, parabolic equation; FD, finite difference; FE, finite element.

As shown in Figure 5, an it a priori assumption about the environment being range independent leads to solutions based on spectral techniques (WI) or normal modes (NM); both of these techniques can, however, be extended to treat range dependence. Ray, PE, and FD/FE solutions are applied directly to range varying environments. For high frequencies (a few kilohertz or above), ray theory, the infinite frequency approximation, is still the most practical, whereas the other five model types become more and more applicable below, say, a kilohertz in shallow water. Models that handle ocean variability in three spatial dimensions have also been developed, but these models are used less frequently than 2-D versions because of the computational cost involved.

The starting point for all the acoustic models is the wave equation in cylindrical coordinates (r, z, φ) for a harmonic point source with time dependence $\exp(-i\omega t)$,

$$\nabla^2 p(r,z) + \left[\frac{\omega}{c(r,z)}\right]^2 p(r,z) = -\frac{\delta(r)\,\delta(z-z_0)}{2\pi r}, \tag{5}$$

where we have assumed azimuthal symmetry and hence no dependence on the φ-coordinate. Here, p is the pressure, ω the angular frequency of the source placed on the axis of symmetry at depth z_0, and $c(r,z)$ the sound speed in the medium. This partial differential equation must be solved with the appropriate pressure release and radiation boundary conditions.

Despite the continuous advances in computer technology, a direct numerical integration of Eq. (5) by finite-difference and finite-element techniques results in prohibitive computation times, except for extremely low-frequency, short-range problems.

Consequently, simplifying assumptions about the environment (2-D spatial variability, fluid bottoms only) or approximations to the acoustic wave equation must be introduced to arrive at practical acoustic models.

In the following, we briefly describe the derivation of three popular modeling techniques (WI, NM and PE) for time-harmonic problems, starting with range-independent wave theory.

4 • Time-Harmonic Solutions of Separable Problems

If we assume the ocean waveguide to be horizontally stratified, i.e., the sound speed $c(z)$ varies only with depth, Eq. (5) can be solved by separation of variables (1-D problem) in an orthogonal coordinate system. This class of problems typically consists of a series of plane-parallel fluid–solid layers terminated above by a vacuum halfspace and below by a solid halfspace.

4.1 Wavenumber Integration

A full-spectrum solution for the acoustic field in a plane-parallel waveguide can be obtained through a decomposition of the field into an infinite set of horizontally propagating cylindrical waves

$$p(r,z) = \int_0^\infty g(k,z)\, H_0^{(1)}(kr)\, k\, dk, \tag{6}$$

where k is the horizontal wavenumber, $g(k,z)$ is the depth-dependent Green's function, and $H_0^{(1)}(kr)$ is the Hankel function of order zero for outgoing waves. By substituting Eq. (6) into Eq. (5) we see that $g(k,z)$must satisfy the depth-separated wave equation

$$\frac{d^2g}{dz^2} + \left[\left(\frac{\omega}{c(z)}\right)^2 - k^2\right] g = -\delta(z - z_0). \tag{7}$$

To obtain the acoustic field at various depths as a function of range, we approximate the Hankel function in Eq. (6) by its asymptotic form and evaluate the integral numerically by means of an FFT. However, before doing this it is necessary to solve Eq. (7) with appropriate boundary conditions for a large number of discrete k-values covering the entire (but finite) spectrum of horizontally propagating cylindrical waves. For separable problems the wavenumber integration technique, also referred to as the fast field program (FFP) technique, constitutes a benchmark solution against which other approximate solutions can be checked.

The first implementations of this solution technique dates back to the 1970s, but several codes with improved computational efficiency are today available within the underwater acoustics community.

4.2 Normal Modes

The alternative to a direct numerical integration of Eq. (6) is to expand $g(k, z)$ into a complete set of normal modes, $g(k, z) = \sum A_m(k_m)\,\Psi_m(z)$, where the Ψ_m's are solutions of the eigenvalue equation

$$\frac{d^2\Psi_m}{dz^2} + \left[\left(\frac{\omega}{c(z)}\right)^2 - k_m^2\right]\Psi_m = 0, \tag{8}$$

which must be solved with the appropriate boundary conditions. The spectrum of eigenvalues generally consists of both a discrete part (trapped modes) and a continuous part (leaky modes). By ignoring the leaky modes, which correspond to sound interacting with the bottom at angles above the critical grazing angle (see Figure 4) and hence are suffering high reflection losses, the field solution can be written as a modal sum

$$p(r, z) = \sum_{m=1}^{M} \Psi_m(z_0)\,\Psi_m(z)\,H_0^{(1)}(k_m r). \tag{9}$$

Although this solution is complete only for waveguides with perfectly reflecting boundaries, it is a useful approximate solution to a wide class of long-range propagation problems in underwater acoustics.

Normal-mode models have been in use since the early 1970s and by now they clearly constitute the most popular wave-theoretic modeling tool for range-independent propagation situations. Several well-documented codes are available in the community, some of which utilize very efficient modal solvers that can provide up to an order-of-magnitude reduction in calculation time compared to standard solution techniques. This, in connection with the general increase in computational power, has promoted the use of normal-mode techniques in long-range propagation modeling in place of classical ray techniques.

5 • Time-Harmonic Solutions of Non-Separable Problems

If we consider non-separable problems where the sound speed $c(r, z)$ is allowed to vary with both depth and range (2-D range-dependent problems), we must distinguish between purely fluid media and mixed fluid–solid media. The first class of problems can be solved by stepwise-coupled normal modes or by the parabolic equation approximation, as discussed in detail below. The class of 2-D problems involving mixed fluid–solid media can be solved only by computer intensive finite-difference or finite-element techniques, which shall not be discussed in this chapter.

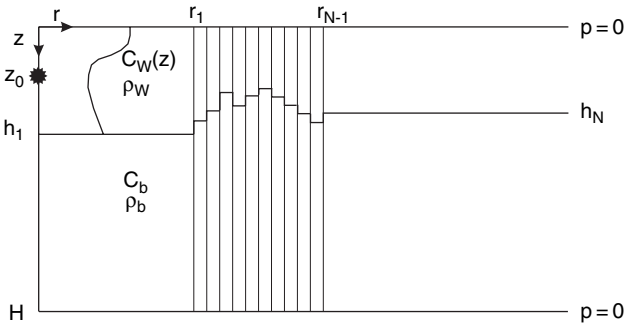


FIGURE 6 Discretization of environment for stepwise coupled mode solution.

5.1 Coupled Modes

A complete two-way solution for wave propagation in fluid media with 2-D inhomogeneity can be formulated in terms of stepwise-coupled normal modes. As shown in Figure 6, the medium variability in the r-direction has been discretized by subdivision into a number of range segments each with range-invariant properties, but with allowance for arbitrary variation of environmental parameters with depth within each segment. The solution for the acoustic field can be formally written as a sum of local modes with unknown excitation coefficients $A_{j,m}$ and $B_{j,m}$ giving the amplitudes of both outgoing and incoming cylindrical waves

$$p_j(r,z) = \sum_{m=1}^{M} \left[A_{j,m} H_0^{(1)}(k_{j,m} r) + B_{j,m} H_0^{(2)}(k_{j,m} r) \right] \Psi_m(z, k_{j,m}), \qquad (10)$$

where index j refers to the segment number and m to the mode number. The functions $\Psi_m(z, k_{j,m})$ are local modes, which are determined by solving an eigenvalue equation similar to Eq. (8) for each range segment. The unknown coefficients $A_{j,m}$ and $B_{j,m}$ are determined by requiring continuity of pressure and horizontal particle velocity across segment boundaries. This involves quite complex expressions with coupling integrals of the form

$$C_{j,mn} = \int_0^H \frac{\Psi_m(z, k_{j+1,m})\, \Psi_n(z, k_{j,n})}{\rho_{j+1}(z)}\, dz, \qquad (11)$$

to be evaluated for all modes at all segment boundaries.

The coupled mode technique does require considerable computational power and is therefore not a practical tool for solving general 2-D propagation problems. However, since the technique provides complete wave solutions with backscattering included, it is useful as a benchmark for checking approximate 2-D solutions.

If we neglect cross-mode coupling, we can derive a simple first-order approximation to the full coupling problem, called the adiabatic modal solution. It is here assumed that the energy contained in a given mode at the source range will stay in that particular mode throughout the range-varying environment and not couple to neighboring modes. The adiabatic solution is accurate only for "weak" range dependence, but it has been found to be applicable to many realistic ocean-acoustic scenarios. The adiabatic mode technique is computationally efficient since no coupling integrals need to be evaluated. Furthermore, the extension of a standard normal-mode code to include the adiabatic solution is trivial.

5.2 Parabolic Equations

An approximate solution to 2-D propagation problems can be obtained from the parabolic wave equation, which is derived by assuming a field solution of the form $p(r,z) = \phi(r,z)\exp(ik_0 r)\, r^{-1/2}$, where k_0 is an average horizontal wavenumber of the propagating wave spectrum. By substituting this expression for p into Eq. (5) in a source-free region and introducing the paraxial approximation, we arrive at the parabolic wave equation

$$\frac{\partial^2 \phi}{\partial z^2} + 2ik_0 \frac{\partial \phi}{\partial r} + k_0^2 (n^2 - 1)\phi = 0\,, \tag{12}$$

where $n(r,z) = k(r,z)/k_0$ is the refraction index. This partial differential equation is amenable to a marching solution based on either split-step FFT or finite-difference techniques. The classical split-step FFT solution takes the form

$$\phi(r+\Delta r, z) = \exp\left(\frac{ik_0}{2}(n^2-1)\Delta r\right)\ \mathcal{F}^{-1}\left\{\exp\left(\frac{-i\Delta r}{2k_0}k_z^2\right)\mathcal{F}\{\phi(r,z)\}\right\}, \tag{13}$$

where $\mathcal{F}$ denotes the Fourier transform from the z-domain to the k_z (vertical wavenumber) domain, and $\mathcal{F}^{-1}$ is the inverse transform.

The parabolic equation technique dates back to the early 1970s and has undergone several refinements and improvements over the years. There are now several parabolic equation (PE) codes available in the community, each solving a particular form of the parabolic wave equation. Today, the PE technique is the preferred wave-theoretic model for solving range-dependent problems.

In its original form, Eq. (12) is an approximate wave equation that neglects backscattering and, more importantly, assumes energy to propagate within a limited angular spectrum of $\pm 30°$ with respect to the horizontal. However, these limitations have

largely been overcome by the formulation of a family of high-angle parabolic wave equations which are applicable also to elastic media [1].

6 • Pulse Solutions by Fourier Synthesis

Work on direct simulation of broadband signal dispersion in an ocean waveguide has been under way for many years now. Within the framework of *linear* acoustics there are fundamentally two different approaches to this modeling problem. The first is to solve the pulse propagation problem directly in the time domain, which requires the development of an entirely new set of propagation codes. Time domain codes exist, but they have not yet matured to the same degree of algorithmic efficiency as single frequency codes. Hence, the alternative approach of generating broadband solutions by Fourier synthesis of harmonic results is the preferred one:

$$p(r, z, t) = \int_{-\infty}^{\infty} S(\omega)\, p(r, z, \omega)\, e^{-i\omega t}\, d\omega, \tag{14}$$

where $S(\omega)$ is the source spectrum and $p(r, z, \omega)$ is the spatial transfer function. This approach is attractive since it requires little programming effort. In fact any of the wave-theory codes described earlier can be linked up with a pulse post-processor which numerically integrates Eq. (14) with an FFT. The spatial transfer functions are generated by repeated execution of the time-harmonic code for a number of discrete frequencies within the frequency band of interest. Clearly, a pulse calculation by Fourier synthesis is computationally slow since typically hundreds (or thousands) of frequency samples are required to synthesize a pulse result. However, there is no indication so far that direct solution of the problem in the time domain has computational advantages. An overview of the many computer models in use in the underwater acoustics community today may be found in [6].

7 • Signal Simulation Examples

Since real ocean environments are too complex to provide test problems with easily interpretable results, simplified propagation situations (range-independent environments, homogeneous halfspace bottoms) are generally considered in numerical examples. This, however, is done purely for illustrative purposes. The strength of numerical models is that they allow for an arbitrary

variation in depth and range of the acoustic parameters of the environment, i.e., wave speed, wave attenuation, and density.

To illustrate the wide variety of acoustic problems that can be solved with current numerical codes, we present three examples. The first deals with long-range signal dispersion in a range-dependent shallow-water waveguide. Several codes (ray, mode, parabolic equation) have been applied to the same problem not only to show consistency in results, but also to address the issue of relative computational efficiency for the different solution techniques. The second example deals with propagation in a leaky surface duct, and the analysis is done here using a normal-mode code. The third and last example deals with global propagation in a 3-D environment, with the analysis done in the frequency domain using adiabatic normal-mode theory to describe the field structure in the vertical, and ray theory to describe horizontal propagation with refraction.

7.1 Shallow Water Waveguide

Although underwater acousticians have traditionally favored spectral analysis techniques for gaining information about the band-averaged energy distribution within a shallow-water waveguide, additional insight into the complication of multipath propagation is obtained by looking at signal transmission in the time domain.

From the schematic in Figure 7 we expect the signal structure measured downrange to consist of a number of arrivals with time delays determined by the path length differences, and individual pulse shapes being modified due to frequency-dependent amplitude and phase changes associated with each

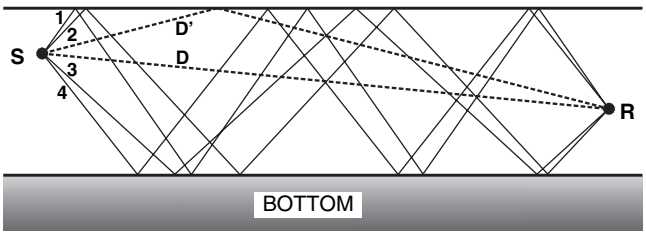


FIGURE 7 Schematic of ray arrivals in shallow-water waveguide. A series of multipath arrivals are expected, represented by *eigenrays* connecting source (S) and receiver (R). The shortest path is the direct arrival D followed by the surface-reflected arrival D'. Next comes a series of four arrivals all with a single bottom bounce; then four rays with two bottom bounces as illustrated in the figure, followed by four rays with three *bottom* bounces, etc.

boundary reflection. From simple geometrical considerations we find the *time dispersion* to be given by

$$\Delta\tau \simeq \frac{R}{\bar{c}}\left(\frac{1}{\cos\theta} - 1\right), \tag{15}$$

where R is the range between source and receiver, $\bar{c}$ is the mean sound speed in the channel, and θ is the maximum propagation angle with respect to the horizontal. This angle will be determined either by the source beamwidth or by the critical angle at the bottom (the smaller of the two). Since the dispersion considered here is solely due to the geometry of the waveguide, it is called *geometrical dispersion*.

Figure 8 is a simple computational example of pulse dispersion for a 100-m deep, isovelocity (1500 m/s) water column overlying a homogeneous halfspace bottom with a speed of 1800 m/s ($\theta_c \simeq 33°$), an attenuation of 0.1 dB/λ and a density of 2000 kg/m^3. The source is positioned at mid-depth (50 m) emitting a single sinusoidal pulse centered around 200 Hz. The receiver is 200 m downrange and consists of a vertical array of hydrophones spaced 10 m apart and covering the entire water column.

As seen in Figure 8 the multipath arrival structure is easily identified in this type of display. Hence, looking at a 30-m receiver depth, the direct arrival D comes first and is a simple replica of the source pulse. Next comes the surface reflected pulse S, which is out of phase with the direct arrival due to the phase shift

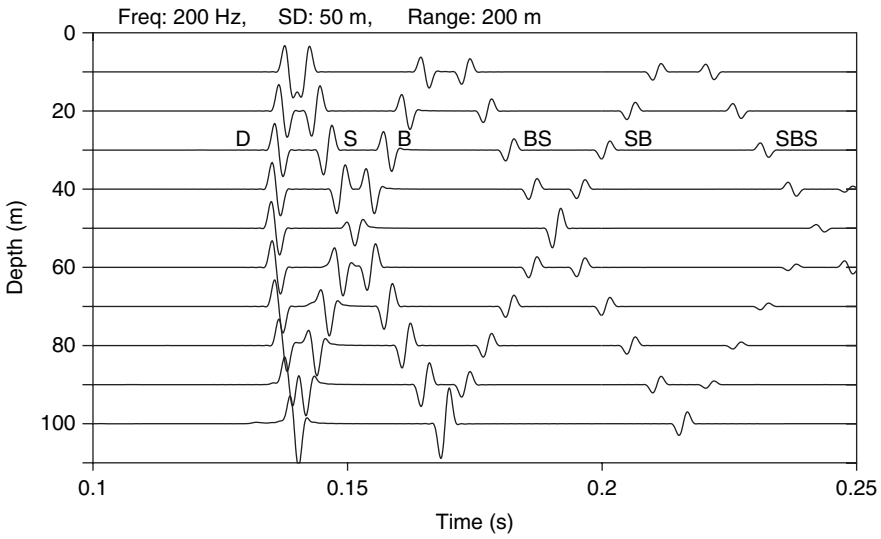


FIGURE 8 Pulse arrivals in a Pekeris waveguide.

of 180° associated with the reflection off the sea surface. Then follows four arrivals all with a single bottom bounce but with an increasing number of surface bounces. Next comes four arrivals with two bottom bounces, etc. Note that the pulse amplitude for the SB pulse is much reduced due to a supercritical incidence on the bottom.

We now proceed to a more complex shallow water scenario as shown in Figure 9. The environment is a symmetric upslope/downslope situations where the water depth varies from 200 m at the deep ends to 100 m at mid range. This geometry corresponds to a bottom slope of 1.15°. The sound-speed profile is downward refracting and is taken to be unchanged along the track. The bottom is a homogeneous fluid halfspace with the properties given in Figure 9.

We consider a broadband pulse emitted by a source at 100-m depth and calculate the received signal on a hydrophone at 20-m depth and at a range of 10 km. The emitted signal is a Ricker pulse, whose shape is the second derivative of a Gaussian, with center frequency of 500 Hz and covering the band 10–1000 Hz. Since the received signal at 10 km is found to have a total time dispersion of nearly 1 s, we know that a frequency sampling of 1 Hz is required to avoid signal wrap-around in the Fourier transformation ($\Delta f = 1/T$). Hence 991 frequency samples must be computed to synthesize the received signal at 10 km.

A comparison of pulse solutions for three different models (ray, coupled mode and PE) is shown in Figure 10. The early arrivals correspond to ray paths near the horizontal and hence with the shortest path length to the receiver. Steeper rays bounce repeatedly between surface and bottom and may arrive up to 1 s later, corresponding to an added path length of 1.5 km between source and receiver. Note in the upper panel that the signal

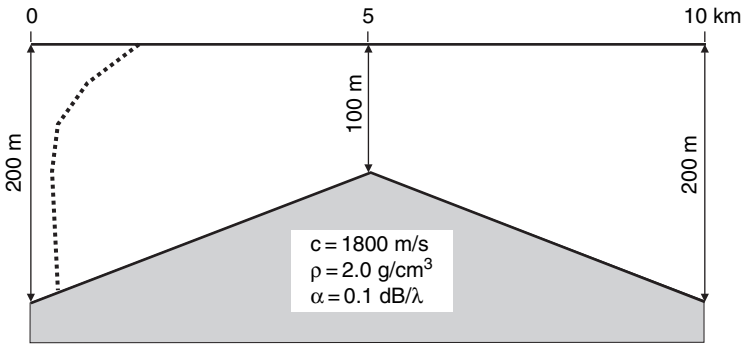


FIGURE 9 Shallow water test environment.

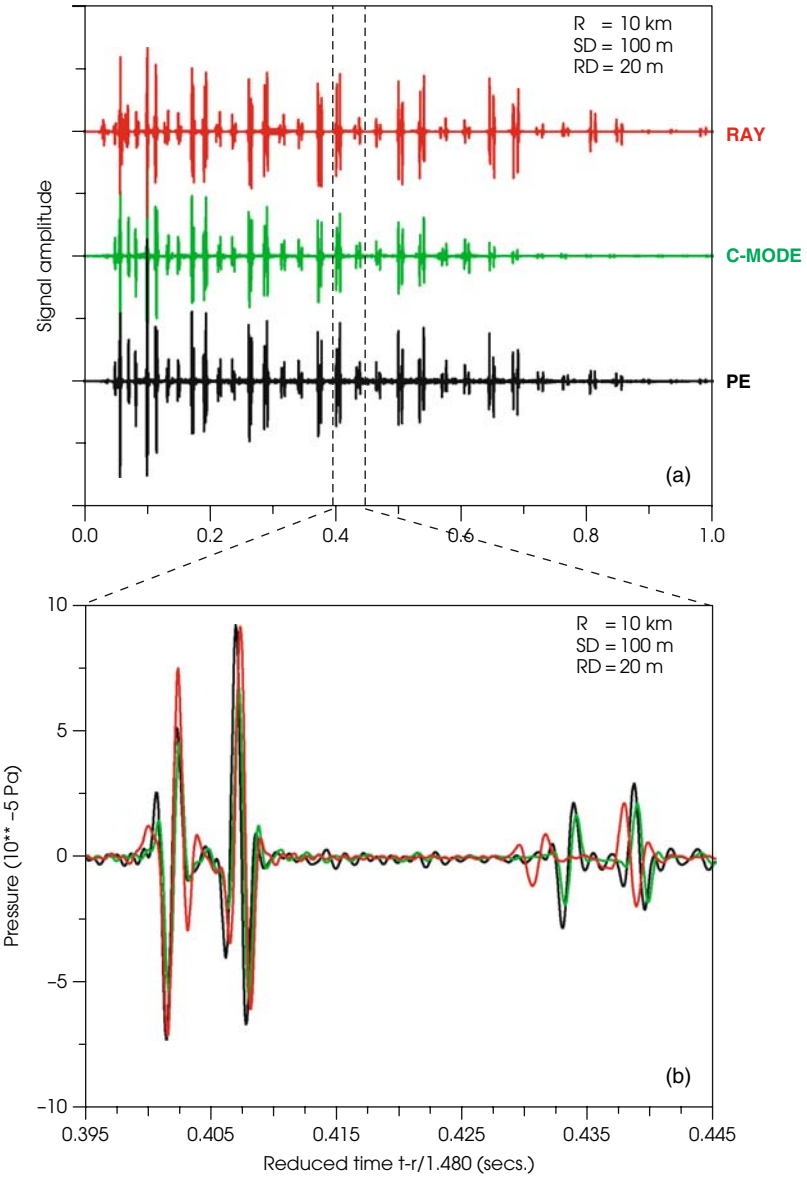


FIGURE 10 Comparison of broadband pulse solutions from three models. The source signal is a Ricker pulse with center frequency 500 Hz.

structure is very similar in all three solutions with a characteristic pattern of two strong arrivals followed by two weak arrivals, etc.

A closer analysis shows that each pair of arrivals consists of four eigenray contributions and, hence, the original pulse shape

is masked by interference effects. It can also be shown that the weak arrivals correspond to rays that pass in the vicinity of the receiver (diffracted energy), which has been accounted for in this particular ray model by assigning a Gaussian intensity distribution to each ray path. At higher frequencies all the weak arrivals disappear entirely from the time series and we are left with only the strong arrivals. These, in turn, can be classified in pairs (four eigenray arrivals) as being associated with an increasing number of bottom bounces. The critical angle effect is evident beyond 0.7 s, where the signal amplitude decays rapidly.

Although the arrival structures in the upper panel of Figure 10 are very similar, an expanded view (lower panel) shows some differences caused either by theoretical limitations in the models or by numerical discretization issues. The computation times on standard PC equipment vary from 10 min for the ray model to 10 h for the wave-theory models. More narrow-band signals can be computed faster, with the computation time being directly proportional to the signal bandwidth. Only the ray model is practical for broadband signal simulations above 1 kHz.

7.2 Leaky Surface Duct

In this second example we study propagation in a surface duct, in particular the existence of ducted "precursors" due to leakage. As shown in Figure 11, we use a generic deep-water profile with a 250-m deep surface duct. The ray diagram for a source at 50-m depth shows energy paths trapped in the duct as well as deep refracted paths being refocused near the surface at a range of 60 km (the convergence zone).

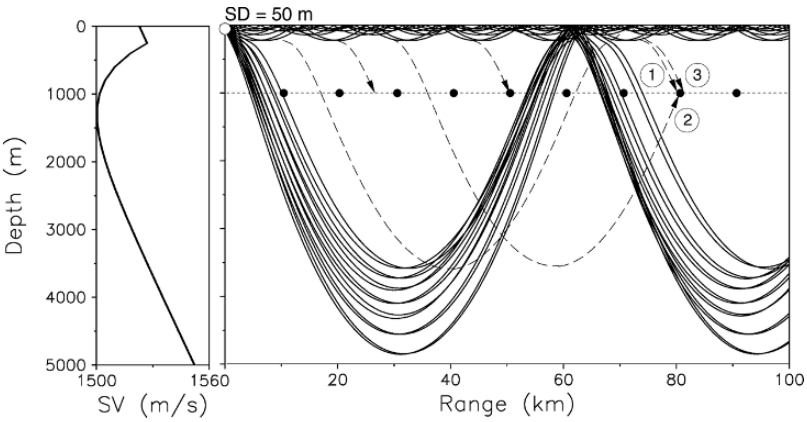


FIGURE 11 Sound-speed profile and associated ray diagram for the leaky surface-duct problem. The leaky paths are indicated by *dashed lines*.

In a wave-theory context, the surface duct allows for only one guided mode at a frequency of 100 Hz. This mode is a leaky (virtual) mode, which continuously sheds energy into the lower medium, as indicated by the dashed lines. Also shown in Figure 11 are examples of two leaky energy paths through the deep ocean (turning depth ~3500 m), which return to the surface duct approximately 60 km down-range, in between the first and the second convergence zone (CZ). Hence, assuming the leakage to be significant, propagation for both source and receiver in the surface duct will consist of purely ducted energy out to the first CZ, and by interference between the ducted mode and the leaky energy beyond the first CZ. In fact, the problem gets increasingly complicated as we move out in range, with an additional leaky arrival appearing after each passing of a convergence zone.

We shall investigate this problem in the time domain by computing the pulse arrivals at various receiver locations. As an example, we note that the field point located at range 80 km and 1000 m depth will be reached by three different arrivals: (1) A leaky arrival directly from the surface duct, (2) a leaky arrival which has passed one lower turning point, and (3) a leaky arrival which has gone through a lower turning point and a surface reflection. The modeling is performed by Fourier synthesis of single-frequency normal mode solutions.

Stacked time series are shown in Figure 12 for a receiver within the duct, RD = 200 m, and for a receiver below the duct, RD = 1000 m. Note that the increase in amplitude with range of the ducted path in Figure 12(a) is an artifact caused by the renormalization of all amplitudes with range r. The first arrival is due to the ducted mode, and this is the only arrival out to the first CZ (~ 60 km). Then we see a strong convergence-zone arrival (at 60 and 70 km), which, however, persists at longer ranges even though the ray diagram in Figure 11 shows that these field points are well beyond the first CZ. These "ghost" arrivals marked L-CZ in Figure 12 are associated with leaky energy having followed deep refracted paths back to the surface duct. Note that the leaky arrivals carry as much, or even more, energy than the ducted arrivals. This, of course, will be true only for particular environmental conditions. However, energy leakage definitely is important in this case where only one mode is trapped in the surface duct.

Looking next at the pulse arrival structure on a receiver at 1000-m depth, Figure 12(b), we see that all arrivals have been identified. The notations L-CZ$^-$ and L-CZ$^+$ refer to leaky CZ paths arriving at the receiver from below and from above, respectively. The three main arrivals are numbered in accordance with the

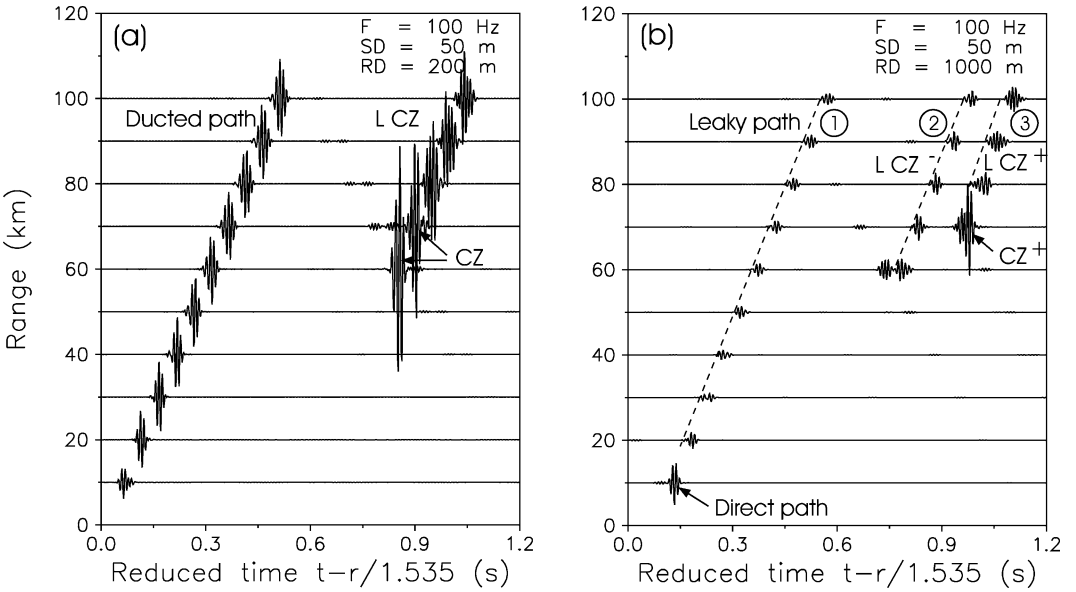


FIGURE 12 Stacked time series versus range at depths (**a**) 200 m and (**b**) 1000 m. The arrivals denoted "L-CZ" indicate leaky surface-duct energy which has followed a deep refracted (or CZ) path to the receiver. The superscripts + and − indicate whether the ray path gets to the receiver from above or from below.

ray-path numbering in Figure 11. Again we note that there is only one arrival out to the first CZ, the direct leaky path "1." At 70 km there is a strong CZ arrival preceded by two weaker arrivals often referred to as the *leaky precursors*. At longer ranges there are three leaky arrivals, and we notice that the last arrival is the strongest. This is because the duct is continuously leaking energy into the lower medium causing a rapid decay of the ducted pulse amplitude with range. As a consequence the leaky path "3" which originates at the shortest range will be the strongest arrival at long ranges.

It is evident that only a wave-theory-based pulse solution allows an unambiguous and complete dissection of a propagation situation as complex as the surface-duct problem considered here.

7.3 Global Propagation

In this last example, we show some results of a very-long-range modeling study of propagation from Perth to Bermuda [7]. On this global scale ($r \sim 20\,000$ km) prediction accuracy is closely tied to the available environmental information on bathymetry and sound-speed profiles along the various propagation paths. Moreover, 3-D effects due to Earth curvature and horizontal

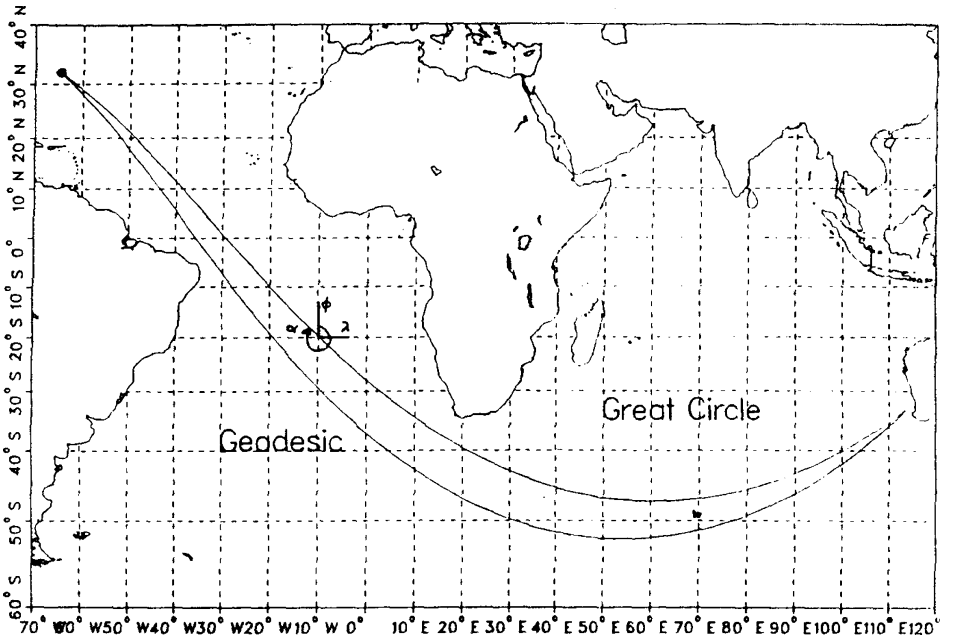


FIGURE 13 Shortest paths between Perth and Bermuda for a spherical Earth (great *circle*) and for an ellipsoidal Earth (geodesic) distorted by rotation.

refraction must be taken into account in order to provide accurate estimates of both travel time and signal dispersion.

The propagation geometry is shown in Figure 13. The difference between a great-circle path (spherical Earth) and a geodesic path (ellipsoidal Earth) is seen to be very significant. The geodesic path is displaced up to 800 km to the south, thus traversing cold, slow waters near Antarctica. The result is an increased travel time of approximately 30 s compared to the the great-circle path. Since the experimental results show a two-pulse arrival structure with the first arrival at 13364 ± 5 s (around 4 h) and the second appearing 30–40 s later, we require a modeling accuracy of at least 5–10 s on travel time in order to explain the experimental data. Consequently, the geodesic path associated with the flattened Earth geometry must be invoked.

The modeling of the Perth–Bermuda propagation experiment was done by adiabatic mode theory in the vertical and by ray theory in the horizontal [7]. For a source frequency of 15 Hz (estimated peak energy of received pulse), the modal structure was calculated on the horizontal grid shown as dots in Figure 14. Both sound-speed profiles and bathymetry were available from

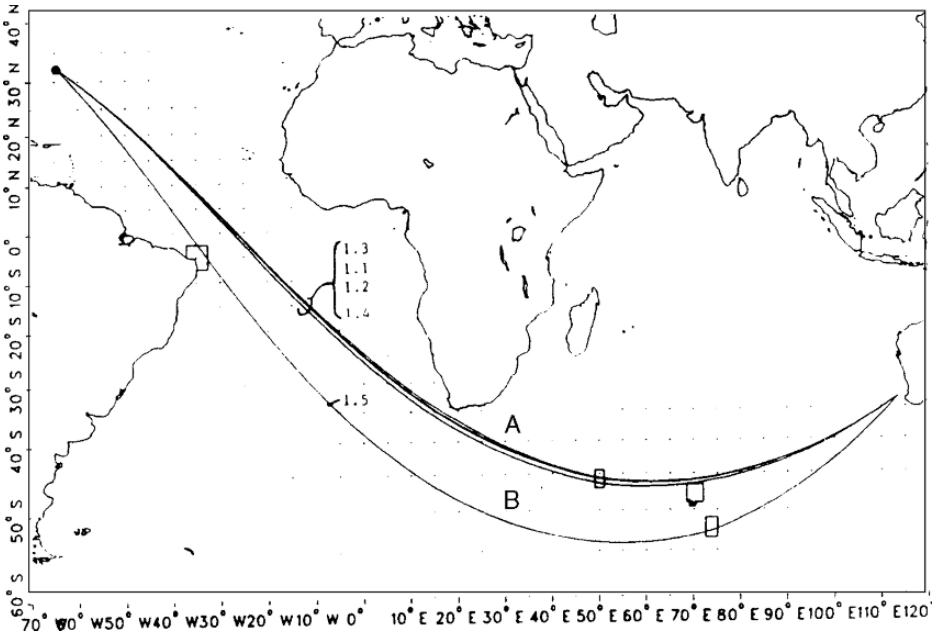


FIGURE 14 Five successful mode 1 rays between Perth and Bermuda. Sound-speed profiles are specified on the longitude–latitude grid points shown.

data bases on the same 230 grid points. Next, a horizontal ray tracing was performed for each mode and for a series of horizontal launch angles to find unobstructed propagation paths from Perth to Bermuda. This ray modeling includes horizontal refraction due both to transverse sound-speed gradients and to bottom interaction near land masses.

Examples of successful mode 1 paths are shown in Figure 14. There are five paths connecting Perth with Bermuda, four paths to the north (group A) and one path almost 1000 km further south (group B). Intermediate rays are blocked by islands. The higher-order modes show exactly the same behavior, with six unobstructed paths for mode 2 (three in A and three in B), and one each for modes 3 and 4.

The signal travel time along a ray can be shown to relate to the modal properties as

$$t_m = \int_0^S \frac{\omega}{k_m} \int_0^D \frac{1}{\rho(z)} \left[\frac{\Psi_m(z)}{c(z)} \right]^2 dz\, ds, \tag{16}$$

where ω is the angular frequency, k_m the modal eigenvalue, ρ the water density, Ψ_m the normalized modal eigenfunction, and c the sound speed over depth z.

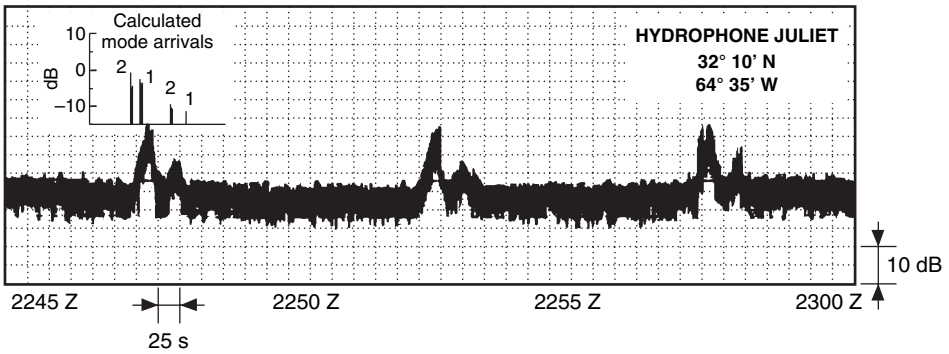


FIGURE 15 Data record of 21 March 1960 for three successive 300-lb explosive charges. The *upper left* insert gives arrival times and relative strengths for successful mode 1 and mode 2 arrivals.

Figure 15 displays the recorded time series at Bermuda from three consecutive explosions near Perth. The horizontal time scale is 25 s per tick mark, while the vertical dB scale is 10 dB per two tick marks. Note the characteristic double-pulse arrival structure for all three shots. In the upper left inset is shown computed mode arrivals with amplitudes normalized to 0 dB. First we see a set of mode 2 arrivals followed by several mode 1 arrivals, all pertaining to the faster group A rays. Approximately 30–40 s later comes the slower (and weaker) group B rays composed again of both mode 2 and mode 1 arrivals. Considering the environmental uncertainties in the data bases used in the calculations, the overall model–data agreement for this global propagation experiment is more than satisfactory.

This example shows the full potential of 3-D propagation modeling on a global scale, even considering the limited environmental information available. In fact, the success of this modeling exercise has important implications for the planned global warming experiments, which involve long-range acoustic transmissions through several of the World's oceans.

8 • Summary

Numerical models have become indispensable tools for both data analysis and sonar performance predictions in ocean acoustics. In particular, these models are used for Monte Carlo studies of acoustic signal fluctuations in the ocean, for tomographic time-domain inversions of ocean structures, for wideband geoacoustic inversions with global search algorithms, for designing underwater acoustic communication systems, and for testing signal

processing algorithms in sonar systems in general. For many practical applications, the computational effort involved in using broadband models is still excessive and more efficient solution techniques are continuously being developed.

References

1. F.B. Jensen, W.A. Kuperman, M.B. Porter and H. Schmidt, *Computational Ocean Acoustics* (Springer-Verlag, New York, NY, 2000).
2. L.M. Brekhovskikh and Yu. Lysanov, *Fundamentals of Ocean Acoustics*, 3rd edn. (Springer-Verlag, New York, NY, 2003).
3. X. Lurton, *An Introduction to Underwater Acoustics: Principles and Applications* (Springer-Praxis, London, UK, 2002).
4. H. Medwin and C.S. Clay, *Fundamentals of Acoustical Oceanography* (Academic Press, San Diego, CA, 1998).
5. R.J. Urick, *Principles of Underwater Sound,* 3rd edn. (Peninsula Publishing, Los Altos, CA, 1996).
6. P.C. Etter, *Underwater Acoustic Modeling and Simulation,* 3rd edn. (Spon Press, London, UK, 2003).
7. K.D. Heaney, W.A. Kuperman, and B.E. McDonald, "Perth–Bermuda sound propagation (1960): Adiabatic mode interpretation," J. Acoust. Soc. Am. **90**, 2586–2594 (1991).

94

Statistical Signal Processing

Edmund J. Sullivan

EJS Consultants, Portsmouth, RI, USA

1 • Background

The transfer of acoustic information can be represented as consisting of three parts: the source, the medium, and the receiver. The transmitted information (signal) arrives at the receiver distorted by the medium and corrupted by noise. Thus, even when the signal is deterministic in nature, a complete description of the received signal must be a statistical one. That is to say, any processing carried out on the received energy must contain the best characterization of the distortion by the medium and corruption by the noise that is available, and this can only be done in terms of statistics.

Signal processing is conventionally divided into three tasks: detection, estimation, and classification. Detection is defined as the determination of the existence or non-existence of a postulated signal at the receiver, and in its simplest form, is a simple binary (yes/no) decision. Estimation is the determination of the values of certain parameters of the signal, the source or the medium. A simple example is the determination of the bearing (angle) of an acoustic source at a receiving array of hydrophones. At its most complex, estimation leads to a class of problems called Inverse Problems, an example of which could be the extraction of the value of some property of the ocean bottom (density, sound speed, etc.) from the signal, or the characterization of a source or

scatterer, commonly called classification. Inverse Problems will be discussed in detail in the following chapter.

The quantification of detection and estimation is where statistics plays its major role. In estimation, the *quality* of the estimate is usually measured by its variance, which is a measure of the statistical spread of the estimate about its true value. In the case of detection, the performance is measured in terms of the probability of detection for a given probability of false alarm. This is embodied by the so-called Neyman-Pearson detector, which will be discussed here. There are more sophisticated approaches to detection based on formal hypothesis testing, and information on these approaches can be found in [1,2].

As mentioned above, some estimation problems fall into the class of inverse problems, which involves not just the receiver, but the medium and the source. In these cases, it is convenient to divide the processing task into two classes: parametric and non-parametric. What is meant by non-parametric is that the processing does not concern itself with other than the detection decision or the value of the estimate. Parametric, on the other hand, attempts to assign values to certain characteristics of the signal, as in time series analysis [2,3], or the source and the medium, as in inverse problems such as Matched Field Problems (MFP) [4,5] or Ocean Acoustic Tomography (OAT), both of which are discussed in the following chapter. These parameters may or may not be easily identified as directly representing actual physical parameters. When they are, it is convenient to call this type of parametric processing Model-Based Processing (MBP) [6,7], an example of which is MFP. In any case, parametric processing is a means of improving the results by introducing structure (models) into the signal, the medium or, the source. These ideas are depicted in Figure 1.

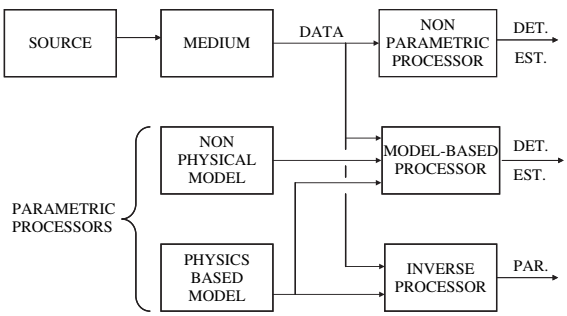


FIGURE 1 Classification of processors.

2 • Detection

In the underwater signal processing world, the Binary Hypothesis describes most detection problems. That is, we are concerned with determining whether a signal does or does not exist. And in turn, these problems are most often dealt with by the Neyman-Pearson (N-P) detector [1,2]. The N-P detector delivers the highest probability of detection for a given probability of false alarm. The performance of such a detector is most conveniently described by a set of curves called the Receiver Operating Characteristics or "ROC Curves" (see Chapter 96). Note that this approach allows for two kinds of errors. Deciding that the signal is present when it is not is called an error of Type I, and deciding that the signal is not present when it is, is called an error of Type II.

As a simple example of the N-P detector, consider the problem of detecting the presence of a non-zero DC level in white Gaussian noise. The binary Hypothesis is stated as follows. Designating the DC level as v, the white Gaussian noise as n with variance σ^2 and the received signal as y, we have

$$\begin{aligned} H_0 &: y = n \\ H_1 &: y = v + n \end{aligned} \tag{1}$$

The associated Likelihood Functions are given by

$$p(y/H_0) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-y^2/2\sigma^2}, \tag{2}$$

and

$$p(y/H_1) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(y-v)^2/2\sigma^2}. \tag{3}$$

The N-P detector is based on the *likelihood Ratio*, defined by

$$L = \frac{p(y/H_1)}{p(y/H_0)} = \frac{e^{-(y-v)^2/2\sigma^2}}{e^{-y^2/2\sigma^2}}. \tag{4}$$

The N-P detector decides that H_1 is true if $L > \gamma$, where γ is determined by the desired probability of false alarm, and that H_0 is true otherwise. In practice, the logarithm of L is used so that the test statistic, lnL, is compared with $ln\gamma$. Thus, for our example, Eq. (4) tells us that H_1 is true if

$$\frac{yv}{\sigma^2} > \frac{v^2}{2\sigma^2} + ln\gamma. \tag{5}$$

This is cast in terms of probability as follows. The probability of detection, i.e., the probability that H_1 is true, is given by

$$P_D = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_T^\infty e^{-(y-v)^2/2\sigma^2} dy, \tag{6}$$

and the probability of false alarm, P_{FA}, is given by

$$P_{FA} = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_T^\infty e^{-y^2/2\sigma^2} dy, \tag{7}$$

where T is the threshold. The recipe for designing the N-P detector now follows immediately. The desired P_{FA} is chosen, which defines the threshold T via Eq. (7). Given T, Eq. (6) is then used to compute P_D.

2.1 Known Signals

The above example is based on a single sample. In the case of a known, discretely sampled time-varying signal v_n with N independent samples, the above argument can be generalized as follows. Assuming Gaussian white noise (GWN), Eq. (2) becomes

$$p(\mathbf{y}/H_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=1}^N y_n^2}. \tag{8}$$

Correspondingly, Eq. (3) generalizes to

$$p(\mathbf{y}/H_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=1}^N (y_n - v_n)^2}. \tag{9}$$

Here, $\mathbf{y}$ represents the vector of data samples with elements $\{y_n\}$. Note also that Eqs. (8) and (9) are simply the joint probabilities of the N independent samples for the two cases. For this case Eq. (5) generalizes to

$$\frac{1}{\sigma^2}\sum_{n=1}^N y_n v_n > \frac{1}{2\sigma^2}\sum_{n=1}^N v_n^2 + ln\gamma. \tag{10}$$

Here we see that, as one would expect, the threshold increases with the signal energy, which is given by

$$E = \sum_{n=1}^N v_n^2. \tag{11}$$

Upon rearranging Eq. (10), we finally get

$$\sum_{n=1}^N y_n v_n > \frac{1}{2}\sum_{n=1}^N v_n^2 + \sigma^2 ln\gamma. \tag{12}$$

Note that the term on the LHS is the correlation between the data and the signal. When the form of the signal is known a priori, this test statistic can be computed. This is the well-known matched filter, and in the form shown, is known as the *Replica Correlator*. The replica correlator is discussed in more detail in Chapter 96.

2.2 Unknown Signals

For the case of unknown signals, the approach is the same as for known signals. The likelihood ratio is computed and a threshold is set based on a desired P_{FA}. Consider the case of N samples of a Gaussian white signal, s_n, with unknown variance σ_s^2, in Gaussian white noise, n_n with variance σ_n. Then

$$\begin{aligned} H_0 &: y_n = n_n \\ H_1 &: y_n = s_n + n_n \end{aligned} \tag{13}$$

The likelihood functions follow as

$$p(\mathbf{y}/H_0) = \frac{1}{(2\pi\sigma_n^2)^{N/2}} e^{-\frac{1}{2\sigma_n^2}\sum_{n=1}^{N} y_n^2}. \tag{14}$$

and

$$p(\mathbf{y}/H_1) = \frac{1}{(2\pi[\sigma_n^2+\sigma_s^2])^{N/2}} e^{-\frac{1}{2(\sigma_n^2+\sigma_s^2)}\sum_{n=1}^{N} y_n^2}. \tag{15}$$

and the log likelihood follows as

$$lnL = \left(\frac{N}{2}\right) ln\left[\frac{\sigma_n^2}{2(\sigma_n^2+\sigma_s^2)}\right] + \left(\frac{1}{2\sigma_n^2}\right)\left[\frac{\sigma_s^2}{2(\sigma_n^2+\sigma_s^2)}\right]\sum_{n=1}^{N} y_n^2. \tag{16}$$

Here we see that the relevant detection statistic is $\sum_{n=1}^{N} y_n^2$. This is known as an energy detector, since it essentially computes the energy of the received data and compares it to a threshold. It is treated in detail in [1].

3 • Estimation

As mentioned in Section 1, estimation is the determination of the values of certain parameters of the signal, the source or the medium. It is convenient to classify the estimation task into two types: parametric and non-parametric. Parametric estimation can be further separated by identifying those cases where the parameters have a clear physical basis, such as ocean bottom density, sound speed, etc., and those cases where this is not necessarily so, which can be the case in time series analysis. To put in a different way, parametric estimation consists of providing some kind of parameter-based structure to the signal. This is done for either of two reasons: to improve the performance of the estimator or to extract estimates of certain parameters of the source or the medium.

3.1 Estimator Structure

The simplest example of a non-parametric estimator is the linear least-squares (LLS) estimator. Here one seeks the value of x that best characterizes the data by minimizing the mean squared error. In this case, the data (received signal) is modeled as

$$y_i = ax + n_i; \quad i = 1, 2, \cdots, N, \tag{17}$$

where n_i represents a noise sequence. The estimator seeks the minimum of J, where

$$J = \sum_{i=1}^{N} (y_i - ax)^2. \tag{18}$$

From $dJ/dx = 0$, the estimate of x is easily seen to be

$$\hat{x} = \frac{1}{a}\left[\frac{1}{N}\sum_{i=1}^{N} y_i\right] = \frac{1}{a}\hat{y}_N. \tag{19}$$

Thus, the estimate is based on the mean of the data set. More generally, the signal may be a vector of measurements and the model may contain more than one parameter. Then the (linear) system becomes

$$\mathbf{y}_i = A\mathbf{x} + \mathbf{n}_i, \tag{20}$$

where in general, A is an $N \times M$ matrix. $\mathbf{y}_i$ and $\mathbf{n}_i$ are the ith realizations of the N-dimensional data (measurement) and noise vectors, respectively, and $\mathbf{x}$ is the M-dimensional parameter vector. If $N = M$, the solution for $\mathbf{x}$ for a given realization simply requires the inversion of A. If $N < M$ the problem is underdetermined, and cannot be solved without more information (see Chapter 95). When $N > M$, the least-squares solution is given by

$$\hat{\mathbf{x}}_i = (A^T A)^{-1} A^T \hat{\mathbf{y}}_i = P\hat{\mathbf{y}}_i \tag{21}$$

where P is sometimes referred to as the Moore–Penrose inverse, or the pseudoinverse [8,9]. It is a linear estimator and is based on a linear signal model. In general, the matrix A may contain relevant parameters, so that Eq. (21) can represent a parametric estimator. However, this leads to the subject of non-linear estimation, which is beyond the scope of this chapter.

Equation (21) can be viewed from two different points of view. If there is only one realization of the data vector, then Eq. (21) is the least-squares solution for $\mathbf{x}$. If, on the other hand, we assume

that i labels time and we have a sequence of data vectors, Eq. (21) can be considered to be a *preprocessor* which produces a sequence of estimates of the parameter $\mathbf{x}$, i.e., $\{\mathbf{x}_i\}$. In the stationary case, this is simply an extension to two dimensions of the least-squares problem. In the non-stationary case, however, the form of Eq. (21) leads to a generalization known as the Kalman filter, which will be discussed in Section 3.4.1.

At this point there two issues that must be addressed. The first is the question of optimality, and the second is how one selects an appropriate estimator for a particular application.

3.2 The Cramér–Rao Lower Bound

The quality of an estimate is most commonly measured in terms of its variance. For the case of Eq. (17), with $a = 1$ and $M = 1$, the estimate of x, which is the mean of y, is given by

$$\hat{x} = \frac{1}{N}\sum_{i=1}^{N} y_i = \frac{1}{N}\sum_{i=1}^{N}(x + n_i) = E\{x\} + \frac{1}{N}\sum_{i=1}^{N} n_i. \tag{22}$$

The variance on the estimate follows as

$$\sigma_x^2 = E\{(\hat{x} - E\{x\})^2\} = E\left\{\left(\frac{1}{N}\sum_{i=1}^{N} n_i\right)^2\right\} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E\{n_i n_j\}, \tag{23}$$

and since we have assumed Gaussian white (uncorrelated) noise, this reduces to

$$\sigma_x^2 = \frac{1}{N^2}\sum_{i=1}^{N} E\{n_i^2\} = \frac{\sigma^2}{N}, \tag{24}$$

where, $\sigma^2 = E\{n^2\}$.

The Cramér–Rao Lower bound (CRLB) [10] is the lowest achievable variance on an unbiased estimate for a given likelihood function, and an estimator that achieves this bound is said to be efficient. It is found as follows. The likelihood function for a signal y, that depends upon an M-dimensional parameter vector $\mathbf{x} = [x_1,\ x_2, \cdots, x_M]$, is the probability of y given $\mathbf{x}$. It is written as $p(y/\mathbf{x})$, with $\mathbf{x}$ considered to be fixed. The Fisher Information Matrix (FIM), which is defined as

$$I_{ij} = -E\left\{\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} ln[p(y/\mathbf{x})]\right\}, \tag{25}$$

is then evaluated. The CRLB on the estimate of x then follows as

$$\sigma_{x_k}^2 \geq (I^{-1})_{kk}. \tag{26}$$

For our present example, since n_i in Eq. (17) represents a white Gaussian process,

$$p(y/x) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=1}^{N}(y_n - x)^2}, \tag{27}$$

so that from Eq. (25)

$$I = \frac{N}{\sigma^2}, \tag{28}$$

which leads to the result that the lower bound on the variance of the estimate of x is simply the noise variance divided by N, i.e.,

$$\text{CRLB}(x) = \frac{\sigma^2}{N} \tag{29}$$

But this is the same as found in Eq. (24). Thus, the LLS estimator is efficient.

3.3 Hierarchy of Estimators

When choosing an estimator, two things must be considered. These are (1) the structure of the signal model and (2) the character of the data. At the most general level, estimators fall into two classes. These are minimum variance (MV) and maximum likelihood (ML). When unbiased, the MV estimator (MVU) is optimum; however, it is usually not a simple task to find a MVU that applies to the problem at hand. If the signal model is linear, such as in Eq. (20), then the optimal MVU for Gaussian non-white noise (GNWN) is given by [10]

$$\hat{\mathbf{x}} = (A^T R^{-1} A)^{-1} R^{-1} A^T \mathbf{y}, \tag{30}$$

with

$$p(\mathbf{y}/\mathbf{x}) = \frac{1}{|R|^{1/2}(2\pi)^{N/2}} e^{-(\mathbf{y}-A\mathbf{x})^T R^{-1}(\mathbf{y}-A\mathbf{x})}, \tag{31}$$

and

$$R = E[\mathbf{n}\mathbf{n}^T], \tag{32}$$

with $|R|$ being the determinant of R. For the case of white noise (GWN), Eq. (30) reduces to

$$\hat{\hat{\mathbf{x}}} = (\mathbf{A^T A})^{-1} \mathbf{A^T y} \tag{33}$$

which is the least-squares estimator of Eq. (21).

In the case of non-Gaussian noise, the optimal estimator will, in general, be non-linear. If the optimal processor cannot be found,

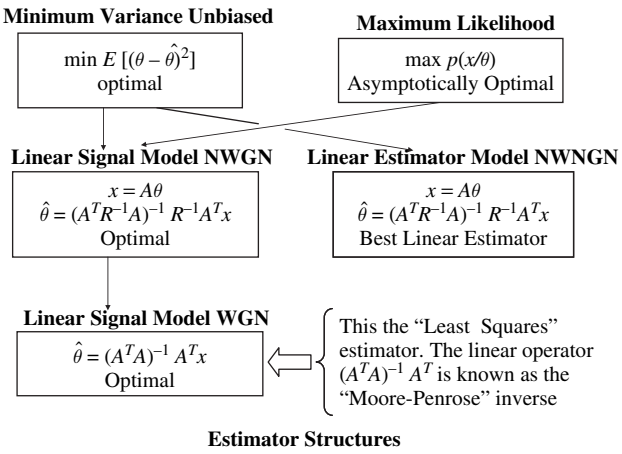


FIGURE 2 Classification of Estimators.

then the "best" is that given by Eq. (30), and is usually referred to as the Best Linear Unbiased Estimator or BLUE [10].

The estimators outlined above are in closed form and require only the noise variance. However, if the noise statistics are not known, complete information on optimality is not available. When $p(y/x)$ is known, the ML type of estimator can be used. The advantage of this is that it is reasonably easy to find numerical solutions by searching for an extremum of the likelihood function and in the case of non-linear models, this is a distinct advantage. In the case of linear models, however, the differences between the two types of estimators becomes moot, since the closed form given in Eq. (30) is obtained in both cases. This is depicted in Figure 2. For the case when no statistical information available, the LS estimator is the only approach available, whether the model is linear or not. Of course, for the non-linear model, a closed form is generally not available, and the solution must be found numerically. A further complication here is that the likelihood surface may not be unimodal, i.e., there may be multiple extrema.

3.4 More General Estimator Structures

The preceding section discussed that group of estimators sometimes referred to as *classical* estimators, since they deal with deterministic parameters. They also form the basis for batch-type processes, since they are employed by first collecting all of the data, and then exercising the algorithm. However, there are certain classes of problems that are best dealt with by using recursive-type processors. For example, if the statistics are not stationary, or the parameters of interest are changing in time,

a batch process clearly cannot be optimal. A simple example of a recursive processor can be constructed from the example in Eq. (21). Consider i to label the time index for a sequence of N measurements of the vector $\{\hat{\mathbf{y}}_i\}$. Upon separating out the *Nth* term, the result is

$$\mathbf{x}_N = P\left[\frac{1}{N}\mathbf{y}_N + \frac{1}{N}\sum_{i=1}^{N-1}\mathbf{y}_i\right]. \tag{34}$$

Making the identification

$$\hat{\mathbf{x}}_{N-1} = P\hat{\mathbf{y}}_{N-1}, \tag{35}$$

where

$$\hat{\mathbf{y}}_{N-1} = \frac{1}{N-1}\sum_{i=1}^{N-1}\mathbf{y}_i, \tag{36}$$

Eq. (34) can be rewritten as

$$\hat{\mathbf{x}}_N = \hat{\mathbf{x}}_{N-1} + K\left[\mathbf{y}_N - A\hat{\mathbf{x}}_{N-1}\right], \tag{37}$$

Where we have defined $K = P/N$. This is referred to as the *gain*. Note that $\hat{\mathbf{y}}_{N-1} = A\hat{\mathbf{x}}_{N-1}$. This we call the *measurement equation* and we shall refer to $\hat{\mathbf{y}}_{N-1}$ as the predicted measurement, i.e., the estimate of $\mathbf{y}$ at time t_i based upon the measurements up to time t_{i-1}.

Equation (37) is now rewritten in a somewhat more general notation in order to clarify the recursive nature of the algorithm.

$$\hat{\mathbf{x}}(t_i|t_i) = \hat{\mathbf{x}}(t_i|t_{i-1}) + K\boldsymbol{\epsilon}(t_i). \tag{38}$$

Here, we have defined the *innovation* as $\boldsymbol{\epsilon}(t_i) = [\mathbf{y}(t_i) - A\hat{\mathbf{x}}(t_i|t_{i-1})]$.

Equation (38) states that the estimate of $\mathbf{x}$ based on the data up to and including time t_i is given by its estimate based on the data up to and including time t_{i-1} plus a correction term, where the correction term is the gain multiplied by the innovation. $\hat{\mathbf{x}}(t_i|t_i)$ is called the *corrected* estimate, $\hat{\mathbf{x}}(t_i|t_{i-1})$ is the *predicted* estimate, and $\hat{\mathbf{y}}(t_i|t_{i-1}) = A\hat{\mathbf{x}}(t_i|t_{i-1})$ is the predicted measurement. Thus, with each new measurement, the updated estimate can be found in terms of the old estimate without the need to recompute the batch. The algorithm is as follows.

- Measurement prediction
  $\hat{\mathbf{y}}(t_i|t_{i-1}) = A\hat{\mathbf{x}}(t_i|t_{i-1})$
- Make new measurement $\mathbf{y}(t_i)$ and compute the innovation
  $\boldsymbol{\epsilon}(t_i) = [\mathbf{y}(t_i) - A\hat{\mathbf{x}}(t_i|t_{i-1})]$

- Correct initial estimate
  $\hat{\mathbf{x}}(t_i|t_i) = \hat{\mathbf{x}}(t_i|t_{i-1}) + K\boldsymbol{\epsilon}(t_i)$
- Return to first step

3.4.1 The Kalman Filter

The form of the recursive least-squares (RLS) estimator given in Eq. (38) leads us in a natural way to the Kalman filter [7,11]. One way to look at the Kalman filter (KF) is as a generalization of the RLS. In Eq. (38) it is implicitly assumed that the parameter vector $\mathbf{x}$ is a constant and that the model is "perfect." The Kalman filter avoids these assumptions by replacing the model of Eq. (20) with a first-order Gauss–Markov model [6]. A complete treatment of the Gauss-Markov model is beyond the scope of this chapter, but for our purposes it is sufficient to think of it as a first-order recursive estimator with Gaussian noise. That is, the estimate at any given point in the data sequence is dependent only on the immediately previous estimate plus a Gaussian noise term. What this does is to permit essentially any linear[1] dynamical model to be employed in the estimator structure. The implications of this are enormous, and the Kalman filter is found in an extremely broad array of applications requiring sophisticated estimation and automatic control [7]. Here, we will outline it in its most common form, commonly referred to as the *predictor–corrector* form [6].

The KF is based on two sets of equations, the state equations and the measurement equations, usually written in the following form, where

$$\mathbf{x}(t_i) = A\mathbf{x}(t_{i-1}) \tag{39}$$

is the state equation and

$$\mathbf{y}(t_i) = C\mathbf{x}(t_i) \tag{40}$$

is the measurement equation. A is the state transition matrix and C is the measurement matrix. Thus, the state transition matrix carries in it any dynamics, and it is the task of the user to determine how to insert his/her model into this form. If the state vector is a constant, then A is simply some multiple of the identity matrix. A common approach taken for more complex cases is to put the dynamics into the form of a first-order differential equation, since the solution can then be represented as a

[1]Here we treat only the linear Kalman filter. The non-linear case is treated by the Extended Kalman Filter (EKF) which is beyond the scope of this chapter.

recursive first-order process. The utility of this is that, via the state-space [6] approach, a linear differential equation of any order, say N, can always be reduced to a set of N first-order differential equations. This then allows the solution to put into the form of Eq. (39). The derivation of the KF is then carried out based on the Gauss–Markov state-space model given by

$$\mathbf{x}(t_i) = A\mathbf{x}(t_{i-1}) + w(t_i), \tag{41}$$

and the measurement model given by

$$\mathbf{y}(t_i) = C\mathbf{x}(t_i) + v(t_i) \tag{42}$$

Here, $w(t_i)$ and $v(t_i)$ are the Gaussian state and measurement noises, respectively. Although we will not show the full algorithm here, the procedure is as follows. An initial guess of the state vector is made. Clearly, the better this is, the more rapid will be the convergence to a solution. This initial state then permits an estimate of the next measurement in the data sequence.[2] The algorithm is as follows.

- Measurement prediction
  $\hat{\mathbf{y}}(t_i|t_{i-1}) = A\hat{\mathbf{x}}(t_i|t_{i-1})$
- Make new measurement $\mathbf{y}(t_i)$ and compute the innovation and gain
  $\epsilon(t_i) = [\mathbf{y}(t_i) - A\hat{\mathbf{x}}(t_i|t_{i-1})]$
- Correct initial estimate
  $\hat{\mathbf{x}}(t_i|t_i) = \hat{\mathbf{x}}(t_i|t_{i-1}) + K(t_i)\epsilon(t_i)$
- Return to first step

This is nearly the same as the recursion given previously for the RLS. The difference here is that the gain, K, is a function of time. This is the heart of the Kalman algorithm. The KF is derived as a recursive minimum-variance estimator for a GM model. As such, the recursive computation of K requires knowledge of the (recursive) values of the relevant statistics in their second-order form. That is,

$$K(t_i) = \frac{C^T \tilde{P}(t_i) C}{R_{\epsilon\epsilon}^{-1}(t_i)}. \tag{43}$$

[2]The data sequence is usually a time sequence, but it could as well be a sequence of spatial measurements.

Here, $\tilde{P}(t_i)$ is the state error covariance matrix and $R_{\epsilon\epsilon}(t_i)$ is the covariance of the innovation ϵ. These are recursively computed by the algorithm. The user must initialize the algorithm with values of the measurement noise covariance R_{vv}, the state noise covariance R_{ww}, and an initial guess for the state error covariance $\tilde{P}_0$.

3.4.2 The Tracking Problem

A common application of the KF is that of tracking of a moving acoustic source using a towed array. Although we shall not actually solve it here, we will formulate it as an example of how one introduces a model into Eqs. (41) and (42), which are the basic KF equations.

This example is sometimes referred to as the "bearings only" tracking problem [12,7], since the only measurements are bearing estimates from a towed array. It is cast as a two-dimensional problem, and it is normally assumed that the source (target) is moving at a constant velocity. This problem can be cast in the form of a first-order linear differential equation, so it fits nicely with the discussion above.

Defining the speed and range of the towed array (observer) as v_0 and r_0, respectively, and the speed and range of the target as v_t and r_t, respectively, we have the following relationship between the speed and range of the target relative to the observer, as depicted in Figure 3.

$$\frac{dr(t)}{dt} = v(t), \tag{44}$$

or

$$dr(t) = v(t)dt, \tag{45}$$

where

$$v_x(t) = v_{tx}(t) - v_{ox}(t), \quad r_x(t) = r_{tx}(t) - r_{ox}(t)$$
$$v_y(t) = v_{ty}(t) - v_{oy}(t), \quad r_y(t) = r_{ty}(t) - r_{oy}(t)$$

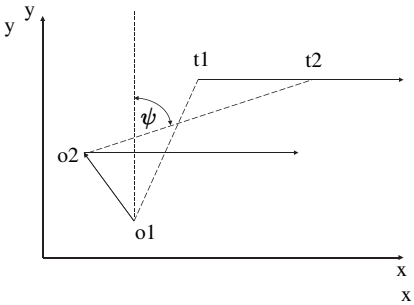


FIGURE 3 Geometry for the tracking problem.

We now discretize the problem with time increment Δt, so that Eq. (45) is approximated by

$$r(t) = r(t-1) + v(t-1)\Delta t. \tag{46}$$

And since we have assumed that the *target* velocity is constant,

$$v(t) = v(t-1) + \Delta v_o(t). \tag{47}$$

The *state vector* is now defined as

$$\mathbf{x}(t) = [r_x(t) \;\; r_y(t) \;\; v_x(t) \;\; v_y(t)]^T, \tag{48}$$

so that Eqs. (46) and (47) can be written as

$$\begin{bmatrix} r_x(t) \\ r_y(t) \\ v_x(t) \\ v_y(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_x(t-1) \\ r_y(t-1) \\ v_x(t-1) \\ v_y(t-1) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \Delta v_{ox}(t-1) \\ \Delta v_{oy}(t-1) \end{bmatrix}, \tag{49}$$

or more concisely,

$$\mathbf{x}(t) = A\mathbf{x}(t-1) + B\mathbf{u}(t-1), \tag{50}$$

where $\mathbf{u}(t) = [\Delta v_{ox}(t) \;\; \Delta v_{oy}(t)]^T$ is referred to as the *source* term, B is the source throughput matrix and A is the *state transition matrix*. The bearing angle to the target from the array is given by

$$\psi(t) = \arctan(r_x(t)/r_y(t)). \tag{51}$$

This is the measurement equation and, unlike the state equation, it is non-linear. That is, it cannot be written in the form of Eq. (40), and therefore the linear KF algorithm cannot be used. The solution using the Extended Kalman Filter (EKF) is carried out in [6]. This problem requires two independent bearing measurements for a solution. This can be seen by inspection of Eq. (51), where it is seen that, for a given bearing angle, ψ, the range is unobservable [6]. This is due to the fact that the bearing estimation assumes a plane wave model for the signal. In practice, this problem is dealt with by carrying out a maneuver and performing a second (and therefore independent) bearing measurement. Referring to Figure 3, the points labeled $o1$ and $o2$ are the respective starting point and maneuver point for the observer, and the points labeled $t1$ and $t2$ are the corresponding points for the target. ψ is the bearing just after the maneuver.

We conclude this example with its Gauss–Markov model representation, which includes the system and measurement noise terms.

$$\mathbf{x}(t) = A\mathbf{x}(t-1) + B\mathbf{u}(t-1) + \mathbf{w}(t)$$
$$\psi(t) = c[\mathbf{x}(t)] + v(t)$$

The measurement model is written in a symbolic form to emphasize the fact that it is non-linear. As it turns out, most problems of interest are non-linear, either in the state equation or in the measurement equation, or in both. As mentioned above, such problems can be treated with the Extended Kalman Filter (EKF). It is based on a first-order Taylor series expansion of the state and/or measurement equation, and is therefore an approximation. Nevertheless, it is widely used with great success.

3.4.3 Discussion

The Kalman Filter constitutes a significant improvement over the classical estimator, since it has an inherent self-correction capability. The measurement noise plays the role of noise as it is commonly understood, i.e., the noise corrupting the measurement. The state noise, sometimes called the *plant* noise, is a means of accounting for modeling errors. Rather than being a noise in the usual sense, it is a statistical measure of our degree of ignorance of the complete details of the model. As such, it permits a "graceful" degradation of the solution in light of incomplete knowledge of the model, rather than a failure to get any solution at all. This is related to the so-called "mismatch problem" [4,13] encountered in Matched-Field Processing (MFP) (see Chapter 95). The introduction of a model into a signal processing procedure can have a significant impact on the solution. Indeed, this is why it is done. As a result, when a *deficient* model is introduced into an estimation procedure, the estimator may perform worse than if no model at all had been used. In the case of MFP, this results in no solution at all or, at best, in an incorrect one. This is sometimes referred to as the *mismatch* problem [4,13]. In the case of a Gauss–Markov based recursive estimator, this problem is greatly ameliorated. Instead of the estimator failing, the adjustment of the input noise covariance terms, $\tilde{P}_0$, R_{vv} and R_{ww}, allows the estimator to still converge to a valid solution. For example, increasing the value of R_{ww} will cause a more rapid convergence at the cost of a noisier estimate, whereas lessening it will result in slower convergence, but a less noisy estimate, provided that the reduction in R_{ww} is not sufficient to expose the model's limitations. The proper selection and manipulation of these input covariances is sometimes referred to as "tuning" the filter.

Another asset of the KF is that new parameters can be "augmented" into the processing scheme. As long as these parameters are observable, i.e., as long as they impact the data, the KF will estimate them jointly along with the other parameters in the state vector. This means that model parameters that are poorly known or actually unknown can be included in the state vector, thereby lessening, or indeed eliminating, the mismatch problem.

The innovation sequence is another valuable part of the KF. It provides a direct evaluation of the "goodness" of the model. That is, if the model were perfect, the innovation sequence would be zero mean and have a white spectrum. This means that the model contains components associated with all of the observables [6] in the data, so that it is better at "capturing" the relevant data in the measurement, causing it to be subtracted in the innovation calculation. Thus, by testing the innovation sequence, even in the case when a solution is achieved, information on the completeness and correctness of the model is provided, and thereby information on how to improve it.

Derivations of the KF, the EKF and numerous examples of their applications (including a complete solution of the above tracking problem) can be found in references [6,7] and numerous examples of tracking studies can be found in [14].

4 • Spatial Processing and Arrays

4.1 Fundamentals

An acoustic array provides a means of controlling the directional properties of acoustic transmission and reception. Here, we consider only receiving arrays, since most transmitting arrays (see Chapter 96) are rather simple in form and, indeed, knowledge of a receiving array allows one to describe most[3] transmit arrays having the same configuration. At its most fundamental level, an array is simply a group of hydrophones (receivers) arranged in a particular configuration for a particular purpose. Its simplest, non-trivial configuration is when the receivers are equally spaced in a line where all of the receivers are the same and have isotropic sensitivity, i.e., they are equally sensitive in all spatial directions. In this case, the length of the array is given by

$$L = d(N-1), \tag{52}$$

[3]A transmit arrays differs from a receive array in that the intensity of the transmitted field can cause a mutual coupling between its transducer elements that can strongly impact its performance. Also, this phenomenon complicates the analysis of such arrays.

where L is referred to as the aperture. N is the number of elements in the array and d is the element spacing. A line array is depicted in Figure 4. The angle θ is the angle of the propagation direction of the incoming plane acoustic wave as measured from the normal to the line of the array or the "broadside" direction. The receiver outputs are voltages that are proportional to the instantaneous acoustic signal (pressure). In a conventional array processor, the signals from the hydrophones are weighted and summed, producing the array output. This summation is described by

$$p(\theta, t) = \sum_{n=0}^{N-1} a_n s_n(\theta, t), \tag{53}$$

with the a_n being the weights. Referring again to Figure 4, and assuming a narrow-band signal, we write the signal at the nth hydrophone in complex form, as

$$s_n(\theta, t) = a_n e^{i(\omega t + \tau_n)}, \tag{54}$$

where $\omega = 2\pi f$ is the radian frequency and τ_n is the time delay of the wavefront associated with the nth receiver. It is easily seen from Figure 4 that, by defining the delay at the first ($n = 0$) receiver to be zero, and recognizing that the wavefront is normal to the propagation direction, the relative delay associated with the nth receiver is

$$\tau_n = n(d/c) \sin\theta, \tag{55}$$

with c being the speed of sound. Note that the time dependence can be factored out of the sum, thereby allowing us to consider the array output in terms of its *spatial* processing performance only. In the case of a single-frequency signal, such as we have here, it is convenient to write the time delay in terms of its associated phase shift. Defining the wavenumber as $k = 2\pi f/c$,

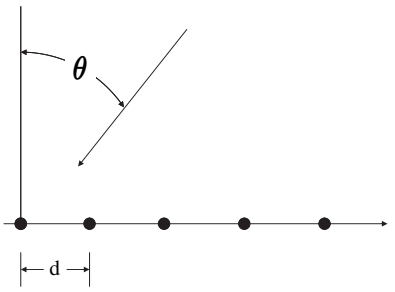


FIGURE 4 Configuration for a line array.

and ignoring the time dependence, we see that the phase shift at the nth receiver is given by $nkdsin\theta$. Equation (53) now becomes

$$p(\theta) = \sum_{n=0}^{N-1} a_n e^{inkd\sin\theta}. \tag{56}$$

In the case where all of the elements have directivity $f(\theta)$, this generalizes to

$$p(\theta) = f(\theta) \sum_{n=0}^{N-1} a_n e^{inkd\sin\theta}. \tag{57}$$

Note that, since the directivity is the same for all elements, it factors out of the sum. This is known as the *Product Theorem* [15]. The sum over the isotropic elements is usually referred to as the *array factor*. For the case of all $a_n = 1$, Eq. (56) has the closed form

$$p(\theta) = e^{ik\frac{(N-1)}{2}d\sin\theta}\left[\frac{\sin\left(k(\frac{Nd}{2})\sin\theta\right)}{\sin(k\left(\frac{d}{2}\right)\sin\theta)}\right], \tag{58}$$

which has a magnitude given by

$$|p(\theta)| = \left|\frac{\sin\left(k\left(\frac{Nd}{2}\right)\sin\theta\right)}{\sin(k\left(\frac{d}{2}\right)\sin\theta)}\right|, \tag{59}$$

Equation (59) is plotted in Figure 5 as a function of $\xi = N(k\frac{d}{2})\sin\theta$ for the case of $N = 8$ and $d = \lambda/2$.

A great deal can be learned about the properties of linear arrays from Figure 5. First, we see that the maximum response axis (MRA) of the main lobe of the array is at $\theta = 0$. Also, we see that the level of the main lobe is decreased to the half-power of point (or the *amplitude* is 0.707 of the maximum) at $\xi \approx \pi/2$, or at the angles $\pm\theta'$, where

$$k\left(\frac{Nd}{2}\right)\sin\theta' = \pi/2. \tag{60}$$

Defining the beamwidth of the array as $\Delta\theta = 2\theta'$ we find that, for θ' reasonably small, $\sin\theta' \approx \theta'$, so that

$$\Delta\theta = 2\theta' \approx Nd/\lambda. \tag{61}$$

Making the observation that for N reasonably large (see Eq. 52), $Nd \approx L$, and by defining the *acoustic aperture* as

$$L' = L/\lambda, \tag{62}$$

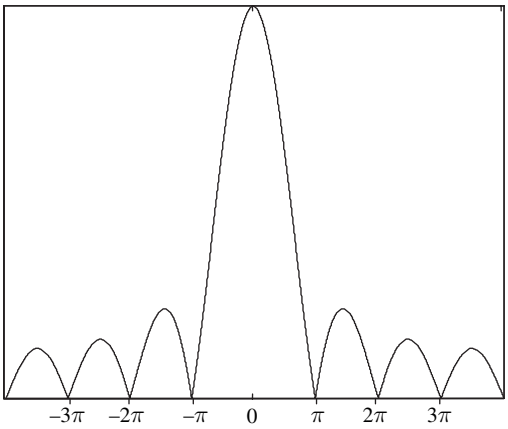


FIGURE 5 Beam pattern of unshaded 8 element line array.

we see that

$$\Delta\theta L' \approx 1. \tag{63}$$

This is the *beamwidth-aperture product*, and shows that the beamwidth of a linear array in radians is, to a good approximation, equal to the reciprocal of the acoustic aperture. This tells us two things. First, to achieve a small beamwidth, we must have a large aperture, and second, a highly useful measure of the aperture of an array is its length in wavelengths.

The discussion up to this point has assumed that the receiver elements were spaced at $d = \lambda/2$. That is, the array has half-wavelength spacing. By inspection of Eq. (56), it is seen that $p(\theta)$ is a periodic function that repeats at values of θ satisfying the relation

$$\sin\theta = \pm m\lambda/d, \tag{64}$$

with m being an integer. Note that, if $d > \lambda/2$, Eq. (56) will "alias," i.e., there will be other maxima appearing in the visible region [4]. For example, if $d/\lambda = 1$, beside the maximum at $\theta = 0$, there are maxima at $\theta \pm 90°$. As d/λ grows larger than unity, more of these maxima will occur. These undesirable extra maxima are

[4]The visible region is where $|k \sin\theta| \leq k$.

traditionally referred to as *grating lobes*. They are avoided when $d/\lambda \leq 1/2$. Indeed, this is the spatial version of Nyquist[5] sampling.

4.2 Beamforming

Returning again to Eq. (56), we see that the maximum of this function occurs when the relative phase between the elements is zero. This means that the main lobe can be steered to the angle θ_m by the subtraction of the phases

$$\delta_n = nkd \sin\theta_m, \tag{65}$$

from the appropriate phase terms in Eq. (56), resulting in

$$p(\theta) = \sum_{n=1}^{N} a_n e^{inkd\sin\theta - i\delta_n} = \sum_{n=1}^{N} a_n e^{inkd(\sin\theta - \sin\theta_m)}. \tag{66}$$

In the case of the line array, beamforming is directly analogous to the discrete Fourier transform. By defining the vector wavenumber as

$$\mathbf{k} = k\sin\theta, \tag{67}$$

Eq. (66) becomes

$$p(\mathbf{k}, \mathbf{k_m}) = \sum_{n=1}^{N} a_n e^{ind\mathbf{k} - ind\mathbf{k_m}}. \tag{68}$$

More generally, by defining the input signal spatial samples as

$$A_n = a_n e^{ind\mathbf{k}}, \tag{69}$$

and making the identification

$$d\mathbf{k_m} = 2\pi\left(\frac{m}{N}\right), \tag{70}$$

Eq. (68) becomes

$$p_m = \sum_{n=1}^{N} A_n e^{-i2\pi\frac{nm}{N}}. \tag{71}$$

Thus, we see that in this context, beamforming can be represented as a discrete Fourier transform of the spatial signal, i.e., the signal sampled over the aperture. This leads to the so-called $k-\omega$ beamformer, which is discussed in detail in Chapter 97. Note that the wavenumber spectrum is equally spaced in $\sin\theta$ and not in θ itself. This requires that an interpolation process be carried out in order to provide beams that are equally spaced in angle [16].

---

[5]The Nyquist criterion is the requirement that a signal must be sampled at a rate that is at least twice the highest frequency. Here the relevant frequency is k, the spatial frequency.

4.3 Array Optimization

Array optimization can be viewed in two rather different ways. One way is to try to obtain sidelobes as low as possible, given a desired beamwidth. The other approach seeks the maximum SNR for a given spatial noise distribution. Both approaches reduce to finding the set of coefficients $\{a_n\}$ in Eq. (56) consistent with this goal. In the underwater acoustics community, this process is sometimes referred to as "shading" the array [17,18]. In the first approach, for the case of a spatially uniform noise field, this usually results in a set of "tapered" coefficients, i.e., large in the center of the aperture and monotonically decreasing toward the ends of the aperture. As a general rule, the lower the sidelobes the wider the beamwidth for a given aperture. This is consistent with Eq. (63), since tapered shading will result in a shorter effective value for L'.

The second approach, seeking the maximum SNR, entails expressing G as a function of the shading coefficients $\{a_n\}$, and finding the set of coefficients that maximize it. That is, we maximize

$$\mathrm{SNR}_{\mathrm{out}} = G(\{a_n\}) \times \mathrm{SNR}_{\mathrm{in}}. \tag{72}$$

The output SNR of the array can be expressed as

$$\mathrm{SNR}_{\mathrm{out}} = E\left[\frac{\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m \int s_n^* s_m d\Omega}{\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m \int \eta_n^* \eta_m d\Omega}\right]. \tag{73}$$

Here s_n and η_n are the signal and noise amplitudes at the nth receiver, respectively, and $*$ denotes the complex conjugate. The integral in the denominator is the noise covariance matrix, R, where

$$R_{nm} = E\left\{\int \eta_n^* \eta_m d\Omega\right\}. \tag{74}$$

Now consider the case for the uniformly spaced line array for spatially uniform uncorrelated noise, i.e., spatially white noise. Since the noise is uncorrelated, the noise covariance matrix reduces to a diagonal, where each term on the diagonal is simply the power on a single element. That is

$$R_{nm} = P\delta_{nm}, \tag{75}$$

so that the denominator of Eq. (73), which is the total noise power, P_T, reduces to

$$P_T = \sum_{n=1}^{N} a_n^2 |\eta|^2 = P\sum_{n=1}^{N} a_n^2. \tag{76}$$

This is simply the weighted sum of the (identical) powers, P, on each receiver. To compute the numerator, we assume that the signal is a plane wave and that the array is steered in its direction. Since it is a plane wave, it is then perfectly correlated over the aperture of the array, so that the total signal power, S_T, is given by

$$S_T = \sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m s_n^* s_m = \sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m S_n = S\sum_{n=1}^{N}\sum_{mM=1}^{N} a_n a_m, \qquad (77)$$

with S being the (identical) signal powers at the each receiver. For this case, then, Eq. (73) reduces to

$$\mathrm{SNR}_{\mathrm{out}} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m}{\sum_{n=1}^{N} a_n^2}\left(\frac{S}{P}\right). \qquad (78)$$

Identifying S/P as SNR_{in}, the SNR at the individual element level, we see from inspection of Eq. (72) that

$$G = \frac{\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m}{\sum_{n=1}^{N} a_n^2} \qquad (79)$$

It is easy to show that this maximizes to the value N when all of the $\{a_n\}$ are equal. In this case, Eq. (79) reduces to

$$G = N. \qquad (80)$$

That is, for spatially white noise, the gain of a uniformly shaded array is simply equal to the number of elements in the array. For this case, i.e., spatially white noise, the gain is called the directivity factor and is usually expressed in dB's as the "Directivity Index" or DI. That is,

$$\mathrm{DI} = 10\log G. \qquad (81)$$

In the case of more general noise fields, the gain can vary widely. For example, consider the existence of a strong interferer in the region of the main lobe. This will greatly reduce the gain. On the other hand, when such an interferer is near a null in the beam pattern, the gain can be quite high.

4.4 Spatial Processing Issues

Referring to Eq. (72), if one desires to maximize detection performance then G, the gain of the array, should be maximized. In the case of multiple signals, however, the *spatial resolution* of the arriving signals' angles becomes the prime importance. There are a class of beamformers that are collectively referred to

"high-resolution" beamformers. One example is the Minimum Variance Distorsionless Response (MVDR) method [19]. This is derived by minimizing the power into the array, over the solid angle of interest, with the constraint that the response at the MRA is fixed. This means that the power into the array is minimized everywhere except at the MRA. The optimal weights for the MVDR can be expressed concisely by defining the signal matrix $S = [s_1, s_2, \cdots, s_N]^T$. The optimal weight vector is then given by

$$\mathbf{a} = \frac{R^{-1}S}{S^* R^{-1} S}. \tag{82}$$

High-resolution beamformers are generally not optimal in a SNR sense, but can provide the ability to resolve sources that are closely spaced in angle, where the conventional beamformer would fail. A complete discussion of these types of beamformers is found in [20,21]. Note that if one has knowledge of the noise covariance as a function of time, Eq. (82) provides a means of adapting the weights to maintain performance. Because of this, the MVDR is sometimes referred to as an adaptive beamformer. It is not unique in this sense, however, since any covariance-based algorithm that allows adaptation of the weights leads to an adaptive array.

An example of the role of arrays in spatial estimation is the bearing estimation problem. In the passive sonar problem, the estimation of the bearing of a source is a task often required of a towed line array. Assuming that range of the source is large as compared to the aperture of the array, it can be assumed that the arriving signal is a plane wave. Also, to a very good approximation, this can be considered to be a two-dimensional problem, as depicted in Figure 4. The problem then is to estimate the angle θ in Eq. (56). For the narrow band case, this can be done by observing that the maximum likelihood estimate of the frequency is found from the peak of the periodogram $I(f)$, where the periodogram is defined as the normalized magnitude squared of the frequency spectrum [8], which is given by,

$$I(f) = \left(\frac{1}{N}\right) \left| \sum_{n=0}^{N-1} a_n e^{inkd \sin\theta} \right|^2. \tag{83}$$

Since the bearing is related to the frequency through the phase term, i.e.,

$$kd \sin\theta = 2\pi (f/c) d \sin\theta, \tag{84}$$

we can use the maximum likelihood invariance[6] theorem [22] to conclude that the maximum of $I(f)$ is the estimator for the bearing angle θ. Although we do not derive it here, it is of interest to look at the CRLB for the bearing estimate. This is given in [8] as

$$\sigma_\theta^2 = \frac{3}{(\mathrm{SNR})\pi^2 N \frac{N+1}{N-1}(L' \sin\theta)^2}, \tag{85}$$

which, for N reasonably large, is well approximated by

$$\sigma_\theta^2 = \frac{3}{(\mathrm{SNR})\pi^2 N(L' \sin\theta)^2}. \tag{86}$$

Note that the acoustic aperture L' (see Eq. 62) appears in the denominator. As expected, the CRLB on the estimate of θ decreases with an increase in L'. Also, it decreases with an increase in the *SNR*, and in N, the number of elements in the array. Recall also, that the beamwidth of the array is well approximated by the reciprocal of the acoustic aperture, so we see that the variance on the estimate of the bearing decreases with the beamwidth, as expected. The term $L'sin\theta$ is usually referred to as the *projected aperture*. Thus, the CRLB on the estimate of θ *increases* as the bearing moves away from broadside, since the projected aperture *decreases*.

4.5 Three-Dimensional Arrays

Although the line array plays a major role in underwater signal processing, three-dimensional arrays can arise in many applications. For example, arrays may be mounted on the hulls of Autonomous Undersea Vehicles, or AUV's leading to conformal arrays. Submarines and surface ships carry spherical and cylindrical arrays. Although the mathematical modeling of such arrays is easily obtained from the generalization of the previous discussion of the line array, the optimization of such arrays becomes much more problematical, and usually must be treated numerically.

The general expression for a three-dimensional array, in a standard polar coordinate system, is given by

$$p(\theta, \phi) = \sum_{n=1}^{N} a_n e^{i\mathbf{k}\cdot\mathbf{r_n}}. \tag{87}$$

---

[6]This theorem states that given an invertible relation between two variables, say $y = f(x)$, the maximum likelihood estimate of x, designated as $\hat{x}$, can be found by finding the maximum likelihood estimate of y, designated as $\hat{y}$, and solving the relationship for $\hat{x}$. That is, $\hat{y} = f(\hat{x})$.

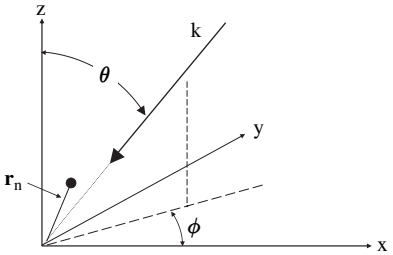


FIGURE 6 Coordinates for a three-dimensional array.

Referring to Figure 6, $\mathbf{k}$ is the wave vector of the incoming plane wave signal. It points in the direction of propagation and has the magnitude $|\mathbf{k}| = k = 2\pi/\lambda$. θ and ϕ are the polar angles of $\mathbf{k}$ and $\mathbf{r_n}$ is the radial coordinate of the nth isotropic receiver element.

To see the structure of the phase term in Cartesian coordinates, it is useful to expand it as follows.

$$\mathbf{k} \cdot \mathbf{r_n} = k[x_n \sin\theta\cos\phi + y_n \sin\theta \sin\phi + z_n cos\theta]. \tag{88}$$

Observe that for a line array on the x-axis, y_n, z_n, and ϕ are all zero and Eq. (88) reduces to

$$\mathbf{k} \cdot \mathbf{r_n} = kx_n \sin\theta, \tag{89}$$

which is the phase term for the line array pattern in Eq. (56) for $x_n = nd$. Note that an element directivity can be included in Eq. (87), just as in the line array case, but the product theorem no longer holds, since these directivities will not all be in the same rotational positions, except for the case of a planar array, where all the z_n are zero.

One case of interest is the *conformal* array, i.e., an array that conforms to a surface. By defining the *Projected Planar Array* (PPA) to be the array that results by placing one axis of the coordinate system, say the z axis, along the maximum of the main lobe, and setting the z coordinates of all the elements to zero, some insight into the performance of the associated conformal array can be obtained. First, it can be shown [23] that the main lobe of both the conformal array and its associated PPA are approximately the same. Second, the apertures of the PPA are inversely related to the beamwidths in the $x-z$ and $y-z$ planes of the PPA. However, the sidelobe behavior of the conformal array is generally worse than that of the associated PPA. By worse we mean that, for a given set of shading coefficients, the sidelobes of the conformal array are higher than those of the PPA, and the difference increases as the observation angle moves away from the MRA.

Examples of beamforming for three-dimensional arrays are given in [23,24,25].

References

1. S. M. Kay, *Fundamentals of Statistical Signal Processing; Vol II, Detection Theory*, Prentice-Hall, New Jersey, 1998.
2. L. L. Scharf, *Statistical Signal Processing; Detection, Estimation, and Time Series Analysis*, Prentice-Hall, New Jersey, 1993.
3. G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.
4. E. J. Sullivan and David Middleton, "Estimation and detection issues in matched-field processing," *J. Oceanic Eng.*, 18, 3, July 1993.
5. A. B. Baggeroer, W. A. Kuperman and H. Schmidt, "Matched-field processing: Source localization in correlated noise as an optimum parameter estimation problem," *J. Acoust. Soc. Am.*, 83, 2, 571–587, 1988.
6. J. V. Candy, *Signal Processing:The Model-Based Approach*, McGraw-Hill International Editions, 1986.
7. J. V. Candy, *Model-Based Signal Processing*, John Wiley & Sons, 2006.
8. E. H. Moore, "On the reciprocal of the general algebraic matrix," *Bull. Am. Math. Soc.*, 26, 394–395 1920.
9. R. Penrose, "A Generalized Inverse for Matrices," *Proc. Cambridge Phil. Soc.*, 51, 406-413, 1955.
10. S. M. Kay, *Fundamentals of Statistical Signal Processing; Estimation Theory*, Addison-Wesley, 1991.
11. R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, 82, 1960.
12. V. J. Aidala and S. M. Hammel, "Utilization of modified polar coordinates for bearingd-only tracking," *IEEE Trans. Autom. Contr.*, 28, 283–294, 1983.
13. A. Tolstoy, *Matched Field Processing for Underwater Acoustics*, World Scientific, 1991.
14. Y. Bar-Shalom and X. Li, *Estimation and Tracking: Principles, Techniques and Software*, Norwood MA: Artech House, 1993.
15. R. J. Urick, *Principles of Underwater Sound in the Sea*, McGraw-Hill, New York, 1975.
16. B. Maranda, "Efficient digital beamforming in the frequency domain," *J. Acoust. Soc. Am., 86 (5), 1989.*
17. W. S. Burdic, *Underwater Acoustic Systems Analysis*, Prentice Hall, New Jersey, 1991.
18. T. T. Taylor, "Design of line source antennas for narrow beamwidth and low side lobes," *IRE Trans.*, AP-3,1955.
19. J. Capon, R. J. Greenfield and R. J. Kolker, "Multidimensional maximum-likelihood processing of a large aperture seismic array," *Proc. IEEE*, 55 (2), 1967.
20. R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*, John Wiley, New York, 1980.
21. H. L. van Trees, *Optimum Array Processing*, John Wiley, New York, 2002.

22. P. W. Zehna, "Invariance of maximum likelihood estimators," *Ann. Math. Stat.*, 37 (3), 1, 1966.
23. E. J. Sullivan, "The Sidelobe Behavior of Conformal Arrays", *J. Acoust. Soc. of Am.*, 71 (2), February, 1982.
24. E. J. Sullivan, "Amplitude Shading of Irregular Acoustic Arrays", *J. Acoust. Soc. of Am.*, 63 (6), June, 1978.
25. J. G. Kelly, E. J. Sullivan and J. Salisbury, Comments on "Conformal Array Beam Patterns and Directivity Indices" [J. Acoustical Society of America, 63, 841–847 (1978)] *J. Acoust. Soc. of Am.*, 65, (2), February, 1979.

95

Inverse Problems in Underwater Acoustics

N. Ross Chapman

University of Victoria, Victoria BC, Canada

1 • Introduction

Inversion is the process of inferring information about a physical system from measurements of physical quantities that are obtained from an interaction with the system. In underwater acoustics, the physical system is generally some aspect of the ocean environment, including the ocean bottom, and/or the experimental system of sources and receivers. This general definition for the inversion process, although appealing simple, implies that we can infer properties of the real ocean environment. Unfortunately, this is not the case. Instead, we are only able to infer the parameters of a model that we design to describe the physical system. The inversion is constrained by the model, which may be inaccurate or incomplete, and by the data, which contain noise.

The formal definition of an inverse problem can be developed in terms of the relationship between the model parameters, $\boldsymbol{m} = [m_1, m_2, \ldots m_n]^T$, and the data, $\boldsymbol{d} = [d_1, d_2, \ldots d_k]^T$. In underwater acoustics, the model parameters may be the sound speeds, densities and attenuations of a geoacoustic model of the ocean bottom, and the data of interest are measurements of the acoustic field at sensors in the water (Figure 1). Consequently,

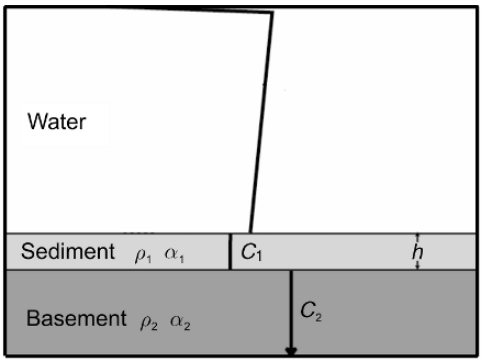


FIGURE 1 An example of a simple geoacoustic model consisting of a water layer, a single sediment layer of thickness h, and a basement half-space. The model describes the variation with depth of the sound speed, c, attenuation, α, and density, ρ, in the sediment.

the data and model are related through the wave equation that describes acoustic propagation. The expression

$$\boldsymbol{d}_0 = F(\boldsymbol{m}) \tag{1}$$

describes the forward problem that predicts the data that would be observed in an ideal, perfectly accurate experiment in which the environment is described by a specific set of model parameters. This problem has a unique and stable solution for the field at a receiver system in the ocean, and there are many numerical recipes for generating highly accurate solutions. The inverse problem of inferring the set of model parameters that generated the observed data is expressed by the relationship

$$\boldsymbol{m} = F^{-1}(\boldsymbol{d}). \tag{2}$$

This problem is generally very difficult to solve, assuming that a solution exists. Existence of the solution is addressed by constructing a model that provides an adequate fit to the data, within some known uncertainty. However, the solution is nonunique. Due to incomplete or inaccurate data, there are many models that will fit the data, and the solution is generally unstable since small errors in the data can lead to large changes in the estimated model parameters.

All measured data contain uncertainty or errors that we can describe as additive noise:

$$\boldsymbol{d} = \boldsymbol{d}_0 + \boldsymbol{n}, \tag{3}$$

where $\boldsymbol{d}_0$ is the "accurate" data according to (1) and $\boldsymbol{n}$ is the noise. Errors in the data arise from two sources: measurement errors due for example to inaccurate readings or ambient noise, and theory or model errors due to inaccurate or incomplete parameterization of the physical system, or approximations in the physics of the forward propagation problem. The data errors are not known explicitly, but instead, we can assume that $\boldsymbol{d}$ is a random variable and assign known statistical distributions for the errors. Usually, a Gaussian distribution is assumed. The theory errors are more difficult to estimate, and for some problems, they can be the dominant source of uncertainty in the inversion.

An important distinction is whether the inverse problem is linear or nonlinear. Although many of the practical inverse problems in underwater acoustics are nonlinear, in some cases, the problems are only weakly nonlinear or can be linearized locally. Linear methods are very appealing because (1) a complete analytical theory exists for linear inverse problems with Gaussian errors and (2) linear methods are computationally much faster. Nonlinear inverse problems are usually cast in terms of optimization problems, employing numerical global search algorithms and importance sampling techniques to explore the model parameter space. In general, the model parameter space is multi-dimensional and characterized by many local minima. Especially for noisy data, the inversion process often becomes trapped in one of the local solutions.

In underwater acoustics, we are primarily interested in using the amplitude and phase information in pressure field data. However, in some instances, it may be more practical to use quantities derived from the pressure field, especially if this leads to a linearized inverse problem. Some examples of these quantities include intensity or energy density versus range (transmission loss); wavenumbers of propagating modes; ocean bottom reflection coefficients or bottom loss; dispersion relationships for propagating modes; and travel time of transmitted signals. We first describe linear inverse problems and the methods for solving them, and then discuss the complete solution for inverting acoustic field data that involve nonlinear methods.

2 • Linear Inverse Problems

In linear inverse problems, the functional relationship between the data and the model described in equation (1) is replaced by a direct numerical relationship [1,2]:

$$\boldsymbol{d} = G\boldsymbol{m}. \tag{4}$$

The rows of the $N \times M$ matrix G consist of linear equations linking the data and the model parameters, and containing all the physics of the forward problem and the geometry of the experimental arrangement. The system of equations (4) has a unique and exact solution

$$G^{-1}\boldsymbol{d} = \boldsymbol{m} \tag{5}$$

if the matrix G is nonsingular. This is known as the even-determined case, in which $N = M$ and the matrix is full rank, i.e. the equations are all linearly independent. Physically, this implies that the experimental data contain information about all the model parameters. If there are more data than model parameters, i.e. the matrix G has more linearly independent rows than columns, the problem is over determined. In this case, there is no solution that fits noisy data uniquely, and the challenge is to determine a model that fits the data well in some sense. If the data do not provide sufficient information to determine all the model parameters, the problem is said to be underdetermined and there is no unique solution. In addition, all these types of linear inverse problems can also be ill-conditioned, leading to unstable solutions. Ill-conditioning can arise if several of the data contain similar information about the model parameters. The challenge is to develop reasonable solutions for these types of problems.

Nonuniqueness can be addressed in overdetermined linear inverse problems by solving for the maximum likelihood solution. This solution maximizes the likelihood of a model given the measured data. Assuming Gaussian distributed errors, the likelihood function, $L(\boldsymbol{m})$, is expressed in terms of the misfit between measured and modeled data, $\boldsymbol{d} - G\boldsymbol{m}$:

$$L(\boldsymbol{m}) \propto \exp[-1/2(\boldsymbol{d} - G\boldsymbol{m})^T C_d^{-1}(\boldsymbol{d} - G\boldsymbol{m})]. \tag{6}$$

Here, C sub small d is the covariance of (bold face n) in equation (3). The most likely solution is obtained by minimizing the misfit function,

$$E = (\boldsymbol{d} - G\boldsymbol{m})^T C_d^{-1}(\boldsymbol{d} - G\boldsymbol{m}), \text{ is } \boldsymbol{m} = (G^T C_d^{-1} G)^{-1} G^T C_d^{-1} \boldsymbol{d}.$$

For the case of independent data errors, this solution is equivalent to the least squares solution.

For ill-conditioned inverse problems, the solution can be stabilized using the method of singular value decomposition (SVD) [1,2]. In this approach, the $N \times M$ matrix G is decomposed by SVD into an $M \times M$ matrix that has a set of eigenvalues (the singular values) that indicates the amount of linearly independent information about the model parameters in G. Small singular values

correspond to linear combinations of model parameters that are most poorly determined by the data. The SVD solution excludes all but the largest singular values.

An alternative method to truncation is the regularized inversion, which instead adds an additional term to the data misfit function that describes some form of prior knowledge, $E = (\boldsymbol{d} - G\boldsymbol{m})^T C_d^{-1}(\boldsymbol{d} - G\boldsymbol{m}) + \mu(\boldsymbol{m} - \hat{m})^T H(\boldsymbol{m} - \hat{m})$. Here $\hat{m}$ is the prior knowledge, H is a regularization matrix, and μ is a trade-off parameters. The solution attempts to achieve a balance between using the information in the data relative to that in the prior knowledge or regularizing term by adjusting the value of μ. A very simple and widely used implementation of regularized inversion is diagonal loading for stabilizing least squares inversion problems. In this case, $\hat{m} = 0, H = I$, and μ is small compared with the signal power. If prior estimates are known, these can be applied explicitly. Regularizations based on derivatives of the model parameters (gradients) are also used to generate minimum structure or smoothest solutions [3].

3 • Linearized Inversion

In the conventional procedure, we begin by linearizing in the vicinity of an arbitrary reference model $\boldsymbol{m^0}$, assuming that the unknown model is related by a small perturbation $\delta\boldsymbol{m}$ to the reference model. From equation (1), the data are related to the reference model by

$$\boldsymbol{d} = F(\boldsymbol{m^0}) + \frac{\partial F(\boldsymbol{m^0})}{\partial \boldsymbol{m}}\delta\boldsymbol{m}, \tag{7}$$

to first order in $\delta\mathbf{m}$. This leads to the linearized relationship

$$\boldsymbol{d} = F(\boldsymbol{m^0}) + G\boldsymbol{m} - G\boldsymbol{m^0}, \tag{8}$$

where G is an N x M matrix of first derivatives, $\frac{\partial F_i(m^0)}{\partial m_j}$. The model is then computed iteratively to obtain a final model $\boldsymbol{m^f}$ using the relation

$$\boldsymbol{m^{i=1}} = \boldsymbol{m}^i + G^{-1}(\boldsymbol{d} - F(\boldsymbol{m}^i)). \tag{9}$$

The second derivatives $\frac{\partial^2 F_i(m^f)}{\partial m_j^2}$ are used to characterize all the models in the neighbourhood of $\boldsymbol{m}^{\mathrm{f}}$ in order to provide an uncertainty for the solution. This approach has the advantage that computationally fast linear methods can be applied, but there are

many pitfalls. The most serious defect is that one is never sure that the final model is independent of the reference model. The inversion may in fact not converge if the starting model is not close to the solution, or more likely, it may converge to a local minimum. Another serious defect is that because F is nonlinear, it can be very misleading to use only the parameter space near $\boldsymbol{m}^{\mathrm{f}}$ to characterize the solution.

Although the pressure field is not related linearly to environmental model parameters such as sound speed, density or water depth, changes in sound speed can be related linearly to changes in quantities that can be derived from measurements of pressure. There are many examples of applications of this approach. Frisk et al. developed an elegant method for estimating sound speed profiles in marine sediments by linearizing the relationship between changes in the horizontal wavenumbers of propagating modes and changes in the sound speed [4,5]. Dosso et al. have implemented a very efficient algorithm for array element localization, based on linearizing the travel time of signals received at the array [3,6]. Perhaps, the most striking example of linearized inversion is ocean acoustic tomography. The basic technique is described extensively in [2], and specific examples of inversions are found in the works by Worcester et al. [7,8] for acoustic thermometry in the Pacific, Mikhailevsky et al. [9] for similar applications in the Arctic, and Skarsoulis et al. [10,11].

4 • Nonlinear Inverse Problems

Methods for solving nonlinear inverse problems in underwater acoustics have been implemented within the last decade and are in widespread use in applications with experimental data. They are generally formulated as optimization problems that require numerical search algorithms to investigate multi-dimensional model parameter spaces in order to determine the model $\boldsymbol{m}$ that optimizes an objective function, $E(\boldsymbol{m})$. One of the most widely used methods is based on matched field processing (MFP). MFP compares measured acoustic fields at an array of sensors with modelled or replica acoustic fields that are calculated for a specific set of environmental model parameters. An accurate comparison is obtained only if the physics of sound propagation in the environment is well understood and the geoacoustic and oceanographic characteristics are well known. The information about the environmental model parameters, and also the experimental geometry, is contained primarily in the spatial phase relationships at the array. It is useful to retain this very physical

picture of MFP in order to understand the more formal development that follows. The concept of MFP has been known for a very long time, from the first simple experiments of Parvelescu and others [12] at Hudson laboratories in the mid- 1960s and the first formal paper by Homer Bucker in 1976 [13]. However, MFP could not be applied effectively as an inversion method until modern numerical propagation models and fast computers with large storage capacity became available.

We outline first the formal approach for the solution to the inverse problem that is based on the statistical formulation of Tarantola [14–16]. The solution is given in terms of the *a posteriori* distribution or (normalized) probability density (PPD) in model parameter space. From Bayes' rule for conditional probabilities, we can derive a relationship for the PPD:

$$P(\boldsymbol{m}\,|\,\boldsymbol{d}) = P(\boldsymbol{d}\,|\,\boldsymbol{m})P(\boldsymbol{m})/P(\boldsymbol{d}) \tag{10}$$

where $P(\boldsymbol{m}\,|\,\boldsymbol{d})$ is the conditional probability density function (PDF) of $\boldsymbol{m}$ given $\boldsymbol{d}$ (the PPD), $P(\boldsymbol{d}\,|\,\boldsymbol{m})$ is the conditional PDF of $\boldsymbol{d}$ given $\boldsymbol{m}$, $P(\boldsymbol{m})$ is the PDF of $\boldsymbol{m}$, the prior information independent of the data, and $P(\boldsymbol{d})$ is the PDF of $\boldsymbol{d}$. If $\boldsymbol{d}$ represents measured data, $P(\boldsymbol{d}\,|\,\boldsymbol{m})$ is interpreted as the likelihood function of $\boldsymbol{m}$, and equation (10) becomes

$$P(\boldsymbol{m}\,|\,\boldsymbol{d}) \propto L(\boldsymbol{m}, \boldsymbol{d})P(\boldsymbol{m}). \tag{11}$$

The likelihood function can be expressed as $L(\boldsymbol{m}, \boldsymbol{d}) \propto \exp[-E(\boldsymbol{m}, \boldsymbol{d})]$ where E is the objective function or measure of data misfit,

$$E(\boldsymbol{m}, \boldsymbol{d}) = \frac{1}{2}\sum (d_i - d_i(\boldsymbol{m}))^2/\sigma_i^2, \tag{12}$$

with σ_i the standard deviation for d_i. Although the PPD provides the complete solution, it is more practical to represent the solution in terms of the maximum *a posteriori* value (MAP), or by various moments of the PPD, such as the mean, covariance or marginal probability density for each parameter. Numerical methods such as Gibbs sampling have been implemented for efficient evaluation of the various moments [17].

5 • Matched Field Inversion

Methods that are designed to determine the MAP value are equivalent to the optimization techniques such as matched field inversion (MFI) [18,19]. MFI algorithms for estimating parameters of geoacoustic models consist of four basic components:

- *Prior knowledge*, $P(\boldsymbol{m})$, expressed as limits or bounds for the parameter values of a geoacoustic model.
- A *wave propagation model* that describes the forward problem for each model of the environment. Inversions have been reported using ray, normal mode and parabolic equation codes.
- An *objective function* for assessment of the candidate models of the environment according to a selection criterion based on the comparison of measured and calculated acoustic fields.
- An *efficient search algorithm* for investigating the multi-dimensional model parameter space.

The bounds for the model parameters are determined from available information about the experimental location, including ground truth measurements such as sediment cores, and high-resolution seismic survey data. In the absence of any bias, the distribution of allowed values is uniform. It is very important to realize that the inversion results are constrained by the form of the geoacoustic model. The model usually consists of profiles of the sound speed, attenuations and densities of sediment and basement layers.

The objective function is developed using a linear model of the signal in which the calculated field, q, is defined in terms of the transfer function of the acoustic waveguide, w, and the source signal, s. For a single frequency, the calculated field is given by $q(\omega, \boldsymbol{m}) = w(\omega, \boldsymbol{m}s(\omega)$, and the measured data are related to q according to $d(\omega) = q(\omega, \boldsymbol{m}) + n(\omega)$. As stated previously, the noise includes measurement and theory errors. In the assumption that the errors are Gaussian distributed with zero mean and uncorrelated, it can be shown from equation (12) that the objective function is given by [20]:

$$E(\omega, \boldsymbol{m}) = 1 - \frac{q^*(\omega, \boldsymbol{m}) R q(\omega, \boldsymbol{m})}{|d|^2 |q(\omega, \boldsymbol{m})|^2}, \tag{13}$$

where $R = dd^*$ is the cross-spectral matrix of the measured data and $*$ denotes the conjugate transpose operation. The second term in (13) can be recognized as the normalized Bartlett matched field processor for a single frequency, $B(\omega, \boldsymbol{m})$. For multiple frequencies, the objective function can be constructed as an incoherent sum of expressions of single frequency components,

$$E(\omega, \boldsymbol{m}) = 1 - \sum_{l=1}^{L} B_l(\omega, \boldsymbol{m}), \tag{14}$$

where L is the number of frequencies. The expressions in equations (13) and (14) are designed to use the spatial phase

relationships at single frequencies. Applications have also been developed to include information in the phase relationships across different frequencies [21,22]. Fully coherent MFP that makes use of spatial and temporal phase relationships has also been implemented [23,24]. The objective functions for these broadband applications are expressed in terms of correlations between measured and modelled signal waveforms. These applications generally employ ray theory propagation models for computational efficiency.

Because the dimensions of the model parameter spaces are generally very large, simple grid searches are not practical. Algorithms based on importance sampling are required [16]. The most widely used of these methods are simulated annealing (SA) and genetic algorithms (GA). SA is based on an analogy with the thermodynamic process of annealing a crystal. There are two algorithms that simulate this process, the Metropolis algorithm and the heat bath algorithm. The former method is in more widespread use in underwater acoustics. It is designed to test possible models to determine whether the objective function is improved or weakened while a control parameter analogous to temperature in the thermal annealing process is reduced. If the cost function is improved, the model is accepted unconditionally, and the search moves to a more promising region of the parameter space. However, if the cost function is weakened, the model is accepted according to the condition

$$\eta \leq \exp[-\Delta E/T], \tag{15}$$

where η is a random number from the interval [0, 1] and T is the control parameter. At high temperatures, the probability of accepting models that do not improve the cost function is high, and in this way, search process can escape from local minima and explore other regions of the parameter space. The efficiency of the search process depends upon the choice of model parameters, and it is usually more efficient to reparameterize the original system to obtain an orthogonal set [25,26].

GAs use a population of Q models that are coded to represent genetic information. The population evolves from generation to generation in order to maximize the fitness. The basic steps in GA are (1) coding of the model parameters to represent the solutions within the population; (2) selection to determine the subset of models from one generation to serve as parents for the next generation; (3) cross-over to recombine genetic information of a pair of parents; (4) mutation to randomly alter the genetic information and (5) replacement to determine how the offspring replace the parents for the next generation. The SAGA inversion package

developed by Gerstoft introduced genetic algorithm inversion methods in ocean acoustics and remains in widespread use [27].

Hybrid methods are designed to retain some memory of good regions of the parameter space in order to improve the efficiency and speed of random search processes. This approach provides the benefits of local gradient-based algorithms that are computationally fast but are prone to get caught in local minima, with global search techniques that can explore all regions of the parameter space. One of the most effective algorithms is Adaptive Simplex Simulated Annealing (ASSA) [28]. This algorithm combines the downhill simplex method with simulated annealing and provides three significant advantages:

- Instead of a single 'walker' in the parameter space, ASSA allows a simplex of size $M+1$, where M is the number of model parameters, to navigate the parameter space.
- The simplex always retains the best M models that have been found.
- The simplex is effectively a shape that adapts to the topology of features in the parameter space such as valleys or ridges that are related to model parameter correlations. Consequently, the search is more efficient in the original model parameterization.

A very useful measure of the quality of the estimated values in an MF inversion can be obtained by monitoring the objective function values that are tested during the search process. An example of scatterplots of the objective function versus parameter values derived for two parameters in an ASSA inversion is shown in Figure 2. Parameters that are well estimated generally have very tight distributions about the minimum value (left panel),

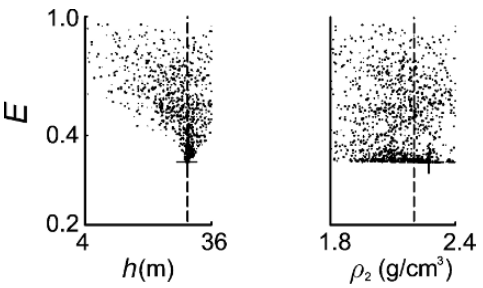


FIGURE 2 *Scatterplots* of the objective function for models that were tested in the search process. The *leftpanel* shows a well-estimated parameter, indicated by the high density of points near the minimum value. The *rightpanel* shows the result for a parameter of lower sensitivity that is not well estimated.

whereas parameters that are not well estimated have flat distributions (right panel). Although the scatterplots do not provide a statistical measure of the uncertainty of the estimate, the qualitative information about parameter sensitivity is very useful.

In general, there is a hierarchy of parameter sensitivities that is often strongly related to the design of the experiment. Acoustic field data are highly sensitive to the water depth and the sound speed profiles in the bottom [29]. However, the depth to which the profile can be estimated accurately in the sediment depends on factors such as sound frequency and experimental geometry. Sediment density and attenuation are less sensitive but can be estimated in carefully designed experiments. It is generally true that the geometrical parameters of the experiment such as source range and depth are highly sensitive, and in some cases these parameters are inverted jointly. However, it is very important to note that there is strong correlation between model parameters, for instance source range and water depth, and it is a wise practice in experiments to obtain accurate independent measures of geometrical parameters.

Nonlinear methods based on global search techniques have also been implemented to invert quantities derived from the acoustic field data such as modal dispersion relationships [30] and reflection coefficients [31].

References

1. W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press, Orlando, 1987.
2. W. Munk, P. Worcester and C. Wunsch, *Ocean Acoustic Tomography*, Cambridge University Press, 1995.
3. S.E. Dosso, M.R. Fallat, B.J. Sotorin and J.L. Newton, Array element localization for horizontal arrays via Occam's inversion, J. Acoust. Soc. Am., 104, 846–859, 1998.
4. S.D. Rajan, J.F. Lynch and G.V. Frisk, Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water, J. Acoust. Soc. Am., 82, 998–1017, 1987.
5. G.V. Frisk, J.F. Lynch and S.D. Rajan, Determination of compressional wave speed profiles using modal inversion techniques in a range-dependent environment in Nantucket Sound, J. Acoust. Soc. Am., 86, 1928–1938, 1989.
6. S.E. Dosso and N.E. Collison, Acoustic tracking of a freely drifting field of sonobuoys, J. Acoust. Soc. Am., 111, 2166–2177, 2002.
7. P.F. Worcester and the ATOC (Acoustic thermometry for Ocean Climate) team, A test of basin-scale acoustic thermometry using a large aperture vertical array at 3250-km range in the eastern North Pacific, J. Acoust. Soc. Am., 105, 3185–3201, 1999.

8. P.N. Mikhalevsky, A.N. Gravrilov and A.B. Baggeroer, The transarctic acoustic propagation experiment and climate monitoring in the Arctic, IEEE J. Oceanic Eng., 24, 183–201, 1999.
9. B.D. Dushaw, Inversion of multimegameter-range acoustic data for ocean temperature, IEEE J. Oceanic Eng., 24, 215–223, 1999.
10. E.K. Skarsoulis, G.A. Athanasoulis and U. Send, Ocean acoustic tomography based on peak arrivals, J. Acoust. Soc. Am., 100, 797–813, 1996.
11. E.K. Skarsoulis, U.Send, G. Piperakis and P. Testor, Acoustic thermometry of the western Mediterranean, J. Acoust. Soc. Am., 116, 790–798, 2004.
12. A. Parvelescu, Signal detection in a multipath medium by M.E.S.S. processing, J.Acoust. Soc. Am., 33, 1674, 1965.
13. H. Bucker, Use of calculated sound fields and matched field processing to locate sound sources in shallow water, J. Acoust. Soc. Am., 59, 368–373, 1976.
14. A. Tarantola, *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam, 1987.
15. M.K. Sen and P.L. Stoffa, *Global Optimization Methods in Geophysical Inversion,* Elsevier, Amsterdam, 1995.
16. M.K. Sen and P.L. Stoffa, Bayesian inference, Gibbs sampler and uncertainty estimation in geophysical inversion, Geophys. Prosp., 44, 313–350, 1996.
17. S.E. Dosso, Quantifying uncertainties in geoacoustic inversion I: A fast Gibbs sampler approach, J. Acoust. Soc. Am., 111, 128–142, 2002.
18. M.D. Collins, W.A. Kuperman and H. Schmidt, Nonlinear inversion for ocean bottom properties, J. Acoust. Soc. Am., 92, 2770–2783, 1992.
19. C.E. Lindsay and N.R. Chapman, Matched field inversion for geoacoustic model parameters using adaptive simulated annealing, IEEE J. Oceanic Eng., 18, 224–231, 1993.
20. P. Gerstoft and C.F. Mecklenbrauker, Ocean acoustic inversion with estimation of a posteriori probability distributions, J.Acoust. Soc. Am., 104, 808–819, 1998.
21. Z.-H. Michaloupoulou, Robust multi-tonal matched field inversion: A coherent approach, J. Acoust. Soc. Am., 104, 163170, 1998.
22. D.P. Knobles, R.A. Koch, L.A. Thompson, K.C. Focke and P.E. Eisman, Broadband sound propagation in shallow water and geoacoustic inversion, J. Acoust. Soc. Am., 113, 205–222, 2003.
23. L. Jaschke and N.R. Chapman, Matched field inversion of broadband data using the freeze bath method, J. Acoust. Soc. Am., 106, 1838–1851, 1999.
24. N.R. Chapman, J. Desert, A. Agarwal, Y. Stephan and X. Demoulin, Estimation of seabed models by inversion of broadband acoustic data, Acta Acustica, 88, 756–759, 2002.
25. M.D. Collins and L. Fishman, Efficient navigation of parameter landscapes, J. Acoust. Soc. Am., 98, 1637–1644, 1995.

26. L.T. Fialkowski, J.F. Lingevitch, J.S. Perkins and M.D. Collins, Geoacoustic inversion using a rotated coordinate system and simulated annealing, IEEE J. Oeanic Eng., 28, 370–379, 2003.
27. P. Gerstoft, Inversion of seismo-acoustic data using genetic algorithms and *a posteriori* probability distributions, J. Acoust. Soc. Am., 95, 770–782, 1994.
28. S.E. Dosso, M.J. Wilmut and A-L. Lapinski, An adaptive hybrid algorithm for geoacoustic inversion, IEEE J. Oceanic Eng., 26, 324–336, 2001.
29. N.R. Chapman, S. Chin-Bing, D. King and R.B. Evans, Benchmarking geoacoustic inversion methods for range-dependent waveguides, IEEE J. Oceanic Eng., 28, 320–330, 2003.
30. G. Potty, J.H. Miller, J. Lynch and K. Smith, "Tomographic Inversion for sediment parameters in shallow water," J. Acoust. Soc. Am., 108, 973–986, 2000.
31. S.E. Dosso and C.W. Holland, Visco-elastic inversion of seabed reflection data. In Press: IEEE J. Oceanic Eng., 2005.

96

Active Sonar

Edmund J. Sullivan

EJS Consultants, Portsmouth, RI, USA

1 • Introduction

An active sonar system is one in which pulses of acoustic energy are launched into the water for the purpose of producing echoes. By examining the echoes of transmitted pulses, it affords the capability of both detecting the presence of and estimating the range, and in certain cases, the bearing, of an underwater target. In its most common arrangement, the transmitter (or projector) and the receiver are colocated. This is known as the monostatic configuration and is depicted in Figure 1. When this is not so, it is known as a bistatic or multistatic configuration. [1] Generally speaking, one desires that the transmitter power be as high as reasonably possible and in most cases, have rather high directivity[1] (See **Statistical Signal Processing**). The receiver is usually in the form of an array, and its configuration may vary widely. The pulse configuration can take several forms, depending on the application. In deep water, reverberation can be low, so that the ambient noise will often be the limiting factor. In this case, increasing the pulse energy is sufficient to overcome the noise. Since there is a power limit to any projector, this requirement can lead to long pulse durations in order to obtain sufficient pulse energy. For single-frequency pulses, i.e., continuous wave (CW)

[1]Although high directivity provides a high sound pressure level (SPL) at the target, it reduces the so-called volume search rate. Thus, its maximum useful value is really tactic dependent.



FIGURE 1 Monostatic sonar. The source and receiver are colocated.

pulses, the range resolution can be low. For example, if the pulse is 0.1 s in length, it has a spatial extent of 150 m, since the speed of sound in water is 1500 m/s. This leads to a *maximum* error in range estimation on the order of 75 m. The actual error will depend upon the signal-to-noise ratio [2]. As will be seen later, this can be improved upon by increasing the bandwidth of the pulse.

In shallow-water scenarios, the problem becomes more complicated because of increased reverberation. Here, the solution does not lie in simply increasing the pulse energy, since this results in a proportional increase in the reverberation, which will eventually replace the ambient noise as the limiting factor. There are two approaches to dealing with reverberation. One is to exploit any Doppler information in the signal. In many cases, the target is moving, thus imposing a Doppler shift on the echo. This Doppler shift can have the effect of moving the echo spectrum into the tail of the reverberation spectrum [3], providing a significant improvement in reverberation performance. In order to resolve the Doppler, a narrow-band signal is required. That is, the bandwidth of the signal must be less than twice the Doppler shift. Thus, if the target is moving too slowly, this approach is not feasible. When Doppler processing against the reverberation is not possible, the second approach, which is to increase the signal bandwidth, is used [4]. Increasing the bandwidth results in a reduction in the spectral density of the reverberation. Associated with this increase in signal bandwidth is an increase in its spatial resolution capability. Beyond a certain point, however, the target return becomes "over-resolved". That is, instead of producing a single high-level detection, the matched filter will produce a set of lower-level detections corresponding to the target highlights [5]. This "energy splitting" can be pronounced when the target is near bow or stern aspect, since this is where the spatial separation of the highlights is greatest.

2 • The Sonar Signal

2.1 The Matched Filter

In **Statistical Signal Processing**, it is shown that for the case of the completely known signal in Gaussian white noise (GWN), the log-likelihood ratio test [6,7] results in the following equation.

$$\frac{1}{\sigma^2}\sum_{n=1}^{N} s_n y_n - \frac{1}{2\sigma^2}\sum_{n=1}^{N} s_n^2 > \gamma. \tag{1}$$

Here, γ is the logarithm of the likelihood ratio, whose value is set by specifying the false alarm probability, σ^2 is the noise power, y_n is the *nth* data sample and s_n is the *nth* signal sample. Denoting n_n as the *nth* noise sample, we have

$$y_n = s_n + n_n. \tag{2}$$

Implicit in these expressions is that we have sampled at a frequency f_s that is *at least* the Nyquist frequency. This means that if we view Equation 1 as a discrete integration, then we must integrate with a sample time increment no greater than

$$\Delta t = 1/f_s = \frac{1}{2B}, \tag{3}$$

where B is the bandwidth of the signal. This insures that N is equal to the number of independent samples in the signal.[2] Equation 1 can now be written as

$$\frac{2B}{\sigma^2}\sum_{n=1}^{N} s_n y_n \Delta t > \frac{B}{\sigma^2}\sum_{n=1}^{N} s_n^2 \Delta t + \eta. \tag{4}$$

Expressing the power spectral density of the white noise as

$$N_0 \equiv \frac{\sigma^2}{B}, \tag{5}$$

rewriting the sums as integrals and recognizing $\sum_{n=1}^{N} s_n^2 \Delta t$ as the signal energy term, we now have

$$\frac{2}{N_0}\int_0^T s_n y_n \mathrm{d}t > \frac{E}{N_0} + \eta, \tag{6}$$

with $T = N\Delta t$.

The integral on the LHS is the correlation between the received data and the signal or "replica", and indeed, the matched filter is sometimes referred to as a "replica correlator". This term is taken as the detection statistic, and maximizing it is equivalent to maximizing the likelihood ratio [6]. By inspection, it can be

[2]The derivation of Equation 1 is based on an assumption of Nyquist sampling and GWN. This allows the pdf of the time series to be represented by a product of Gaussians, since the time samples are independent. At higher sample rates, however, the maximum number of independent samples is still constrained to be 2BT.

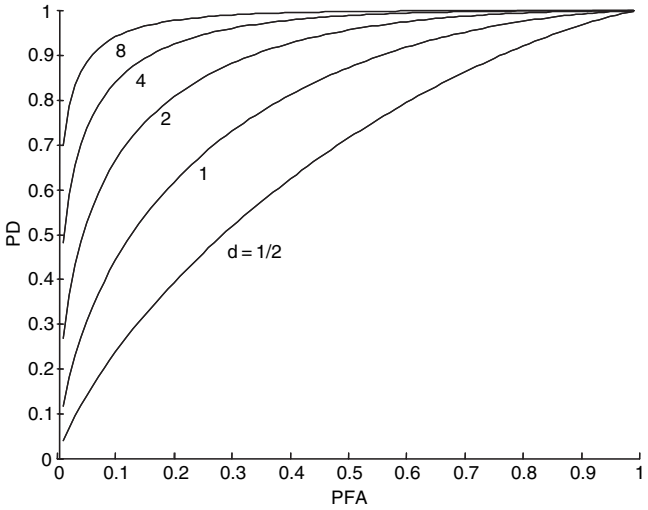


FIGURE 2 Receiver operating characteristics for the case of Gaussian noise.

seen that the maximum of the correlation integral is the signal energy E. This leads to the "detection parameter", given by

$$d = \frac{2E}{N_0}, \tag{7}$$

which is the relevant parameter for the so-called ROC (receiver operating characteristics) [8,9] curves shown in Figure 2.

The detection parameter for the matched filter also leads to an expression for the matched filter gain. Consider a constant-amplitude pulse of length T and amplitude S. Multiplying the numerator and denominator of d by the bandwidth B, Equation 7 becomes

$$d = \frac{|S|^2}{\sigma^2} 2BT = 2BT \times SNR_{\text{in}} = SNR_{\text{out}}. \tag{8}$$

Thus, when the signal is uncorrelated with the noise, the gain of the matched filter is $2BT$, where BT is the *time-bandwidth product*. In fact, $2BT$ is equal to the number of independent samples in the signal [10]. This equation must be interpreted with some caution, since superficially, it appears that increasing the bandwidth of the signal can increase the gain without increasing the pulse energy. But such an increase in bandwidth will produce an associated increase in σ^2, thereby holding SNR_{out} constant. However, in the reverberation-limited case, increasing the signal

bandwidth *can* provide processing gain, since the power spectral density of the reverberation decreases with bandwidth while leaving the signal energy unchanged for a given pulse length.

2.2 The Ambiguity Diagram

The matched filter is an optimum *detector*. However, active sonar is also concerned with range, bearing and Doppler *estimation*. Bearing estimation is primarily a spatial processing task and is dealt with in **Statistical Signal Processing**. Here, the focus is on range estimation and Doppler processing, which is also an estimation process.

As mentioned previously, range estimation performance is related to the bandwidth of the pulse. This can be seen by the following argument. Suppose we desire a range resolution of ΔR meters. This requires that we resolve the pulse to a precision of $\Delta T = \Delta R/c$. For a CW pulse, the time-bandwidth product is approximately unity, so that with the requirement that $\Delta R = 1$ m, we would require $\Delta T = 1/1500$ s, since the required bandwidth of the pulse is 1500 Hz. A pulse of such a short length is highly unrealistic, mainly since its energy would be too low for acceptable detection performance. The high bandwidth, however, is necessary in order to achieve a sufficiently narrow correlation peak from the matched filter to achieve the desired resolution. The obvious solution is to increase the bandwidth of the pulse. A commonly used pulse is the linear FM slide [11]. If, for acceptable detection performance, we desired a pulse length of, say, 0.1 s, an FM slide with a bandwidth of 1500 Hz would then provide the required range resolution while maintaining the pulse energy at the desired level, i.e., it would have a correlation peak approximately equivalent to that of a CW pulse of .00067 s. This increase in bandwidth has effectively "compressed" the pulse, thereby providing the desired range resolution without sacrificing pulse energy.

There is a penalty associated with this type of pulse, however. If one is concerned with Doppler processing, the pulse bandwidth must be of the order of the frequency resolution required. As an example, consider a 1.5 m/s (3 kt) target. Assuming a 5 kHz CW pulse, the maximum Doppler shift one could expect is given by $2f_0v/c = (2 \times 5000 \times 1.5)/1500 = 10$ Hz. Clearly then, this pulse must be longer than $1/20 = 0.05$ s. Obviously, one cannot perform accurate range estimation and Doppler processing with the same pulse form. This example is a special case of a kind of "uncertainty principle" that constitutes a fundamental limitation associated with single-ping active sonar processing.

This uncertainty principle is embodied in the *ambiguity diagram* [12]. It is a generalized representation of the correlation

function of Equation (6). Although the matched filter is implemented as a correlation between the signal and its replica, it can also be thought of as a filter. If the signal is denoted by $y(t)$, then the output at time τ of the matched filter can be represented in the time domain by

$$a(\tau) = \int_0^{t'} y(t')s(\tau - t')\mathrm{d}t'. \tag{9}$$

Now suppose we consider $s(t)$ as being Doppler shifted by f_d. In complex envelope form [7], the signal can now be represented as

$$y_d(t) = y(t)s^{i2\pi f_d t}, \tag{10}$$

and the matched filter output then generalizes to the ambiguity function, given by

$$a(\tau, f_d) = \int_0^{\tau} y(t')s^{i2\pi f_d t}s(\tau - t')dt' \equiv A(\tau, f_d). \tag{11}$$

and is conventionally plotted as in terms of its magnitude squared. An example of an ambiguity diagram is shown in Figure 3. The region labeled "CW" is the case for the 0.1 s CW pulse considered in the previous section. Here, one can see that the ambiguity diagram indicates that the range resolution is poor, but the frequency resolution *and accuracy* are quite good. The region labeled "FM" is the case for the FM slide of the previous section. Here, the pulse length remains at 0.1 s, but the bandwidth

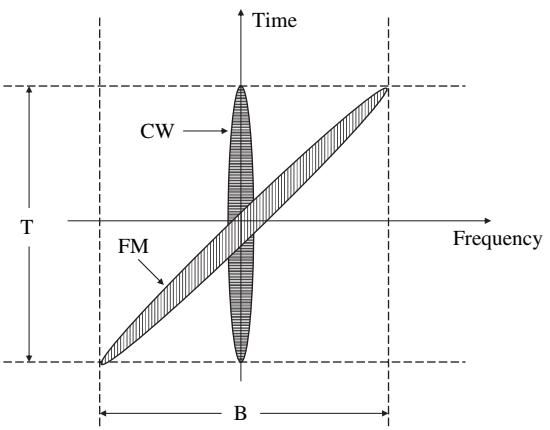


FIGURE 3 The ambiguity diagram showing examples of a CW and an FM pulse.

has been increased to 1500 Hz. In this case, both the frequency and range resolution are good, but the accuracy for both is poor. This is not really important in the case of range estimation, since a high degree of accuracy is not needed. However, in the case of frequency estimation, especially for the case of Doppler processing as we shall see later, good accuracy is also necessary.

3 • Sonar System Design

3.1 Processing Issues

Unlike passive sonar, the active sonar signal configuration is under the control of the user. As a consequence, a great deal of effort has been put into the design of the sonar pulse form [13,14], mainly to obtain particular ambiguity properties. Here, we shall look at the CW and FM pulses only, since they sufficiently exhibit the important properties of the signal for our purposes.

Active sonar is mainly concerned with detection and ranging. That is, both detection and estimation performance are of interest. However, it is not always useful, or even meaningful, to separate the two tasks, since they become intertwined by the role of the sonar. For example, Doppler processing, which is actually estimation, plays a role in the detection process, since it preceeds the detection decision but plays a role in its ultimate success or failure. Another example is the performance of the matched filter. Although it is an optimum detector, it requires a priori knowledge of the signal arrival time. That is, knowledge of the *form* of the signal alone is not sufficient for achieving optimality. However, the arrival time is the parameter of interest in the estimation of range. As a result, range estimation must be carried out by a sequence of detection decisions. Consequently, the processing performance of an active sonar system is really dealing with a mixture of detection and estimation, and this is where the ambiguity diagram plays an important role, since it graphically presents both the detection and estimation behavior of a sonar pulse in a convenient manner.

3.2 Signal Selection

For purposes of discussion, we shall view the signal selection process as a top-down procedure. Beginning with the frequency, it is shown in Reference [1] that, for a given range, there is an optimum frequency. Above this frequency, the absorption loss limits the range. Below the optimum frequency, several parameters contribute to limiting the range, including the *DI* (see **Statistical Signal Processing**), the ambient noise and the bandwidth. Generally speaking, this optimum frequency decreases with range. In the examples to follow, we shall use 5 kHz as the pulse frequency, since it falls within the range of commonly used long-range sonar systems.

Signal energy maximization plays a major role from the outset. This is simply that, as seen in Equation 7, it is often a necessary (but not always sufficient) condition for maximization of detection performance. The next step is determined by the nature of the noise. If the ambient noise is the limiting factor, then pure detection does not require anything beyond the pulse energy to be specified. However, if an estimation of the range is desired, then clearly, the bandwidth plays a critical role, since, as can be seen in Figure 3, it is the major determinant of the range resolution. For most tactical situations, the range *accuracy* will be sufficient, even for unknown Doppler. However, the range *resolution* plays a critical role in the ability to detect in the presence of high reverberation. It should be noted, however, that unknown Doppler results in some performance degradation due to an incorrect replica in the matched filter.

What is clear at this point is that there is no single pulse that can perform all of the tasks equally well. In practice, more than a single pulse is required. A common procedure is the sequential transmission of a high-bandwidth pulse and a low-bandwidth (CW) pulse in order to provide both frequency resolution and range resolution.

4 • Performance Analysis

4.1 The Sonar Equation

The sonar equation is a major tool for sonar system analysis [1]. Fundamentally, it is the modeling of the complete system, source, receiver and propagation channel, in a simple but highly useful representation. For the active sonar case, it is given by

$$DT = SL - 2TL + TS - (NL - DI). \tag{12}$$

The terms in Equation 12 are defined as follows, with all terms being in decibels. DT is the detection threshold, SL is the source level, TS is the target strength, NL is the effective noise level, DI is the directivity index of the receiving array and TL is the transmission loss. In the following, we will be using Equation 12 in the spectral form by expressing the noise term as the power in a 1 Hz band, viz.,

$$DT_s = SL - 2TL + TS - (NL_s - DI), \tag{13}$$

where

$$NL_s = NL - 10\log B \tag{14}$$

where, as before, B is bandwidth. The term $(NL_s - DI)$ is in parenthesis, since its interpretation depends upon the character of

the noise. If the noise is isotropic, its effects can be reduced by increasing the *DI* of the receiver array, where the *DI* is defined as the gain of the receiver array for the case of isotropic noise. However, if the noise is not isotropic, as is certainly true for boundary (surface and bottom) reverberation, the collective term $(NL_s - DI)$ must be replaced by an effective noise or reverberation level and can only be properly computed by using the actual array gain, which requires a priori knowledge of the spatial structure of the noise or reverberation. *DT* is the detection threshold referred to a 1 Hz band. That is

$$DT = 10\log\left(\frac{P}{N_s}\right), \tag{15}$$

where P is the signal power at the receiver. The transmission loss is based on a simple, straight-line model of the propagation and for active sonar, must account for the two-way propagation. As an example, for spherical spreading (as is the case for deep-water propagation), the transmission loss is given by

$$TL = 20\log R + \alpha R, \tag{16}$$

with α being absorption loss in dB/m.

The source level term is defined as follows. For a transmitter power of P_s watts, the acoustic intensity, in W/m^2, at a distance of R meters is defined as

$$I(R) = \frac{P_s}{4\pi R^2} = \frac{|p(R)|^2}{\rho c}. \tag{17}$$

Here, ρ is the density of the medium (in our case water) and c is the speed of sound. The ρc term is referred to as the acoustic impedance. The source level is then defined in terms of the pressure level at distance of 1 m. That is,

$$SL = 10\log\left[\frac{I(1)}{I_{\text{ref}}}\right], \tag{18}$$

with

$$I_{\text{ref}} = \frac{|p_{\text{ref}}|^2}{\rho c}, \tag{19}$$

and the reference pressure is

$$p_{\text{ref}} = 1\mu Pa. \tag{20}$$

From this, it follows [3] that

$$SL = 170.77\ dB + 10\log P_s. \tag{21}$$

If P_s is to be interpreted as the electric power driving the projector, the efficiency must be included. Also, since Equation 21 is valid only for the isotropic case, the directivity of the projector must be added. Thus, denoting the efficiency as η, Equation 21 generalizes to

$$SL = 170.77\ dB + 10\log\eta + 10\log P_s + DI. \tag{22}$$

The sonar equation quantifies the range limitations imposed upon active sonar performance imposed by the environment. There are essentially two types of limitation. The first is that due to the ambient noise. Since the *TL* term increases with range and *NL* represents ambient noise only, at some maximum range, the *NL* term will dominate. Increasing *SL* can overcome this limitation. However, as *SL* increases, the reverberation level increases proportionally, so that at some point, the reverberation becomes the limiting factor. At this point, only an increase in *TS* or the introduction of some sort of processing against the reverberation can help. Although some improvement in performance can be obtained by exploiting knowledge of the propagation physics, such as the multipath character of sonar echoes or other model-based approaches, most effort is directed toward non-parametric processing methods, since sufficiently detailed a priori knowledge to permit modeling of the propogation is usually not available, and even if it were available, exploiting it in an effective manner is no easy matter. Attempts to deal with reverberation in this manner are still at the research level [15].

4.2 Processing Gain

The sonar equation as expressed in Equation 13 does not allow for any temporal processing gain. That is, it simply deals with the power balance. To deal with this, we must have an expression for DT_s in terms of the temporal gain. Determining a meaningful value requires knowledge of the propagation conditions. In the case of the perfectly known signal, a case that essentially never occurs in practice, the gain is found from Equation 8, repeated below, where it is seen that the output *SNR* is given by

$$SNR_{\text{out}} = 2BT \times SNR_{\text{in}}, \tag{23}$$

and since $SNR_{\text{out}} = d$, we find, with the help of Equation 15, that

$$\frac{d}{2T} = B \times SNR_{in} = \frac{P}{N_s}, \tag{24}$$

so that

$$DT_s = 10\log\left(\frac{d}{2T}\right). \tag{25}$$

This can only be considered as an upper bound, since it is never achievable in practice, mainly due to energy splitting. From a theoretical standpoint, the return signal must be considered as a partially known signal, since its amplitude, phase and arrival time are really unknown parameters. Kay [7] shows that, for the CW case, the quadrature detector is the proper detector structure for such a signal, where the statistic, which is the periodogram, is sequentially tested until a maximum is found. Further, since the signal is spread by the channel, there will be several detections. Therefore, a post-detection integrator must be used. Consequently, an exact determination of the detection performance is simply out of question. Those interested in pursuing this issue further are referred to [7,16]. For the purpose of finding a reasonable estimate of the performance of a sonar system, however, an estimate of the channel spreading loss can be made. This is discussed in the following section.

4.3 Energy Splitting Loss

Weston [17] pointed out that as the number of multiple paths of different lengths increases for active sonar, the arrivals of the echoes in time tend to approach a Gaussian distribution. This led to a semi-empirical energy splitting loss (ESL) expression, referred to as the "Bell–Jones" model [18,19], which relates the ESL to the system resolution and the standard deviation (sigma) of the time spread. In this model, the ESL is determined from the ratio of the energy in the system resolution interval with the strongest signal to the total energy in the echo. Even in cases where the arrival time distribution is not strictly Gaussian, the Bell–Jones model still permits a quite useful ESL calculation, given an estimate of the time spread sigma. It also permits an inference of the effect of sonar resolution on time spread and a means of estimating the improvement from processing with differing resolutions. Figure 4 indicates the Bell–Jones ESL model developed from a Monte Carlo simulation based on random draws from a Gaussian distribution of arrival times, where each resolvable arrival is a random combination of multiple paths. The combining process of random-phase signals produces a Rayleigh distribution of intensity sums [20]. That is, for each bin of time spread, the intensity distribution is Rayleigh. The simulation was carried out over a large number of ratios of time spread sigma to system resolution. The model assumes that each resolvable arrival is a realization of a Rayleigh distribution.

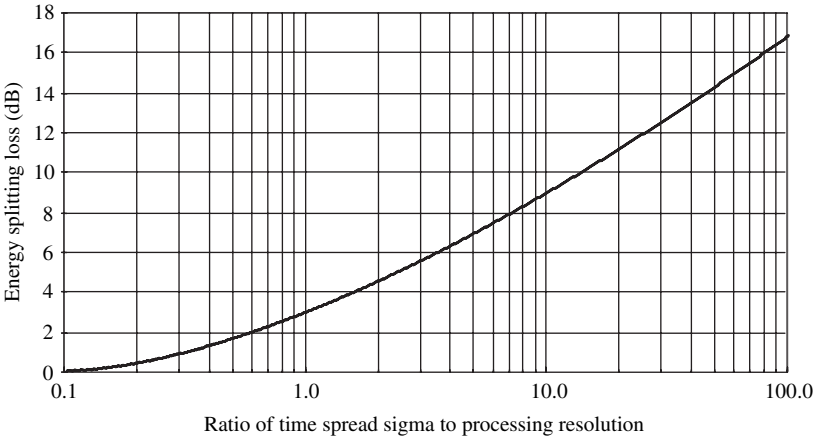


FIGURE 4 Example of the Bell-Jones splitting loss model for a shallow-water channel.

This provides a more accurate estimate of the peak signal arrival than would be expected if the simple average of the intensities of the resolvable arrivals were used to estimate the ESL. Computer ray path arrival time calculations under conditions of many paths, with the mean intensities computed for each resolvable arrival often, show distributions resembling Weston's prediction of a Gaussian tendency. Figure 4 shows the Bell–Jones model for the case of a shallow water multipath environment. The ESL concept is particularly important in this type of environment, where hundreds of two-way paths commonly exist between target and sonar. In practice, a more accurate estimate of ESL, and thus effective time spread sigma, is often obtained from direct ESL measurements. Given time spread sigma, the impact on ESL of changing system resolution may be deduced.

4.4 Examples of Performance Calculations

4.4.1 The Noise-Limited Case

Given a value for DT_s, we are now in a position to calculate the detection range for a specific case. This is done as follows. From the ROC curves in Figure 2, the desired detection performance defines a corresponding value of d, the parameter associated with the specified detection and false alarm probabilities. Given d, the resulting value of DT_s is subsituted into Equation 13 which is

then solved for *TL*. That is, for the spherical spreading case,

$$TL = \frac{1}{2}(SL + TS - NL_s + DI - DT_s) = 20\log R + \alpha R. \tag{26}$$

Consider the case of the perfect replica and a single 5-kHz pulse with a length of 0.1 s. Specifying a detection probability of 0.85 and a probability of false alarm of 0.05, Figure 2 tells us that $d = 8$. Thus, Equations 25 and 26 tell us that

$$\begin{aligned} TL &= \frac{1}{2}[(SL + TS - NL_s + DI - 10\log(\frac{d}{2T})] \\ &= \frac{1}{2}(SL + TS - NL_s + DI - 16). \end{aligned} \tag{27}$$

a source level of 220 dB, a target strength of 10 dB, a *DI* of 20 dB and an equivalent noise level of 50 dB re. 1 Hz,[3] and substituting these values into Equation 27 and eliminating *TL* using Equation 26, we find that

$$20\log R + \alpha R = 92\ dB. \tag{28}$$

Finally, assuming an absorption loss of .4 dB/km (from Figure 5.3 in Reference [1]), we find by numerical solution of Equation 28 that the maximum detection range for this case is

$$R_d = 17.6\,\text{km}. \tag{29}$$

This example is unrealistic in the sense that it has ignored any loss due to energy splitting. From Figure 4, we select an overall ESL of 6 dB. If the pulse bandwidth had been increased for the purpose of reducing the reverberation, increasing it beyond a certain point would cancel any payoff in reverberation processing due to an increase in energy splitting from target over-resolution. In this case, if we ignore any improvement in reverberation processing for the moment, the detection range would be reduced to 14.5 km. Thus, if there is not at least a 6 dB improvement in reverberation processing, there is no payoff in increasing the bandwidth beyond this point. Here, the judgement of the sonar designer must play an important role.

---

[3]This was taken from the Knudsen curves shown in Figure 7.5 of Reference [1], for the case of sea state three.

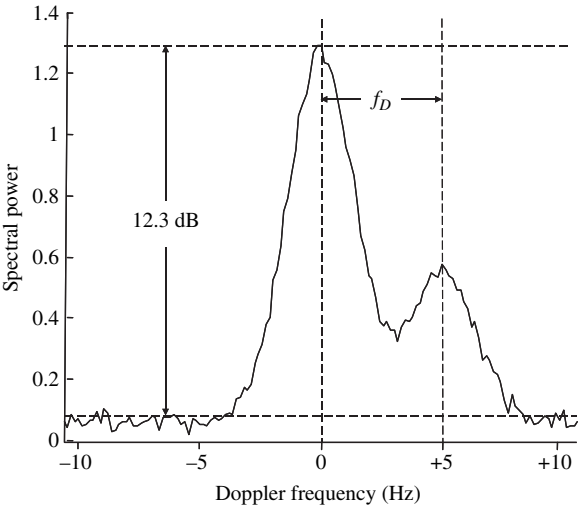


FIGURE 5 Example of Doppler processing for a moving target.

4.4.2 The Reverberation-Limited Case: High-Target Speed

When the target motion is sufficient for Doppler processing, the recipe is to transmit a narrow-band signal, so that the echo spectrum will be shifted away from the peak of the reverberation spectrum. This is depicted in Figure 5. The reverberation spectrum is a slightly broadened version of the signal spectrum. This spreading is a consequence of the finite beamwidths of the transmit and receive beams. Nevertheless, a sufficiently narrow signal bandwidth along with sufficient target motion will cause the echo to move toward the tail of this "reverberation ridge", thereby significantly reducing the reverberation in the band of the return signal. For the case of surface reverberation the insonified area is given by [4]

$$A = \frac{cTR\theta}{2} \tag{30}$$

with θ being the transmit beamwidth in radians [1]. The factor of 1/2 is to ensure full insonification of the area. The reverberation level is given by

$$RL = SL + DI + 10\log(s_s A) - 2TL. \tag{31}$$

[4]We consider only surface reverberation, since volume reverberation, is quite low at the frequencies considered here. For more information on surface reverberation, see reference [1].

Using Equation 30 to eliminate A, Equation 31 becomes

$$RL = SL + DI + S_s + 10\log(cTR\theta/2) - 2TL, \tag{32}$$

with $S_s = 10\log(s_s)$ being the surface backscattering strength. In order to evaluate the system in the reverberation-limited case, the reverberation level is compared to the echo level, where

$$EL = SL + DI + TS - 2TL. \tag{33}$$

Combining Equations 32 and 33, the detection threshold can now be written as

$$DT = EL - RL = TS - S_s - 10\log\left(\frac{cTR\theta}{2}\right). \tag{34}$$

Thus, the detection range is given by,

$$10\log R = TS - S_s - 10\log\left(\frac{cT\theta}{2}\right) - DT. \tag{35}$$

As expected, the source level does not appear, since both the reverberation level and the echo level are proportional to it. The value of S_s is highly variable, depending strongly on frequency, grazing angle and wind speed [1]. Also, we have assumed that the transmission loss for the reverberation and the signal are the same [1]. Here, we pick $S_s = -50\,\text{dB}$ as a representative value. For $TS = 10$ dB, $T = 0.1\,\text{s.}$, $c = 1500\,\text{m/s}$ and $\theta = .25$ radians (about 15°.), Equation 35 becomes

$$10\log R = 10 + 50 - 12.7 - DT. \tag{36}$$

As before, for the case of $P_d = .85$ and $P_{FA} = .05$, Figure 2 specifies a value of $d = 8$ and Equation 25 gives a value of $DT = 16\,\text{dB}$. Including a 4 dB energy spreading loss, Equation 35 now becomes

$$10\log R = 10 + 50 - 12.7 - 16 - 4 = 27.3\,\text{dB}, \tag{37}$$

or

$$R_d = .54\,\text{km}. \tag{38}$$

This is clearly unacceptable. Now suppose that the target has a velocity component of $1.5\,kts \approx 0.75\,\text{m/s}$ in the direction of the source. This means that the echo will be shifted up in frequency by $2 \times 0.75 \times 5000/1500 = 5\,\text{Hz}$. (Here we assume that the platform Doppler is removed by the processor.) Since the 0.1 s. pulse has a bandwidth of 10 Hz, we increase its length to

$0.5\,s.$, resulting in a bandwidth of 2 Hz, thereby moving the echo well to the edge of the reverberation ridge. This is depicted in Figure 5. Due to the increase in the pulse length, the new value for DT is 9 dB. However, the increase in pulse length has increased the insonification area by 3 dB, so that considering these changes alone, the TL remains essentially the same. But since the echo has been shifted by 5 Hz, the reverberation level at the new frequency has been considerably reduced. From Figure 5, we estimate this reduction to be 12.3 dB. Thus, our final result is

$$10\log R = 10 + 50 - 19.7 - 9 - 4 + 12.3 = 39.6\,\text{dB}, \tag{39}$$

resulting in a detection range of

$$R_d = 9.1\,\text{km}. \tag{40}$$

4.4.3 The Reverberation-Limited Case: Low Target Speed

Reverberation has essentially the same spectrum as the signal. This means that by spreading the signal spectrum, the reverberation spectral density is correspondingly reduced. This is seen as follows. Slightly reworking Equation 34, we find that

$$EL - RL = TS - 10\log\left(\frac{s_s cTR\theta}{2}\right). \tag{41}$$

Equivalently, the signal-to-reverberation ratio is

$$SRR = 10^{[TS - 10\log(s_s cTR\theta/2)]/10}, \tag{42}$$

or

$$SRR = \frac{2ts}{s_s cTR\theta} = P_{\text{sig}}/P_{\text{rev}}, \tag{43}$$

from which it follows that

$$P_{\text{rev}} = \frac{s_s cTR\theta}{2ts} P_{\text{sig}} = KTP_{\text{sig}} = KE, \tag{44}$$

where we have used $TS = 10\log(ts)$ and defined $K = s_s cR\theta/(2ts)$. From this, we see that the reverberation power is proportional to the signal energy. Assuming that the ambient noise and reverberation are uncorrelated, the total interference is simply the sum $N_0 + KE$. Indeed, it is shown in Reference [4] that

$$SIR = \frac{2EB}{N_o B + KE}. \tag{45}$$

It now follows that, for the purely noise-limited case,

$$SNR_{\text{out}} = \frac{2E}{N_0}, \tag{46}$$

as before, and for the purely reverberation-limited case,

$$SRR_{\text{out}} = \frac{2B}{K}. \tag{47}$$

Eliminating K from Equation 47, the result is

$$SRR_{\text{out}} = \frac{4tsB}{s_s cR\theta} = \frac{ts}{s_s A}2BT = \frac{ts}{r_s}2BT, \tag{48}$$

where we have defined a surface "reverberation strength" as $r_s = s_s A$. As can be seen, in the reverberation-limited case, increasing the bandwidth will improve the performance, as long as the target is not over-resolved. As previously discussed, increasing the bandwidth beyond this point will begin to decrease the effective target strength.

In order to compute the detection range, we assume a pulse length of 0.1 s. but increase the bandwidth to $200\,Hz$ by using a linear FM sweep. The detection threshold is 16 dB as in the previous case. From Equation 48, the detection threshold is given by

$$DT = 10\log(SRR_{\text{out}}) = TS + 10\log B - S_s - 10\log\left(\frac{cR\theta}{4}\right), \tag{49}$$

from which it follows that

$$10\log R = TS + S_s + 10\log B - 10\log[\left(\frac{c\theta}{4}\right) - 16]. \tag{50}$$

For the specified values, and adding a 6 dB to allow for ESL, we find that

$$10\log R = 10 + 50 + 23 - 19.7 - 16 - 6 = 41.3, \tag{51}$$

which results in a detection range of

$$R_d = 13.5\,\text{km}. \tag{52}$$

This result should be considered optimistic since the calculation has completely ignored ambient noise.

5 • Discussion

It is important to point out that the preceding analyses are subject to a great deal of uncertainty, since the parameters in the sonar equation are never known to a high degree of accuracy. However, they constitute a valuable means of assessment for operational systems. The importance of having an assessment of the detection range, even in cases of uncertain target strength or propagation loss, cannot be underestimated.

In actual practice, at sea, however, the problem is treated somewhat differently. Since one has access to the actual signal and a useful assessment of the noise is available, the actual detection threshold is available. The (minimum) required performance is then specified, usually as a detection probability of 50%. The difference between the actual detection threshold and the detection threshold for this 50% case is defined as the signal excess or SE [21]. Thus, for example, a negative SE indicates a degradation in performance. As seen in the examples in the previous section, a given detection performance leads to a specific detection threshold. Conversely then, a given SE specifies the actual detection performance in terms of a given tactical situation.

The major source of uncertainty is the lack of quantitative knowledge of the propagation conditions. Even given a reasonable propagation model, including it in the sonar processing scheme is a daunting task [15], since it changes with changes in the area of operation. Consequently, semi-empirical methods such as the Bell–Jones model [18,19] provide a useful means of quantifying the impact of the propagation channel.

Finally, it should be pointed out that, although it plays an important role in sonar system performance, the role of the display has not been treated here. This includes the impact of post-detection integration, ripple loss and pixel density on ultimate performance.

References

1. R. J. Urick, *Principles of Underwater Sound in the Sea*, McGraw-Hill, New York, 1975.
2. J. C. Toomay, *Radar Principles for the Non-Specialist*, Van Nostrand, New York, 1989.
3. W. S. Burdic, *Underwater Acoustic System Analysis*, Prentice-Hall, New Jersey, 1991.
4. A. H. Nuttall, "Signal Processing in Reveberation - a Summary of Performance Capability," Naval Underwater Systems Center TM No. TC-173–72, 1972.

5. R. W. Schwenke, "Sensitivity Analysis of an Estimator-Correlator for the Detection of Spread Targets With Multiple Discrete Highlights," PhD Thesis, The Pennsylvanis State University, December, 2000.
6. A. D. Whalen, *Detection of Signals in Noise*, Academic Press, New York, 1971.
7. S. M. Kay, *Statistical Signal Processing; Vol II, Detection Theory*, Prentice-Hall, New Jersey, 1998.
8. W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The Theory of Signal Detectibility," Transactions of the IRE Professional Group in Information Theory, PGIT, 1954.
9. W. W. Peterson and T. G. Birdsall, "The Theory of Signal Detectability," *Univ. Mich. Eng. Res. Inst. Rep.*, 13, 1953.
10. D. Slepian, "On Bandwidth," *Proc. IEEE*, Vol 3, No. 3, 1976, pp. 292–300.
11. J. L. Stewart and E. C. Westerfield, "A theory of active sonar detection," *Proc. IRE*, Vol. 47, pp. 872–881, 1959.
12. P. M. Woodward, *Probability and Information Theory with Applications to Radar*, Pergammon, London, 1952.
13. J. P. Costas, "A study of a class of detection waveforms having nearly ideal range-Doppler anbiguity properties," *Proc. IEEE*, Vol. 72, No. 8 1984, p. 996.
14. D. M. Drumheller and E. I. Titlebaum, "Cross-correlation properties for algebraically constructed Costas Arrays," *IEEE Trans. on Aerospace and Electrical Systems*. Vol 27, No. 1, January, 1991.
15. J.-P. Hermand and W. I. Roderick, "Acoustic Model-Based Matched Filter Processing for Time-Dispersive Ocean Channels: Theory and Experiment" (Invited Paper), *IEEE J. Oceanic Eng. (Special issue on sonar system technology)*, Vol. 18, 1993, pp. 447–465
16. W. S. Knight, R. G. Pridham and S. M. Kay, "Digital signal processing for sonar," *Proc. IEEE*, Vol. 69, pp. 1451–1506, 1981.
17. D. E. Weston, "Correlation loss in echo ranging," *J. Acoust. Soc. Am.*, Vol. 37, No. 1, January, 1965.
18. T. G. Bell, "Predicting and dealing with energy spreading loss," Proceedings of a Seminar on Active Signal Reverberation, Tracor Document T90-01-9532-U, 1989.
19. B. W. Jones, "Measurement of time spread and energy splitting," Proceedings of a Seminar on Active Signal Reverberation, Tracor Document T90-01-9532-U, 1989.
20. G. H. Robertson, "Operating Characteristics for a Linear Detector of CW Signals in Narrow-Band Gaussian Noise," *Bell Syst. Tech. J.*, Vol. 46, No. 4 1967.
21. D. H. Whalen, W. C. Mylander and T. J. Sanders, eds. *Naval Operations Analysis*, Naval Institute Press, 3rd Edition, Annapolis, 1999.

97

Passive Sonar

Brian H. Maranda

Defence Research and Development Canada – Atlantic, Dartmouth, NS, Canada

1 • Introduction

Passive sonar is a method for detecting acoustic signals in an underwater environment, usually the ocean. The difference between passive and active sonar is that a passive sonar system emits no signals; instead, its purpose is to detect the acoustic signals emanating from external sources. From an historical perspective, the main impetus for the research and development of passive sonar has been its military applications, in particular, the acoustic detection of submarines. The pace of military research has abated in recent years, as the introduction of increasingly quiet submarines into service has shifted the focus to active sonar. However, new applications of passive sonar continue to appear; for example, a heightened awareness of environmental issues has spurred the development of passive techniques for detecting whales.

Passive sonar is a mature technology, in the sense that improvements are expected to be incremental in nature. Although the inexorable advance of electronics continues to open up new vistas for system improvement, progress is restrained by difficulty in algorithm development. For instance, there are many functions within current sonar systems that it would be highly advantageous to automate, but progress in algorithm development for automation has been slow to date. The current trend is to improve passive sonar performance by taking a more global view of the sonar problem. For example, other sensor platforms may be used

to feed the sonar with environmental information that can be used to aid performance, at the expense of making the sonar less self-contained. At an even higher level, a number of sonars, both passive and active, may be connected into a network that enables the cooperative detection and tracking of targets.

The acoustic data analysis in passive sonar is often decomposed into three stages: detection, classification, and localization. The task of the detection stage is to sense the presence of acoustic signals of interest in a background of noise, and is a prerequisite for the subsequent processing. When an acoustic source is detected, the sonar is said to be *holding contact* on the source, and the source itself may be described as a contact. If there are many simultaneous contacts, a classification stage is employed to identify which contacts represent sources of interest and which can be ignored. In a military sonar, for example, the classifier may try to distinguish between surface ships and submarines; i.e., between surface and sub-surface contacts. Upon identification of a source of interest, an attempt may be made to localize the source and track it as a function of time. Much detailed information on the signal processing in passive sonar can be found in review articles [1,2] and textbooks [3–5].

2 • Passive Sonar Systems

Figure 1 illustrates several types of passive sonar systems that are in use today. The *towed array* is a neutrally buoyant line array that is towed behind a surface ship or submarine; the signals from the array sensors are telemetered to a shipboard sonar processor through the tow cable. Surface ships and submarines are often equipped with hull-mounted sonars, such as cylindrical or spherical bow arrays and flank-mounted arrays. A *sonobuoy* is an expendable device that can be dropped from an aircraft into the ocean and which transmits its sensor signals via a radio link to the aircraft's sonar processor. For the passive surveillance of coastal areas or harbors, arrays may be installed on the ocean bottom and connected by cable to a shore-based processing station.

The sensors that are immersed in the ocean comprise the *wet end* of the sonar system. The basic sensor is the pressure-sensitive hydrophone, which is usually chosen to have an omnidirectional response at the frequencies of interest; particle-velocity sensors are sometimes used to realize a directional response. One of the main goals when designing the sensor is to make it *ambient-noise limited*; that is, the main contributor to the total noise level should be the waterborne ambient noise that is incident on the

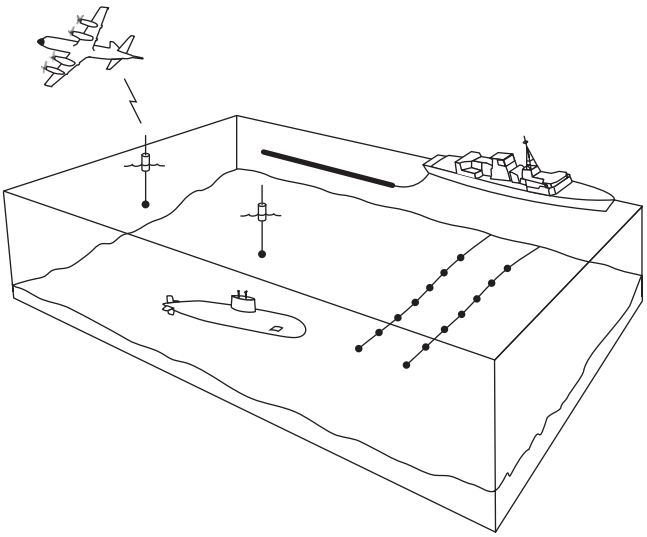


FIGURE 1 Different types of passive sonar in common use today (not drawn to scale).

sensor. This type of noise is intrinsic to the ocean environment and imposes an unavoidable limitation on sonar performance. Other noise components are referred to as *self noise*: examples are the electronic noise of the sensor amplifier, noise ducted into the sensor from its mechanical mounting, and flow noise (arising from water moving past the sensor). Unlike ambient noise, self noise is under the control of the designer and is amenable to reduction through improved sensor construction. Arrays of sensors are frequently used to increase detection sensitivity or to provide directional capability.

The electronics for the processing and recording of data constitute the *dry end* of the sonar system. Modern systems rely almost exclusively on digital signal processing. In the early implementations, custom-built processors were used for the computationally intensive parts of the processing, but commercial equipment has now largely replaced these expensive, specialized processors. The signal-processing software is typically written in a high-level language; a real-time operating system may provide the run-time environment, although with modern hardware the processor loading may be sufficiently light that a standard operating system without a real-time kernel can be used. The main role performed by analog processing in a present-day sonar is the signal conditioning, such as the amplification and filtering, that is required between the sensor and the analog-to-digital

converter (ADC). In early passive sonars the dynamic range of the front-end ADC was insufficient to accommodate the long-term variability in ambient-noise levels, and the analog signal conditioning had to include automatic gain control (AGC). The advent of digitizers having a wide dynamic range (e.g., 24 bits) has made it possible to eliminate the AGC.

Sonars today still require human operators to interpret the results of the data processing, which are presented to the operator aurally or, most commonly, on a visual display. There are several issues that arise from having a human operator in the loop. On the one hand, it is of paramount concern to present the data to the operator in such a way that detection sensitivity is not lost; on the other hand, the volume of processed data has grown immensely with advances in technology, and hence the data must be condensed to avoid overwhelming the operator. Automated aids have been developed as a means to reduce the workload, but automation has not yet progressed to the point that the operator can be supplanted.

3 • The Passive Sonar Equation

The fundamental purpose of passive sonar is to detect signals embedded in noise, and a rough prediction of detection performance can be obtained from the *sonar equation* [3]. One form of the passive sonar equation is

$$\mathrm{SE} = \mathrm{SL} - \mathrm{NL} + \mathrm{AG} - \mathrm{TL} - \mathrm{DT}$$

where the variables are defined as follows:

- SE – the *signal excess* at the input to the detector (in dB);
- SL – the *source level* at a distance 1 m from the acoustic center of the radiating source (a sound pressure level in dB re 1 μ Pa at 1 m);
- NL – the *noise level* at the receiver (a spectrum level in dB re 1 μ Pa^2/Hz);
- AG – the *array gain*, which is the improvement in signal-to-noise ratio (SNR) provided by the directivity of the sensor system relative to an omnidirectional hydrophone (in dB);
- TL – the *transmission loss* between the source and the sonar receiver (in dB);
- DT – the *detection threshold*, or SNR required by the detector to achieve a specified level of performance (in dB).

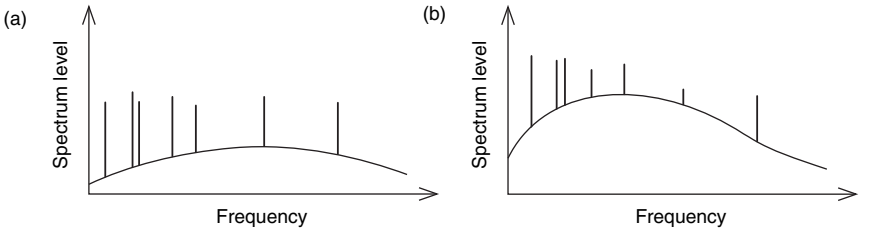


FIGURE 2 Schematic representation of the radiated noise spectrum from a ship or submarine. (**a**) At low speed, (**b**) at high speed (based on a diagram in [3].)

The array gain is often called the *directivity index* (DI) when the noise is isotropic. The signal excess is the amount by which the SNR at the detector input exceeds the detection threshold, and when $SE \geq 0$ the sonar system will meet or exceed the desired level of performance. Note that the array gain and detection threshold are the only terms in the sonar equation that are determined by the sonar signal processing, and consequently they are the main focus of the rest of this article. Other articles in the handbook should be consulted for information on transmission loss and ambient noise level (equal to NL when the sonar is ambient-noise limited); sound sources are briefly treated in the next paragraph.

For the military application of passive sonar, the sound radiated by ships and submarines is of most interest, and the signal characteristics are determined empirically or from theoretical models of the signal-generating mechanisms [6]. The radiated noise spectrum from a ship is a mixed spectrum, containing both discrete and continuous components (Figure 2). The discrete components are the tonal, or narrowband, components produced by such rotating machinery as propellers, engines, compressors, and pumps. These tonal components are described as speed-dependent or speed-independent according to whether or not they change in frequency and/or level when the vessel changes speed. The continuous broadband noise typically originates in propeller cavitation or other flow phenomena, and rises in level with increasing speed. Ross [6] states that a submarine traveling at low speed radiates on the order of only 10 mW of acoustic power.

4 • Spatial Processing and Array Gain

A distributed array of sensors can be used to perform spatial filtering of the signals incident upon it, thereby providing both array gain and directionality information. Baggeroer distinguishes two methods of spatial processing [1]. The first is to form

directional beams, each of which admits an acoustic arrival from a specified direction while suppressing arrivals from other directions. The most important algorithm for forming beams is the *delay-and-sum beamformer* (or *conventional beamformer*), which is usually designed to produce a wideband time series for each beam. Detection processing is applied subsequently by treating the beams as multi-channel time-series data. The second method of spatial processing is to estimate the acoustic intensity as a function of frequency and wave number, a method that borrows techniques from spectral estimation. This second approach does not yield beam time series following the spatial processing, although in common usage any type of spatial processing is referred to as beamforming.

The principles underlying the conventional beamformer will now be discussed in more detail, and optimum and adaptive processing methods will also be reviewed. The emphasis in this section is on beamforming in support of detection, and so those methods that are used primarily for estimation (e.g., of bearing or range) are treated in the section on localization. Several of the formulas will be specialized to the case of a line array, which is the most common array geometry in underwater acoustics; see [7,8] for a treatment of more general array configurations.

4.1 Delay-and-Sum Beamforming

Given an array of N sensors, let $x_n(t)$ denote the waveform received at the nth sensor (for $n = 0, \ldots, N-1$). The output $b(t)$ of the delay-and-sum beamformer (Figure 3) is

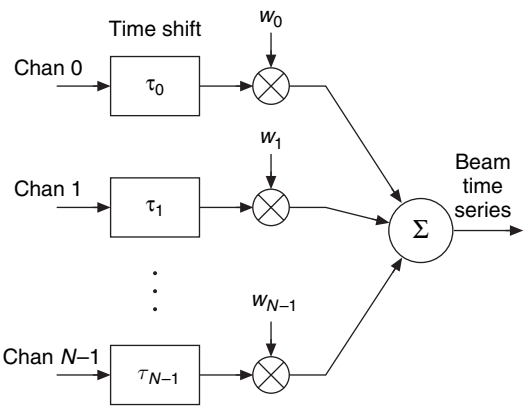


FIGURE 3 The delay-and-sum, or conventional, beamformer. The inputs are the waveforms measured by the acoustic sensors.

$$b(t) = \sum_{n=0}^{N-1} w_n x_n(t + \tau_n), \tag{1}$$

where the τ_n are time shifts inserted by the beamformer and the w_n are weights, or *shading* coefficients, that are used to shape the beam response. When the inserted time shifts are equal to the actual delays of a signal wave front propagating across the array, the signal is summed coherently by the beamformer and is reinforced; signals arriving from other directions do not add coherently and are attenuated. The time shifts τ_n are derived from an assumed model for the shape of the signal wave front, the most important of which is the plane-wave model. A plane-wave signal having temporal dependence *s(t)* is of the form $x(t, \mathbf{r}) = s(t - \mathbf{u} \cdot \mathbf{r})$, where $\mathbf{r}$ is the position in space and $\mathbf{u}$ is the slowness vector ($\mathbf{u}$ points in the direction of propagation and has magnitude $1 / c$, where c is the speed of sound). Then, for an array of omnidirectional sensors located at positions $\mathbf{r}_n$, the wave form received by the nth sensor is $x_n(t) \equiv x(t, \mathbf{r}_n) = s(t - \mathbf{u} \cdot \mathbf{r}_n)$. When the time shifts compensate for the actual time delays, i.e., when $\tau_n = \mathbf{u} \cdot \mathbf{r}_n$ in Eq. (1), $b(t) = s(t) \sum_n w_n$ and the beamformer passes the signal undistorted. For the case of an equispaced line array (Figure 4), the modeled plane-wave delay relative to sensor 0 is given by

$$\tau_n(\theta) = n \frac{d}{c} \sin \theta, \tag{2}$$

where d is the sensor separation and θ is the *steering angle*, i.e., the direction of propagation assumed by the beamformer. Owing to the axial symmetry of the line array, the plane-wave model is parameterized by the single parameter θ; here $\theta = 0°$ corresponds to the *broadside* direction and the angles $\theta = \pm 90°$ to the *endfire* directions. In a *focused* beamformer, the wave front is assumed to be spherical and the time delays are parameterized in terms of range and angle, $\tau_n = \tau_n(r, \theta)$.

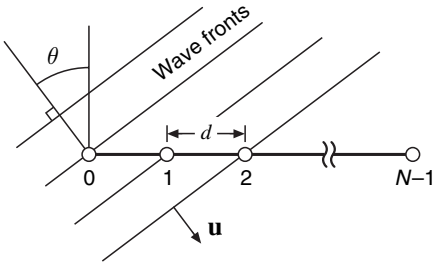


FIGURE 4 The geometry of a plane-wave signal impinging on a line array of sensors. The wave fronts are defined as surfaces of constant phase.

4.2 Frequency-Domain Formulation

Formulating the beamforming operation in the frequency domain has proven to be very fruitful, both for theoretical work and for implementations. In what follows, the Fourier transform of an arbitrary time function $x(t)$ will be denoted by $X(\omega)$, where $\omega = 2\pi f$ is radian frequency. We also introduce the wavenumber vector $\mathbf{k} = \omega\mathbf{u}$, which points in the direction of propagation and has magnitude $k = \omega/c$. Taking the Fourier transform of Eq. (1) and using the shift property of Fourier analysis, we find

$$B(\omega, \mathbf{k}) = \sum_{n=0}^{N-1} w_n X_n(\omega) \exp(i\omega\tau_n) = \sum_{n=0}^{N-1} w_n X_n(\omega) \exp(i\mathbf{k}\cdot\mathbf{r}_n), \quad (3)$$

where the plane-wave model has been used to write $\omega\tau_n = \mathbf{k}\cdot\mathbf{r}_n$. The quantity $|B(\omega, \mathbf{k})|^2$ is a measure of the magnitude of the acoustic intensity at frequency ω and wave number $\mathbf{k}$. If the actual signal is a plane wave $x_n(t) = s(t - \mathbf{u}_0\cdot\mathbf{r}_n)$ having wave number vector $\mathbf{k}_0 = \omega\mathbf{u}_0$, it is readily seen that $B(\omega, \mathbf{k}) = S(\omega)G(\omega, \mathbf{k})$, where the *beam pattern*

$$G(\omega, \mathbf{k}) \equiv \sum_{n=0}^{N-1} w_n \exp[i(\mathbf{k} - \mathbf{k}_0)\cdot\mathbf{r}_n] \quad (4)$$

indicates how the arrival with wave number vector $\mathbf{k}_0$ is spatially filtered when the beamformer is steered to $\mathbf{k}$. For the special case of a line array having uniform weights $w_n = 1/N$,

$$|G|^2 = \left[\frac{\sin(N\Delta\phi/2)}{N\sin(\Delta\phi/2)}\right]^2 \quad (5)$$

with $\Delta\phi \equiv kd(\sin\theta - \sin\theta_0)$. Here θ is the angle at which the beam is steered, and θ_0 the angle at which the signal wave front arrives. One can also write $kd = \pi f/f_\mathrm{d}$ where $f_\mathrm{d} \equiv c/2d$ is the *design frequency* of the line array. This is the frequency for which the sensor spacing is $\lambda/2$ in terms of the wavelength λ, and hence it represents a Nyquist criterion for spatial sampling: when $f > f_\mathrm{d}$ the array may exhibit spatial aliasing (so-called *grating* lobes). Figure 5 shows a plot of $20\log|G|$ for a uniformly weighted line array of 15 sensors steered to $\theta = 30°$. A main lobe appears at the steering angle, along with numerous side lobes; no grating lobes are present. The highest side lobes are immediately adjacent to the main lobe and are approximately -13 dB below it. In most applications the weights w_n are selected to reduce the side lobe levels, at the expense of increasing the width of the main lobe.

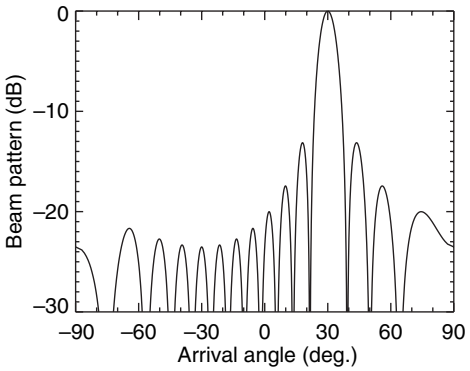


FIGURE 5 The beam pattern of a uniformly weighted line array of 15 sensors steered to $\theta = 30°$. The signal frequency is equal to the array design frequency.

4.3 Implementations

There is a wide range of beamformer implementations, which differ in their computational efficiency and their suitability for a given type of hardware [9]. When implementing the delay-and-sum beamformer in the time domain, the main design problem is to devise a method for accurately realizing the desired time shifts τ_n in the sensor channels. A simple approach is to sample the sensor waveforms at such a high rate that the sampling period itself provides sufficient time granularity (i.e., shifting the channels by an integral number of samples is accurate enough). Another approach is to sample at a lower rate and use time-domain interpolation filters to realize fractional delays [10].

Frequency-domain implementations exploit the computational efficiency of the fast Fourier transform (FFT). The digitized sensor waveforms are first segmented into finite-duration blocks, which are often referred to as data *snapshots*. After the FFT has been applied to the data blocks, the channels are phased and summed as in Eq. (3); computationally efficient methods for the line array are given in [11]. For the discrete-time processing of finite-duration blocks, the phasing operation represents a cyclic shift of the waveforms in the time domain. If it is desired to use an inverse FFT to produce a beam time series having a true wideband response [12], then this wraparound effect must be handled explicitly using an overlap-add or overlap-save algorithm (Figure 6). However, if transformation back into the time domain is not required – for example, if $|B(\omega, \mathbf{k})|^2$ is the desired output from the spatial processing – then the wraparound is often simply ignored. In this latter case, the block length would be chosen such that the maximum possible wraparound is a small fraction (less than 0.1, say) of the block length.

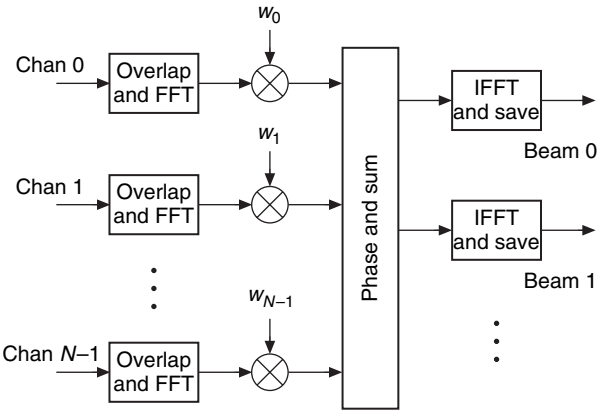


FIGURE 6 A frequency-domain implementation of a wideband conventional beamformer, where the overlap-save algorithm is used to handle wraparound.

4.4 Optimum and Adaptive Array Processing

There are array-processing methods that seek to optimize specific performance criteria (e.g., to maximize the array gain), and hence are termed *optimum* [8]. It is convenient to describe the beamforming operation in vector and matrix notation; in what follows, superscript T and * will denote transpose and conjugate transpose respectively. If $\mathbf{x}^{\mathrm{T}} \equiv (X_0(\omega), \ldots, X_{N-1}(\omega))$ is the vector of sensor Fourier coefficients from a single data snapshot, then the conventional beamformer output can be expressed as $B = \mathbf{w}^*\mathbf{x}$, where $\mathbf{w}^* \equiv (w_0 \exp(i\omega\tau_0), \ldots, w_{N-1} \exp(i\omega\tau_{N-1}))$ is a weighted steering vector. The *cross-spectral density matrix* (CSDM) is defined as $\mathbf{R} \equiv E\{\mathbf{x}\mathbf{x}^*\}$ where $E\{\cdot\}$ denotes expectation. Then $|B|^2 = |\mathbf{w}^*\mathbf{x}|^2 = \mathbf{w}^*\widehat{\mathbf{R}}\mathbf{w}$, where $\widehat{\mathbf{R}} \equiv \mathbf{x}\mathbf{x}^*$ is a sample estimate of the CSDM. In beamforming algorithms that use $\widehat{\mathbf{R}}$ explicitly, the sample estimate is often improved by averaging over multiple data snapshots. The maximum averaging time is limited by source motion, which is usually required to be small during the average.

In a narrowband formulation, the optimum processor specifies a weight vector $\mathbf{w} = \mathbf{w}_{\mathrm{opt}}$ to be used in the beamformer at each frequency ω; it is shown in Section 5.2 how the narrowband results are combined across frequency for broadband detection. A processor that is optimum under several different criteria [8,13] has the weight vector

$$\mathbf{w}_{\mathrm{opt}} = \frac{\mathbf{Q}^{-1}\mathbf{v}}{\mathbf{v}^*\mathbf{Q}^{-1}\mathbf{v}}, \tag{6}$$

where $\mathbf{Q}$ is the CSDM of the noise alone and $\mathbf{v}^* \equiv (\exp(i\omega\tau_0), \ldots, \exp(i\omega\tau_{N-1}))$ is the modeled signal vector. In modern terminology, this is the *minimum-variance distortionless response* (MVDR) beamformer. Note that in practice the matrix $\mathbf{Q}$ will be unknown (and time-varying). It is generally the case that the formulas defining optimum processors depend on unknown parameters of the signal and/or noise field, and hence their practical implementation requires *adaptive* or *data-dependent* algorithms that estimate the parameters, perhaps recursively, from the array data [8,14,15]. The estimation procedure can pose difficulties; for example, estimating the noise-only matrix $\mathbf{Q}$ would require removing the signals, but this cannot be done reliably in passive sonar given the inherent uncertainty about signal characteristics. For MVDR processing, the sample estimate $\widehat{\mathbf{R}}$ of the signal-plus-noise matrix is usually used in place of the noise matrix $\mathbf{Q}$, a substitution that has ramifications for the performance of the modified algorithm.

Adaptive beamforming has not made extensive inroads into operational sonar systems [16], although the situation may be changing [17]. For a long line array operating in approximately isotropic noise, and in the absence of discrete interfering sources, the improvement in array gain offered by optimum beamforming is modest near broadside; it is significant only near the endfire directions, where conventional beams are wide [18]. A more attractive feature of optimum processing is its ability to reject strong interfering signals, although the realization of a robust adaptive algorithm has proven difficult for passive sonar. As examples of the problems that occur [8,15], a desired signal may be rejected if it is highly correlated with an interfering signal at another angle; and, if there is a mismatch between the signal model and the actual received signal, an adaptive algorithm may regard the signal as noise and suppress it. Much research has been conducted on trying to improve the robustness of adaptive algorithms [8].

4.5 Array Gain

For the conventional beamforming of a uniformly shaded line array in isotropic noise, the array gain at the design frequency is $\mathrm{AG} = 10\log_{10} N$. Conventional beamforming is in fact optimum under the stated conditions, since isotropic noise is uncorrelated from sensor to sensor (i.e., $\mathbf{Q}$ is the identity matrix) at the design frequency, and $\mathbf{w}_{\mathrm{opt}}$ in Eq. (6) becomes $\mathbf{w}_{\mathrm{opt}} = \mathbf{v}/N$. See [4] for methods of computing the array gain at other frequencies or in arbitrary noise fields. Also, the above formula for AG assumes that the signal wave front is perfectly coherent across the array, whereas in reality the signal coherence will fall off with distance

owing to such propagation effects as scattering from the ocean boundaries and distortion caused by local inhomogeneities in the medium. Hence there is a practical limit on the array gain that can be achieved by beamforming a single long aperture [19].

5 • Detectors

In passive sonar, the signals lie outside the control of the sonar designer and vary widely in their characteristics (e.g., the bandwidth may vary greatly from signal to signal). Because of the uncertainty concerning the received signals, the basic detector used in passive sonar is the energy detector (Figure 7). The detector input (which may be a beam in a beamformed system) is filtered, squared, and integrated to produce the detector output, which is usually displayed to a sonar operator in some fashion. The standard approach to mitigating the inherent uncertainty concerning the signals is to construct a bank of energy detectors, and to tune the filter bandwidth and the integration time of each detector for a certain type of signal. Figure 8 illustrates a conceptual sonar system built along these lines. The narrowband and broadband detectors, being the most important, are given the most extensive description below.

5.1 Narrowband Detector

The schematic for a narrowband detector is shown in Figure 9(a), in which a bank of narrowband filters is realized in parallel via the FFT. The sensor or beam data are segmented into overlapping blocks, a weighting function is applied to each block, and an FFT is performed. The purpose of the weighting function is to suppress unwanted side lobes by shaping the FFT bins. A sample spectrum, or periodogram, is produced by taking the squared magnitude of the bins; next, spectra from successive FFTs are integrated (summed) to reduce the variance [20]. The integration time may be a few seconds, or in some applications

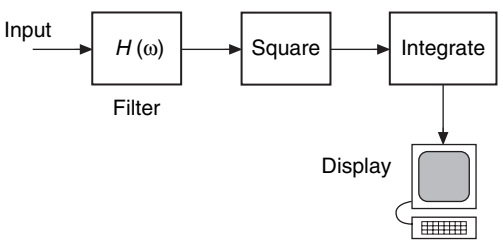


FIGURE 7 The energy detector for a single channel. The input comes from either a sensor or beam.

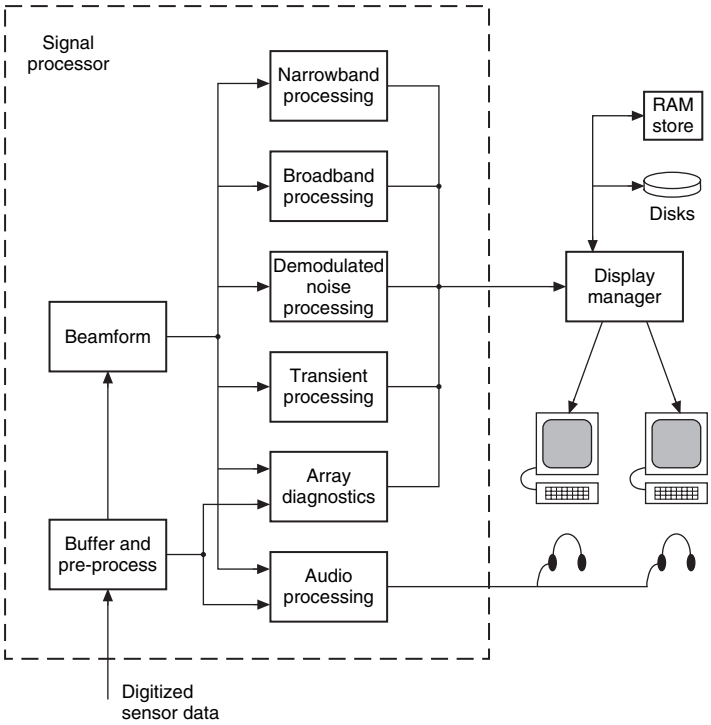


FIGURE 8 The organization of the signal processing in a conceptual passive sonar system. Descriptions of the various detectors can be found in Section 5.

extend over a minute. The spectra are usually normalized [21,22] before display to make the detector's operating point independent of the background noise level; i.e., to achieve a constant false alarm rate (CFAR). The most common form of display is the intensity-encoded spectrogram (called a *lofargram* in military usage), in which successive spectra are stacked vertically on the display surface; see Figure 9(b). A narrowband spectral component appears as a line on such a display. Noise-normalized spectrograms provide sensitive detection performance, while also furnishing such classification information as the frequency and bandwidth of the lines, their dynamic behavior, and the harmonic relationship of multiple lines.

Methods for computing the detection threshold of the narrowband detector can be found in [5]. To maximize detection sensitivity, the spectral analysis bandwidth should be matched to the signal bandwidth, which may vary considerably from signal to signal. To effect the desired matching, a bank of

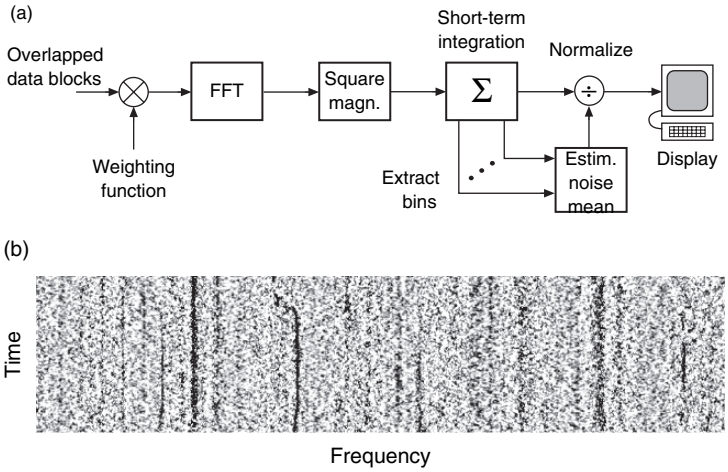


FIGURE 9 The narrowband detector. (**a**) Schematic of an FFT implementation, and (**b**) an intensity-encoded spectrogram, or lofargram, of acoustic data recorded from a sonobuoy (data courtesy of F. Desharnais).

narrowband detectors can be implemented with different analysis bandwidths, or frequency resolutions. In some existing sonar systems, the narrowband processing provides both low-resolution spectrograms, covering the entire frequency band of interest, and reconfigurable high-resolution spectrograms, or *verniers*, covering operator-selectable portions of that band. The present trend in sonar design is to exploit the capability of modern computer technology by computing and storing the data for all the desired signal-processing options, even those not currently being displayed; the sonar operator can then roam through the database, examining the spectrograms at the most appropriate resolution.

5.2 Broadband Detector

Considered next is the detection of a broadband signal, such as the broadband sound radiated from a ship. We first discuss the single-channel detector, whose input may come from a single hydrophone or, for a multi-hydrophone array, from a beam. In this first arrangement, the beamformer and the detector are functionally separate. Later we discuss the case when the broadband detection problem is directly formulated for a multi-channel array, so that the beamforming is optimized as part of the detector. In all cases, the signal and noise will be modeled as stationary stochastic processes that are Gaussian distributed.

We thus start with the single-channel detector. The exact theory of the optimum detector is complex, but the detector in Figure 7

is close to optimum for long observation times [23]. A frequency-domain implementation is straightforward if the system already produces periodograms as described in Section 5.1: for a pre-detection filter with frequency response $H(\omega)$, the detector output is obtained at each integration time by weighting an unnormalized periodogram by $|H(\omega)|^2$ and summing across frequency (an application of Parseval's theorem). One well-known form of the pre-detection filter is the *Eckart filter* [4] given by $|H(\omega)|^2 = S_s(\omega)/S_n^2(\omega)$, where $S_s(\omega)$ and $S_n(\omega)$ are the power spectral densities of the signal and noise at the detector input (i.e., in a beam for a beamformed system). Since the spectra $S_s(\omega)$ and $S_n(\omega)$ are in general unknown, the pre-detection filter is usually chosen to have a simple rectangular passband located where the signal energy is presumed to lie. For a beamformed array, a standard broadband display is the *bearing–time intensity* (BTI) plot, having bearing and time on the axes of the display surface and showing the detector output level as intensity. Normalization can be carried out by estimating the background level across the beams.

The broadband detection problem has also been analyzed when the detector input is multi-channel array data [5,13,24]. Now let $S_s(\omega)$ and $S_n(\omega)$ denote the power spectral densities of the signal and noise at a sensor. If the observation time T is long enough that $S_s(\omega)$ and $S_n(\omega)$ vary little over a frequency width $2\pi/T$ (the broadband assumption), the detector derived from the likelihood ratio consists of the MVDR beamformer followed by a certain pre-detection filter. Specifically, the detector output is the sum over frequency of $|H(\omega)|^2|\mathbf{v}^*\mathbf{Q}^{-1}\mathbf{x}|^2$, where

$$|H(\omega)|^2 = \frac{S_s(\omega)/S_n^2(\omega)}{1+g(\omega)S_s(\omega)/S_n(\omega)}$$

and $g(\omega) \equiv \mathbf{v}^*\mathbf{Q}^{-1}\mathbf{v}$ [cf. Eq. (6)]. The structure of the optimum processor is that assumed previously: a beamformer followed by a detector of the form shown in Figure 7. The optimum filter again depends on characteristics of the signal and noise that are unknown in practice.

The output of the broadband detector is approximately Gaussian distributed when its time–bandwidth product is large, and the detection threshold can be determined from standard curves [3,5].

5.3 Other Detectors

In some circumstances, a source producing broadband noise modulates it in such a way that its power rises and falls harmonically. As the name implies, demodulated-noise processing (often called DEMON processing [25]) seeks to detect this modulation

through a form of demodulation. The reader is referred to works by Nielsen [5,25] for details on this specialized detector.

The detectors discussed so far are tailored to signals that persist much longer than the update time of the corresponding sonar displays, which may be on the order of 10 s. In this context, transient signals are those signals whose duration is shorter than a single update time, and hence which typically last for a few seconds or less. The main problem faced in the design of a transient detector is the uncertainty concerning the time duration and frequency extent of the signal; to achieve adequate detection performance against a wide variety of signal types, the energy detector may once again be employed. Methods for performance analysis can be found in [26]. (Although transient signals may be broadband, they are usually viewed as unknown deterministic signals and thus the methods of analysis differ from those used for the broadband detector of Section 5.2.) More sophisticated transient detectors are described, and their performance is compared, in [27].

Most of the detection processing in modern passive sonar culminates in a visual display; however, the human ear is a sensitive detector that can be exploited. The processing done to the audio stream before presentation to the human listener may be quite simple, consisting of little more than removing out-of-band noise by filtering the sensor or beam output. Most of the processing is done digitally, the analog signal being produced at the last step with a digital-to-analog converter.

6 • Classification

When a passive sonar system is deployed in an ocean area with a large amount of ship traffic, its displays will often be cluttered with a large number of contacts. The narrowband display of a towed-array sonar may show hundreds of spectral lines, many associated with the same source. Hence there is a need to classify the contacts and isolate those of particular interest; in military surveillance, the goal may be to identify a submarine intruder in the presence of a large number of surface vessels.

Narrowband processing provides much of the information that leads to the classification of sea-going vessels: each class of vessel has its own narrowband signature (specific pattern of narrowband acoustic emissions) as determined by the drive train and other equipment on the vessel. For example, a number of lines may be harmonically related, a property that suggests they originate from the same source. A trained sonar operator is capable of "reading" the patterns on a narrowband spectrogram

and classifying the type of vessel. Additional classification information comes from those localization methods that can determine the source depth.

7 • Localization

In this article, the term *localization* refers to any technique for obtaining positional information about a source, which may be its bearing alone; the term is sometimes used only if the source range is also found. When the localization stage of the sonar processing is reached, the emphasis changes from detection to parameter estimation. Much research [24,28,29] has been done on constructing optimum estimators for bearing and range, and on deriving bounds on the attainable accuracy in the presence of noise. The Cramér–Rao lower bound (CRLB) has been used extensively for this latter task. The optimum estimators generally assume knowledge of unknown parameters of the signal and noise field, and the emphasis here will be on the sub-optimum algorithms that are used in practice. In many cases, the sub-optimum processors perform only slightly worse than the theoretically optimum processors.

Many localization algorithms rely on estimating time delay, and it is worthwhile to give a short summary of the method used for broadband signals. In the simplest case, there are two received signals, $x_1(t)$ and $x_2(t)$, which are modeled as

$$x_1(t) = s(t) + n_1(t),$$
$$x_2(t) = s(t - \tau_d) + n_2(t),$$

where $s(t)$ is a wideband stochastic process, $n_1(t)$ and $n_2(t)$ are noise terms, and τ_d is the unknown delay. The standard estimator of τ_d cross-correlates $x_1(t)$ and $x_2(t)$, and then uses the position of the peak in the correlogram as the delay estimate [30,31]. Note that time-delay estimation is problematic for a narrowband signal because the correlogram will exhibit multiple peaks spaced at the signal period, giving rise to ambiguity [32].

7.1 Bearing Estimation

Although it is common to speak of bearing estimation, in actuality one estimates the *direction of arrival* (DOA), i.e., the direction from which the acoustic energy arrives at the receiver. (In many cases, an estimated DOA can be used as an estimate of the source bearing with little error.) In what follows, it is assumed that the receiver is a line array, unless otherwise indicated. It will also be assumed for the DOA problem that the source is in the far field of the array, justifying the use of a plane-wave model for

the signal wavefront. This assumption breaks down for sonobuoy fields, where the sensors may be separated by over a kilometer and for which other techniques are used.

7.1.1 Broadband Signals

The *split-beam* correlator is a standard method for DOA estimation with a line array when the signal is broadband [4,33]. The array is split into two sub-arrays, which are individually beamformed. The time series from beam pairs pointed in the same direction are then cross-correlated to identify time delays, and from each time delay a bearing angle to the corresponding source is deduced. For tracking purposes, a Hilbert transformer may be inserted into one of the two beam channels to provide a zero-crossing instead of a peak in the correlogram.

One drawback of the split-beam correlator is the complexity of its implementation. Also, although the angle estimates provided by the split-beam method are close to optimum under certain conditions [33], the detection capability may be less than that obtained by beamforming the full aperture. A simpler method for estimating the DOA for a broadband source is to find the angular location of a peak in a BTI plot (see Section 5.2), if necessary refining the estimate by interpolating between adjacent beams.

7.1.2 Narrowband Signals

For a narrowband plane-wave signal at frequency ω impinging on a line array, the Fourier coefficients $X_n(\omega)$ as a function of sensor number n have the form of a complex sinusoid $A\exp(-in\phi_0)$, where $\phi_0 \equiv kd\sin\theta_0$ is a normalized spatial frequency. DOA estimation for the line array is therefore mathematically equivalent to frequency estimation for a complex-valued time series, and spectral-analysis methods can be brought to bear [7,34]. The data record length is the number of sensors N, and thus is short even in the best case ($N \sim 100$). The conventional processing in Eq. (3) is the discrete Fourier transform (DFT); that is, $|B|^2$ as a function of $\phi \equiv kd\sin\theta$ is the periodogram, and will attain a local maximum when $\theta = \theta_0$. For a single source, or multiple sources that are well separated in angle, DOA estimates derived from the periodogram will approach the limit of achievable accuracy as given by the CRLB [35,36].

As the angular separation between two sources decreases, the DFT method will eventually fail to resolve them, and *high-resolution* methods can be used to provide DOA estimates at closer spacings. The MVDR algorithm appears to have a greater ability to resolve two closely spaced sources than conventional processing; indeed, Capon [37] considered the method primarily as means to improve resolution in frequency–wave number spectral estimation. Several modern methods of

spectral analysis [38,39] appear to have better resolution than the periodogram on short data records (and at high SNR), although caution must be exercised in evaluating their resolution capability [40]. The maximum entropy method, based on autoregression modeling, was the focus of much early research. More recent methods are based on eigenanalysis of the sample matrix $\widehat{\mathbf{R}}$; examples are the MUSIC (multiple signal classification) and ESPRIT (estimation of signal parameter via rotational invariance techniques) algorithms [8]. The *IEEE Transactions on Signal Processing* is one of the main journals for current research in this area.

7.1.3 Synthetic Aperture Processing

Work has been done on applying the principle of synthetic aperture processing to passive sonar by coherently combining data at a pre-detection stage [41]; one goal is to increase bearing resolution by synthesizing a longer array aperture. Although well established for active processing, particularly in radar, synthetic aperture techniques are more controversial for passive sonar, where the large number of unknown source parameters may vitiate the technique (especially when the source itself is moving). Theoretical bounds have been derived on the accuracy of frequency and bearing estimation that can be achieved with a moving line array [42], and it is found that array motion is advantageous only in certain cases.

7.2 Range Estimation

Determining the range to a radiating acoustic source through purely passive means is difficult, for in general there is no way to determine the propagation time from source to receiver. The numerous ranging methods that have been developed are so varied that it is hard to categorize them systematically; broadly speaking, however, the range information is derived by (1) using a spatially extended receiver to exploit the structure of the acoustic pressure field produced by a radiating source, or (2) exploiting receiver or source motion. Some methods can provide a range estimate on the basis of a small number of data snapshots, whereas those requiring motion may take much longer. The primary ranging techniques are now described.

7.2.1 Wave front Curvature Ranging

Used with a spatially extended array, wave front curvature ranging is simple in concept: the acoustic wave front is assumed to undergo spherical spreading, and the source position is estimated from the time delays measured across the receiver aperture. For a broadband source, one mechanization is to use three sub-arrays that are arranged co-linearly in the horizontal plane; Figure 10(a) shows the arrangement for a submarine-based

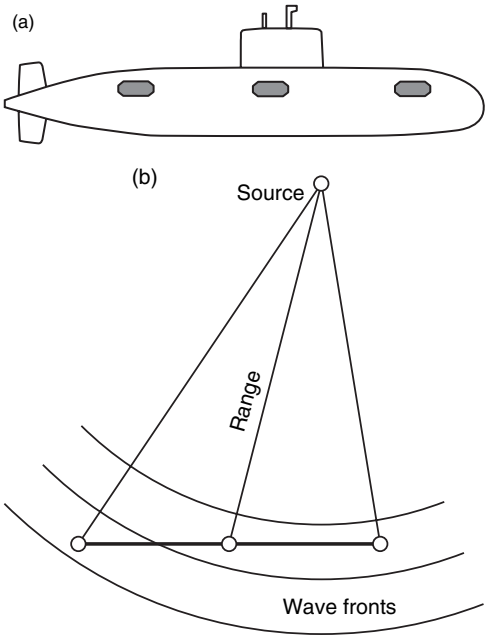


FIGURE 10 Passive ranging via wave front curvature. (**a**) Sensor placement on a submarine, and (**b**) ranging to the source by triangulation (plan view).

system. The relative time delays between the signal arrivals at pairs of sub-arrays are estimated, and from these the range can be deduced by triangulation, Figure 10(b). A detailed treatment of this ranging method, with references to the literature, can be found in [43]. For the aperture lengths that are used in practice, this form of wave front curvature ranging is accurate only for near-range sources. If the source is narrowband, estimating the time delays becomes difficult and other realizations of wave front curvature ranging must be sought; one approach is to use a narrowband focused beamformer.

7.2.2 Matched-Field Processing

The advantage of the spherical wave front model for ranging is its analytical simplicity. In reality, however, the pressure field generated across an array by a radiating source often has a complicated structure, and *matched-field processing* (MFP) is a technique for exploiting that structure [44,45]. MFP can be viewed as a generalized beamforming technique in which the weight vector **w** is derived by solving the acoustic wave equation with the appropriate boundary conditions (**w** is often called the *replica* vector). We consider MFP for passive narrowband local-

ization using a vertical line array, for which the replica vector generated by the acoustic propagation model is parameterized in range and depth, $\mathbf{w} = \mathbf{w}(r, z)$. In the most common implementation, an *ambiguity surface* is generated by computing the normalized quadratic form $B(r, z) = (\mathbf{w}^* \hat{\mathbf{R}} \mathbf{w})/(\mathbf{w}^* \mathbf{w})$ on a two-dimensional grid of hypothesized range and depth values, where as before $\hat{\mathbf{R}}$ is the sample spectral matrix. The ambiguity surface is a measure of the correlation between the measured and predicted acoustic pressure fields, and will exhibit a peak where the hypothesized values of range and depth correspond to the actual source values. The method provides an estimate not only of range, but of depth as well; note, however, that an accurate knowledge of the propagation environment is required.

7.2.3 Multipath Ranging

Like MFP, multipath ranging exploits the structure of sound propagation in the ocean, but is based on a simpler model that requires much less knowledge of the environment: the model assumes that the sound radiated from a distant source arrives at the sensor via multiple paths having simple geometric interpretations (e.g., surface and bottom reflections). The propagation time along each path depends on the range and depth of the source, as well as on the characteristics of the ocean waveguide. For a broadband signal, the cross-correlation between a pair of sensors will show peaks at the relative arrival times. When the different arrival paths involved in the propagation can be identified, the estimated time delays can be used to localize the source; in [46], a source track is produced from a time sequence of computed correlograms. In some scenarios, it may be adequate to compute the autocorrelation at a single sensor [4].

7.2.4 Target Motion Analysis

Target motion analysis (TMA) is a method that builds up localization information over time through the motion of the source (i.e., the target) and/or the receiver. The passive receiver makes direct measurements of such quantities as bearing or frequency (see Table 1) and, through the time evolution of these data, implicitly observes such state parameters of the target as its position and velocity. Which state parameters are observable depends on the properties of the specific measurement set [47]. Bearings-only TMA [48], for example, does not provide a unique ranging solution unless the receiver itself undergoes a change in motion. In contrast, if a stationary omnidirectional sensor observes the Doppler shift of a narrowband source as it moves through the *closest point of approach* (CPA), the target range and speed at CPA can be estimated. Clearly the TMA estimation problem depends strongly on the type of measurement data; it

TABLE 1 Standard types of data used in target motion analysis. For a moving narrowband source the frequency observed by the receiver changes with time owing to the Doppler effect

	Omnidirectional receiver	Directive receiver
Broadband source	–	Bearing
Narrowband source	Frequency	Bearing and frequency

also depends on the model for the target motion (constant velocity or maneuvering) and other specific problem features.

Since noise interferes with the measurement process, an additional determinant of TMA performance is the accuracy of the bearing and/or frequency estimates that are used in forming the TMA solution. To aid the sonar operator, automatic trackers have been developed that extract the bearing and/or frequency data that will be fed into the localization algorithm; they are sometimes given such names as *bearing trackers, frequency trackers*, or *signal followers*. For noisy measurement data, the CRLB can be used to bound the achievable localization accuracy for a particular scenario.

TMA processing may be organized as a *batch* process, wherein a large number of bearing and/or frequency measurements are first accumulated and subsequently fed as a batch into the localization solver. The solution may be found by least-squares techniques, grid searches, or other methods. Once an initial localization estimate has been made, it is convenient to update recursively the estimated source position instead of repeatedly running a batch solver on data sets of steadily increasing size. Consequently, extensive work has been done on target tracking based on such techniques as extended Kalman filtering [43,49].

References

1. A.B. Baggeroer, "Sonar Signal Processing," in A.V. Oppenheim (ed.), *Applications of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1978.
2. W.C. Knight, R.G. Pridham and S.M. Kay, "Digital Signal Processing for Sonar", Proc. IEEE, vol. 69, no. 11, 1981, pp. 1451–1506.
3. R.J. Urick, *Principles of Underwater Sound*. 3rd edition. New York: McGraw-Hill, 1983.
4. W.S. Burdic, *Underwater Acoustic System Analysis*. 2nd edition. Englewood Cliffs, NJ: Prentice-Hall, 1990. Reprinted 2002 by Peninsula Publishing, Los Altos, CA.

5. R.O. Nielsen, *Sonar Signal Processing*. Norwood, MA: Artech House, 1991.
6. D. Ross, *Mechanics of Underwater Noise*. New York: Pergamon Press, 1976.
7. D.H. Johnson and D.E. Dudgeon, *Array Signal Processing*. Upper Saddle River, NJ: PTR Prentice-Hall, 1993.
8. H.L. van Trees, *Optimum Array Processing*. New York: John Wiley, 2002.
9. R.A. Mucci, "A comparison of efficient beamforming algorithms," IEEE Trans. Acoust., Speech, and Sig. Proc., vol. ASSP-32, 1984, pp. 548–558.
10. R.G. Pridham and R.A. Mucci, "A novel approach to digital beamforming," J. Acoust. Soc. Am., vol. 63, 1978, pp. 425–434.
11. B.H. Maranda, "Efficient digital beamforming in the frequency domain," J. Acoust. Soc. Am., vol. 86, 1989, pp. 1813–1819.
12. M.E. Weber and R. Heisler, "A frequency-domain beamforming algorithm for wideband, coherent signal processing," J. Acoust. Soc. Am., vol. 76, 1984, pp. 1132–1144.
13. D.J. Edelblute, J.M. Fisk, and G.L. Kinnison, "Criteria for optimum-signal-detection theory for arrays," J. Acoust. Soc. Am., vol. 41, 1967, pp. 199–205.
14. R.A. Monzingo and T.W. Miller, *Introduction to Adaptive Arrays*. New York: John Wiley, 1980.
15. N.L. Owsley, "Sonar Array Processing", in S. Haykin (ed.), *Array Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
16. N.L. Owsley, "Overview of adaptive array processing," in H.G. Urban (ed.) *Adaptive Methods in Underwater Acoustics*, NATO ASI Series. Dordrecht, Holland: D. Reidel, 1985.
17. S. Stergiopoulos, "Implementation of adaptive and synthetic-aperture processing schemes in integrated active-passive sonar systems," Proc. IEEE, vol. 86, 1998, pp. 358–396.
18. J.B. Lewis and P.M. Schultheiss, "Optimum and conventional detection using a linear array," J. Acoust. Soc. Am., vol. 49, 1971, pp. 1083–1091.
19. H. Cox, "Line array performance when the signal coherence is spatially dependent," J. Acoust. Soc. Am., vol. 54, 1973, pp. 1743–1746.
20. P.D. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," IEEE Trans. on Audio and Electroacoustics, vol. AU-15, June 1967, pp. 70–73.
21. W.A. Struzinski and E.D. Lowe, "A performance comparison of four noise background normalization schemes proposed for signal detection systems," J. Acoust. Soc. Am., vol. 76, 1984, pp. 1738–1742.
22. W.A. Struzinski and E.D. Lowe, "The effect of improper normalization on the performance of an automated energy detector," J. Acoust. Soc. Am., vol. 78, 1985, pp. 936–941.
23. H.L. van Trees, *Detection, Estimation, and Modulation Theory*. Pt. III, Radar-Sonar Signal Processing and Gaussian Signals in Noise. New York: John Wiley, 1971.

24. P.M. Schultheiss and E. Weinstein, "Detection and estimation: a summary of results," in L. Bjorno (ed.), *Underwater Acoustics and Signal Processing*. Dordrecht, Holland: D. Reidel Publishing, 1981.
25. R.O. Nielsen, "Cramer-Rao lower bounds for sonar broad-band modulation parameters," IEEE J. Oceanic Eng., vol. 24, July 1999, pp. 285–290.
26. H. Urkowitz, "Energy detection of unknown deterministic signals," Proc. IEEE, vol. 55, 1967, pp. 523–531.
27. Z. Wang and P. Willett, "A performance study of some transient detectors," IEEE Trans. Sig. Proc., vol. 48, 2000, pp. 2682–2685.
28. W.R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation," J. Acoust. Soc. Am., vol. 58, 1975, pp. 201–207.
29. G.C. Carter, "Variance bounds for passively locating an acoustic source with a symmetric line array," J. Acoust. Soc. Am., vol. 62, 1977, pp. 922–926.
30. G.C. Carter, "Coherence and time delay estimation," Proc. IEEE, vol. 75, 1987, pp. 236–255.
31. G. Clifford Carter (ed.), *Coherence and Time Delay Estimation*. Piscataway, NJ: IEEE Press, 1993.
32. S.-K. Chow and P.M. Schultheiss, "Delay estimation using narrow-band processes," IEEE Trans. Acoust., Speech, Sig. Proc., vol. ASSP-29, 1981, pp. 478–484.
33. V.H. MacDonald and P.M. Schultheiss, "Optimum passive bearing estimation in a spatially incoherent noise environment," J. Acoust. Soc. Am., vol. 46, 1969, pp. 37–43.
34. D.H. Johnson, "The application of spectral estimation methods to bearing estimation problems," Proc. IEEE, vol. 70, 1982, pp. 1018–1028.
35. D.C. Rife and R.R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," IEEE Trans. Info. Th., vol. IT-20, 1974, pp. 591–598.
36. D.C. Rife and R.R. Boorstyn, "Multiple tone parameter estimation from discrete-time observations," Bell System Tech. J., vol. 55, 1976, pp. 1389–1410.
37. J. Capon, "High-resolution frequency-wavenumber spectrum analysis," Proc. IEEE, vol. 57, 1969, pp. 1408–1418.
38. S.L. Marple, *Digital Spectral Analysis with Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
39. S.M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
40. S. Kay and C. Demeure, "The high-resolution spectrum estimator – a subjective entity," Proc. IEEE, vol. 72, 1984, pp. 1815–1816.
41. IEEE J. Oceanic Eng., vol. 17, no. 1, January 1992. (Special issue on acoustic synthetic aperture processing.)
42. G.S. Edelson and D.W. Tufts, "On the ability to estimate narrow-band signal parameters using towed arrays," in [41], pp. 48–61.
43. J.C. Hassab, *Underwater Signal and Data Processing*. Boca Raton, FL: CRC Press, 1989.

44. A.B. Baggeroer, W.A. Kuperman, and P.N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," IEEE J. Oceanic Eng., vol. 18, 1993, pp. 401–424.
45. IEEE J. Oceanic Eng., vol. 18, no. 3, July 1993. (Special issue on detection and estimation in matched-field processing.)
46. E.K. Westwood and D.P. Knobles, "Source track localization via multipath correlation matching," J. Acoust. Soc. Am., vol. 102, 1997, pp. 2645–2654.
47. C. Jauffret and D. Pillon, "Observability in passive target motion analysis," IEEE Trans. Aerosp. and Electr. Systems, vol. 32, 1996, pp. 1290–1300.
48. S.C. Nardone, A.G. Lindgren, and K.F. Gong, "Fundamental properties and performance of conventional bearings-only target motion analysis," IEEE Trans. Auto. Control, vol. AC-29, 1984, pp. 775–787.
49. Y. Bar-Shalom and T.E. Fortmann, *Tracking and Data Association*. Orlando, Florida: Academic Press, 1988.

98

Transducers for Sonar Systems

Kim Benjamin

Naval Undersea Warfare Center, Newport, RI, USA

1 • Transduction Materials

1.1 Non-centro-symmetry and the Curie Temperature

Strictly speaking, a transducer is a device that converts energy from one form to another. For sonar transducers, electrical energy is converted to mechanical energy and vice versa through the use of various natural and man-made materials that exhibit piezoelectric or magnetostrictive properties. This section contains a brief discussion of the crystallographic asymmetry of these interesting materials, and how such materials can be exploited to form acoustic transducers.

The piezoelectric phenomenon, first discovered in 1880 by Jacques and Pierre Curie, was observed in naturally occurring single crystals such as quartz. In more recent times, it was found that piezoelectricity could be induced in polycrystalline (ferroelectric) materials by applying a high static electric field. Whether natural or man-made, a transduction material consists of a crystalline arrangement of atoms; both metal and non-metal. Within their internal structures these atoms either share electrons (covalent bond) or transfer them permanently (ionic bond) [1]. In the latter case, the atoms become ionized carrying a net charge plus (+) or minus (−). This resulting offset between charges, or dipole distance, will change when external forces are applied; either mechanical or electrical. The applied forces can be static, or dynamic.

Static pressure transducers and displacement actuators are a good example of the former and sonar transducers of the latter.

Below the temperature known as the *Curie point* of the material, the unit cell exhibits the non-centro-symmetry described above. However above that temperature, its crystal lattice structure changes from asymmetric to symmetric and the piezoelectric effect is lost. Cooling back below the *Curie point* temperature will re-establish the electric dipoles but their orientation will be random and the material will need to be re-polarized in a large static electric field to regain the effect.

In quartz the dipole charges are permanently oriented. For man-made piezoceramics, the dipole charges can be reversed through the applied polarization field. These types of materials are called ferroelectric. This name comes about from the hysteresis effect [1] common to ferromagnetic materials and has nothing to do with iron. A list of PZT (lead zinconium titanate) transduction materials along with their important properties is can be found in references [2–4] as well as in the vendor catalogs [5] of various PZT products.

1.2 Solid PZT Elements: Plates, Cylinders, and Spheres

For practical sonar devices, the intrinsic transduction found in the materials described above is enhanced through geometrical design or by the addition of passive materials that cause the structure to resonate in a preferred vibrational mode at particular frequency. One of the most common materials used in current sonar systems is piezoceramic which as its name implies is processed and produced using techniques similar to those used to make ceramic pottery. There many varieties of PZT available and the electromechanical properties of the individual types are varied by controlling the amount and organization of the constituent elements (i.e. lead, zircon, titanium, barium, oxygen, etc.).

The basic raw materials are ground into fine powder, mixed with a plasticizer (water, wax, etc.), shaped (plate, tube, hemisphere), dried, and fired in a kiln. The art of fabrication involves controlling the dimensional changes during the thermal cycling since the material (like pottery) shrinks during the drying process. Through the molding and finish grinding processes the piezoceramic can be formed into basic geometric shapes as shown in Figure 1. The most common are plates, rings, tubes, and hemispheres; the latter are used to form spheres.

In order to distribute the electric charge evenly over the formed ceramic part, electrodes consisting of a silver loaded glazing material are painted over the piezoceramic surface and baked

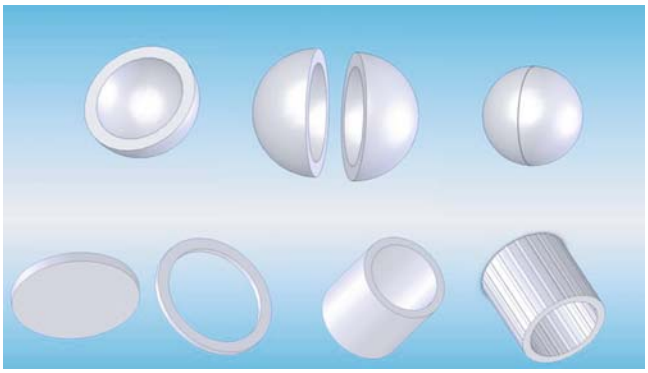


FIGURE 1 Examples of solid PZT elements.

on in a kiln. The striped tube, farthest to the right of Figure 1, shows the inner and outer longitudinal electrodes that are used to tangentially polarize the cylinder; thus enhancing the circumferential excitation.

After manufacture and plating, the part is subjected to a very large electric field of approximately 2 MV/m (50 V/mil) in oil, or some other electrically inert material, to impart a remnant polarization. The remnant polarization and the other electromechanical properties, (i.e. dielectric constant, electromechanical coupling coefficient, compliance, etc.), change logarithmically at a rate of roughly 2% per decade time [3].

In general, when designing transducers with solid PZT elements, the mechanical dimensions in the three orthogonal directions must be chosen such that the resulting resonant vibrational modes are well spaced in frequency. In the case of a rectangular plate for example, the length and width modal resonance frequencies, as well as their first and second harmonics, should be well separated from the desired thickness resonance to insure a smoothly varying response over the operational frequency range. For circular plates, tubes, and spheres the unwanted modes are usually related to the radial dimension.

The modal resonance frequencies are determined from the mechanical dimensions and the speed of sound in the piezoceramic. The sound speed is related to the elastic modulus which is both anisotropic and complex. The imaginary part of the material's modulus is proportional to the mechanical losses and is a function of frequency. Figure 2 shows a typical equivalent circuit for a thickness expander disc [6,7]. Here the purely electrical portion (applied voltage, V and blocked capacitance, C_0) is coupled to the purely mechanical portion (Z_1 and Z_2) with an ideal transformer of N turns where N is determined from the piezoelectric material's

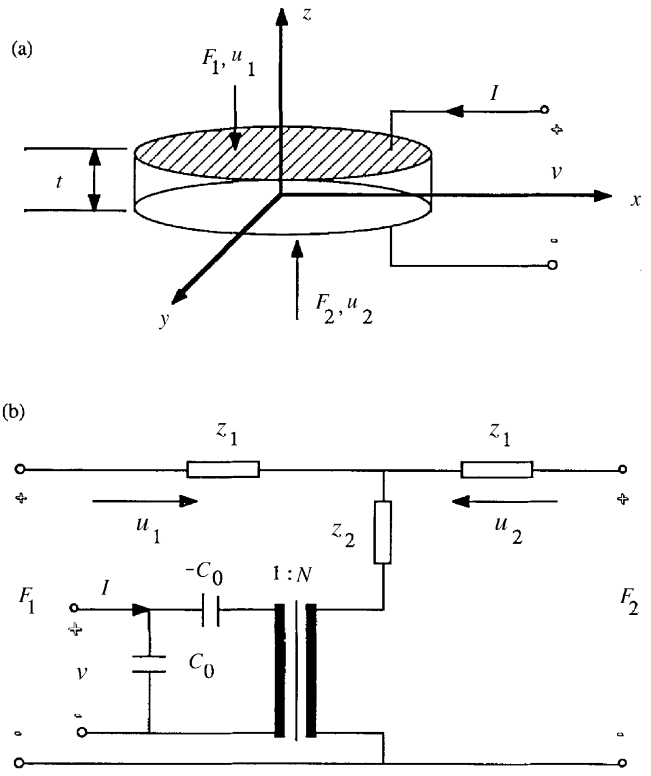


FIGURE 2 Distributed parameter equivalent circuit used for modeling the vibratory response of a thickness expander disc [6].

coupling, dielectric, and stiffness coefficients. The quantities F_1, u_1, F_2, and u_2 represent the forces and velocities of the major vibrating faces respectively. Near the modal resonance the mechanical components such as mass, stiffness, and damping can be represented by simple electrical analogues such as inductors, capacitors, and resistors respectively. However for an exact representation over many vibrational modes, the mass, stiffness, and damping are modeled with trigonometric impedance functions that include complex sound speed quantities based on the complex modulus discussed above. References [6] through [9] offer the reader a more comprehensive treatment of equivalent circuits and there application to transducers.

1.3 Piezoelectric Polymers

Piezoelectric polymers such as polyvinylidene fluoride (PVDF) were introduced in the 1980s [10–13] and have found their use in sonar primarily as broadband large area receivers [14,15]. Being pure polymer, with intrinsic low modulus and high mechanical

loss, it is an attractive transducer material providing suppression of unwanted lateral elastic waves. Its acoustic impedance is well matched to water and it can be fabricated in large sheets having nominal thicknesses that range between 25 and 500 μm (0.001–0.02 in). Figure 3 shows a photograph of the processing apparatus invented and used for PVDF production by Raytheon Company's research division [16,17].

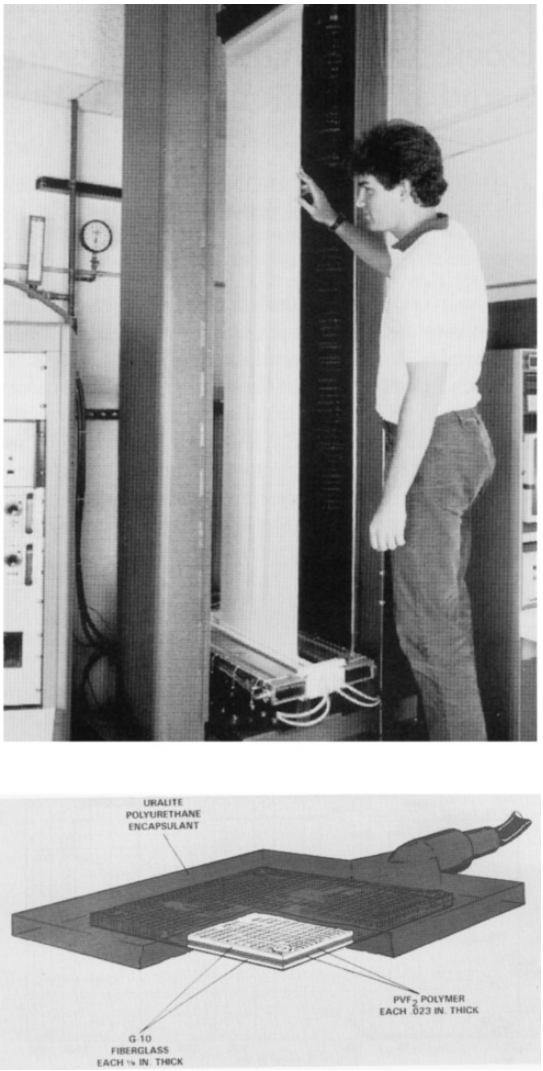


FIGURE 3 PVDF fabrication (*above*) (photo courtesy: Raytheon Co. Lexington, MA), example of large area sensor with stiffeners (*below*) (from Reference [17]).

In biomedical devices [18,19], operating in the megahertz range, the thinner materials are formed into membranes and sections are metalized to form electrodes. For lower frequency applications, like sonar, it can be applied to tubes and plates [20] and [21]. Here the piezopolymer follows the modal vibration of the structure it is bonded to in order to enhance displacement amplitude or to suppress flexural shear modes.

The major advantage of PVDF is its relatively low cost per area. This is due to the method of fabrication and the raw materials required. However, the fabrication of an area sensor requires other passive components (stiffeners) to eliminate extraneous vibrational modes. Typical large area array sensors consist of two or more copper-plated PVDF sheets aligned with their stretch directions at right angles, and sandwiched between metal, or glass-reinforced epoxy stiffener plates as shown in Figure 3.

Current research is exploring the use of electron irradiation techniques to increase the electromechanical coupling and relative dielectric constant of poly(vinylidene fluoride-trifluoroethylene) copolymer [22] which is another type of ferroelectric polymer.

1.4 Piezoceramic–Polymer Composites

Piezoelectric polymer composites were first invented and developed by Robert E. Newnham and Eric L. Cross at the Materials Research Laboratory at Pennsylvania State University [23]. Generally referred to as "piezocomposites", they are a combination of PZT structures arranged within a polymer. The polymer itself could be a composite of various polymers in a specific arrangement as well.

Piezocomposite found immediate use in the biomedical community [24–26], where they enhanced the performance of ultrasonic single elements and arrays that are used for tissue imaging. The major improvement piezocomposites have over their solid PZT counterparts is the suppression of elastic waves in the material that cause unwanted response characteristics. The mechanical decoupling of the PZT structures occurs due to the attributes of the relatively soft polymer matrix, such as high loss and low modulus. This allows the active layer to have lateral dimensions that are several wavelengths long while still providing a resonance-free response above and below its thickness mode frequency. Unlike the pure piezopolymers discussed above which are essentially receive only, piezocomposites can have thicknesses appropriate for transmit as well as receive applications in the sonar frequency range.

Early 1-3 and 2-2 piezocomposites were fabricated by dicing and backfilling solid PZT. In the case of a 1-3 composite, the active material is partitioned as rods and the ceramic material extends only one direction (the thickness direction) and the polymer matrix extends through out the substrate in three directions. For the 2-2 case, the ceramic consists of a series of plates with the polymer matrix sandwiched between them. Thus both ceramic and polymer extend in two directions. The reader is referred to reference [23] for the description of composites with other types of connectivity (i.e. 0-3, 3-3, 3-1, etc.)

Although this approach worked well for the small-sized biomedical transducers the dice and fill approach was inefficient for larger apertures used in sonar and limited the shape and relative placement of rods. One of the major advancements in 1-3 piezocomposite technology was the development of injection molded piezocomposites [27]. Using techniques employed the plastics industry, PZT could now be formed into a variety of useful shapes in a cost-effective way. Figure 4 shows an injection molded PZT perform. These parts could be tiled and backfilled with polymer mixtures to create large sheets of 1-3 piezocomposite.

1.5 Ferroelectric Single Crystals

The recent introduction and characterization of single-crystal piezoelectric materials has generated much interest in the sonar transducer community and promises to significantly improve the performance of current systems [28,29]. These materials consist of single crystals (as opposed to polycrystalline) of lead magnesium niobate (PMN) or lead zinc niobate (PZN) in solid

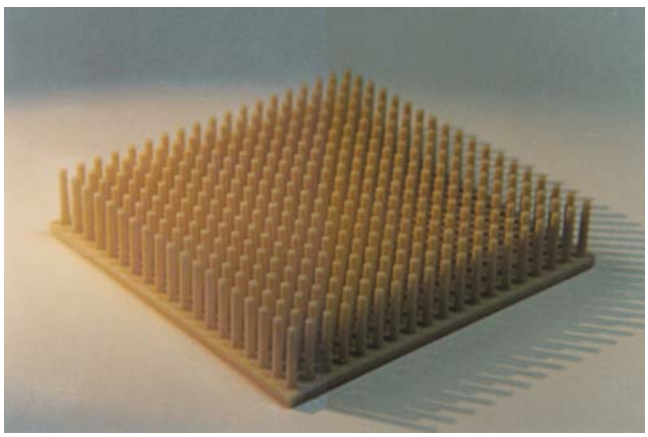


FIGURE 4 Injection-molded piezoceramic pre-form $5 \times 5 \times .9$ cm ($2 \times 2 \times 0.37$ in) (photo courtesy: Materials Systems Inc. Littleton, MA).

solution with lead titanate (PT) (i.e. PMN–PT or PZN-PT). The materials have greater shear and thickness mode coupling coefficients than their polycrystalline piezoceramic counterparts [30]. This and their low modulus of elasticity make it attractive for high-power, low-frequency acoustic transmitters.

2 • Conventional Designs and Modeling Approaches

2.1 Longitudinal Resonators

The longitudinal resonator [31] is one of the most common transducer configurations for generating high sonic intensities over a relatively narrow band. They range in size and complexity based on their application and can be designed to operate in both the sonar and ultrasonic frequency regimes. Their high response amplitude is accomplished by increasing the mechanical quality factor (Q) of the operational resonance. Higher frequency designs (50 kHz and upward), consist of laminated discs or plates of PZT, plastic, and metal (i.e. steel, aluminum, tungsten) [32].

2.2 Spring–Mass (Tonpilz) Designs

The tonpilz designs were developed for frequency ranges (1–50 kHz) lower than practical for a pure longitudinal resonator. This class of transducer provides huge dynamic displacements over a narrow band while occupying a small package size. The word *tonpilz* is the conjunction of two German words *ton* (sound) and *pilz* (mushroom) and physically depicts this type of underwater mechanical oscillator [33].

In general, active piezoelectric [34,35] or magnetostrictive [36,37] layers are combined with passive mechanical components to attain the high-amplitude resonance condition. The active elements, plates or rings, are stacked between two masses, and are kept in compression between the two masses with a slender high strength rod (tierod). The weight ratio between the head (radiating) and tail (stationary) masses along with the stiffness of the stack, which acts as a spring, govern the resonance frequency and Q of this type of oscillator. Figure 5 shows a cross-sectional view of a tonpilz active sub-assembly packaged along with a tuning transformer for underwater use [38].

Over the years, there have been many innovative design modifications to the basic tonpilz configuration. For example, the use of mechanical head flexure can be employed to broaden the element's bandwidth [39]. Typically, the head mass is designed to be light and stiff with its bending modes placed well above the operating frequency band. Yet, exciting the head flexural resonance can in some cases produce a second usable frequency band of operation.

Another interesting approach to increasing tonpilz bandwidth is the multi-resonant design. Here the active driver could consist

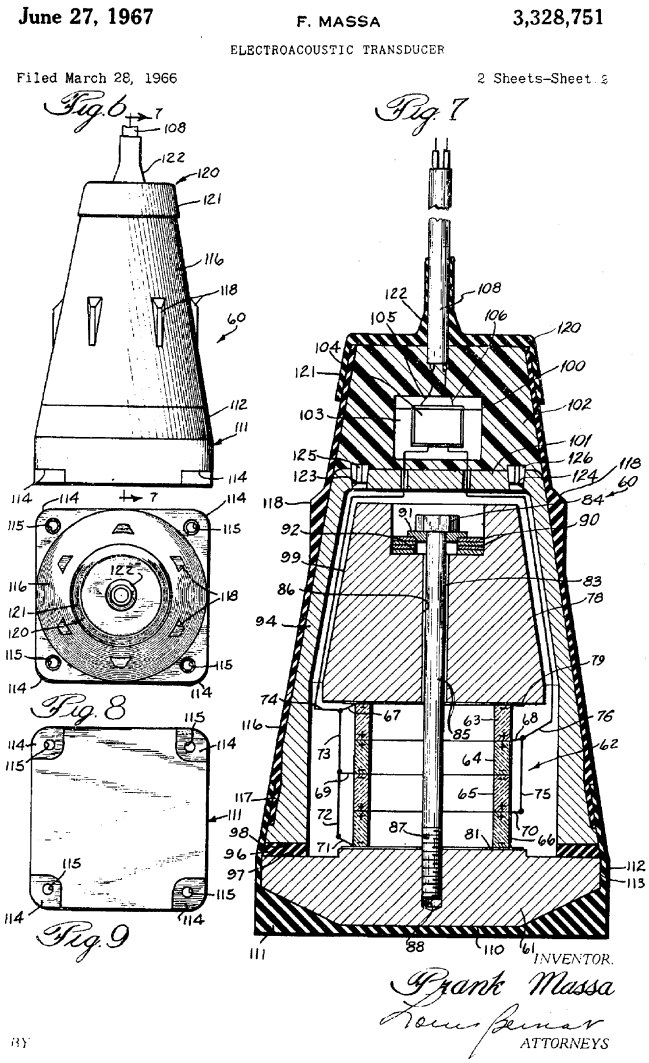


FIGURE 5 Early tonpilz transducer for US Navy applications [38].

of two different PZT stacks tuned to two separate resonance frequencies as shown in Figure 6 and explained further in references [40] and [41], or the stack may be a clever combination of different drive materials as demonstrated by Butler and Tito [42], for example, piezoceramic and the high strain magnetostrictive rare-earth iron compound TERFENOL developed by Clark [43].

2.3 Flexural Resonators

As the operational frequency continues to fall (100 Hz to 1 kHz) transducer designs move from piston sources to flexural devices which rely on a flexural resonance of the radiating structure.

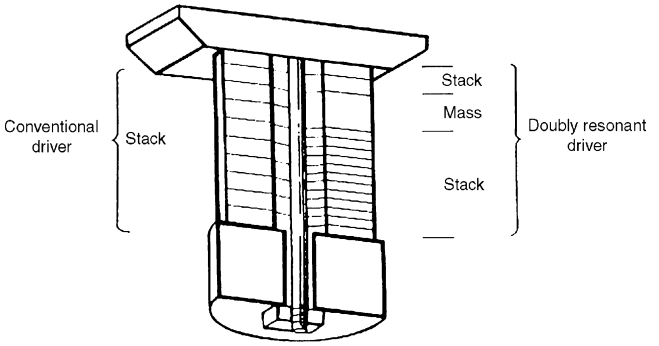


FIGURE 6 Example of a doubly resonant tonpilz element (from Reference [40]).

The lower flexural modulus (and associated lower sound speed of the flexing material), allows one to construct a transducer that delivers low frequency and high displacement within a relatively small package size. Flexural resonator designs include *bender bars* [44–46], *flextensionals* [47], [48–51], *slotted cylinders* [52], and *barrel staves* [53]. In all cases, a piezoceramic material excites a mechanical resonance, in the transducer's flexing component(s) and it is the radiating plate or shell geometry and material properties that primarily govern the fundamental mode. The various designs described below are examples of such combinations of active and passive components that achieve lower frequency flexural vibrations.

2.3.1 Bender Bars

The bender bar transducer uses piezoceramic lateral strains to excite the flexural mode of a rectangular plate or circular disc. The dynamic flexing occurs since the exciting strain is offset from the nodal mid-plane of the composite laminations. In the case shown in Figure 7, stacks of PZT bars are placed on one or both sides of a flexible beam [45]. As the PZT bars expand and contract at the first bending mode the resulting motion produces high volume displacement at low frequency. The mounting attachments simulate free, clamped, or simply supported edge boundary conditions. There are bi-laminar or tri-laminar bender bar configurations. These transducers are air-backed. For a comprehensive theoretical discussion of electroacoustic bender bar transducers, the reader is referred to reference [46] by Woollett.

2.3.2 Flextensionals

There are five different classes of flextensional transducers and a descriptive list can be found in reference [48] by Rolt. In general, a stack of PZT bars or rings drive an enclosing shell.

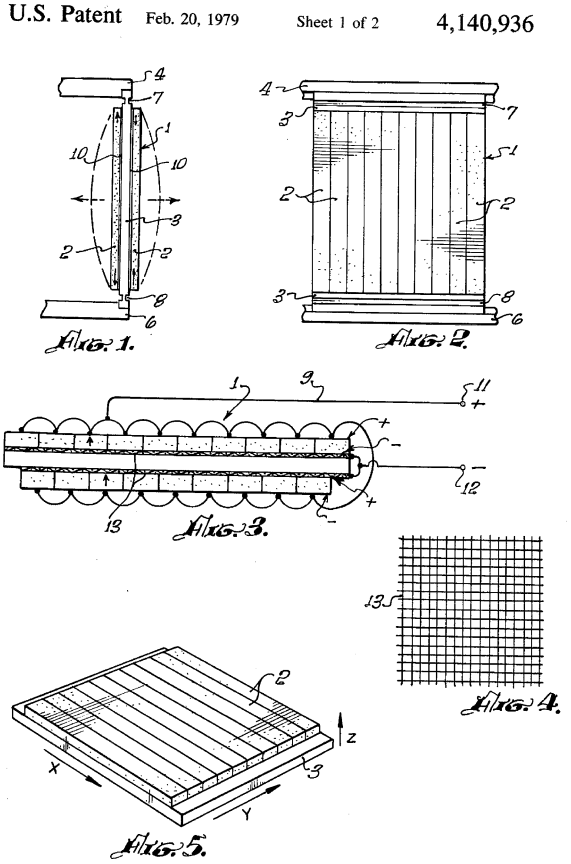


FIGURE 7 Bender bar configurations [45].

For efficient acoustic radiation the shell cross sections are oval rather than circular. Shell materials typically consist of metal or glass reinforced epoxy depending on the depth requirement. In all cases, these devices operate with air behind the radiating structure (air-backed) to attain efficient transfer of energy. The depth limitation results from the decrease of mechanical pre-stress on the PZT stack. High pre-stress is required to maintain the piezoceramic under compression at all times during its dynamic stress cycling. In some classes of flextensionals, a tierod will provide the compressional force; as in the tonpilz design discussed earlier. For others, the shell provides the necessary mechanical bias. Increasing the depth rating of these types of flexural transducers requires applying greater pre-stress on the

Oct. 4, 1966 W. J. TOULIS 3,277,433

FLEXURAL-EXTENSIONAL ELECTROMECHANICAL TRANSDUCER

Filed Oct. 17, 1963 3 Sheets-Sheet 3

FIG. 4

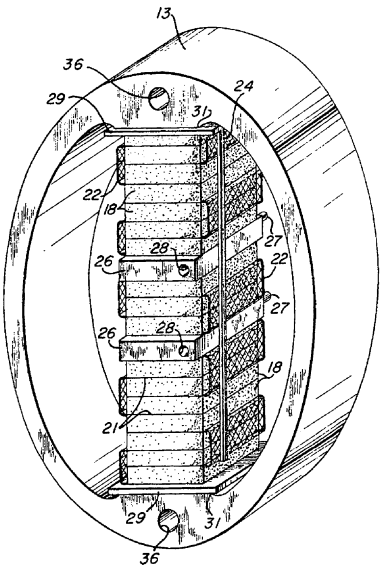


FIGURE 8 Illustration of a Class IV flextensional transducer [47] (end caps and boot not shown).

ceramic with the shell. Figure 8 shows one of the earliest flextensionals that was designed specifically for underwater use [47].

2.3.3 Slotted Cylinders

As shown in Figure 9, the original slotted cylinder design was conceived by Harris [52] and consists of a slotted tube driven by a slotted piezoceramic tube. Lateral displacement of the piezoceramic in the circumferential direction excites the slotted tube's fundamental resonance generating large amplitude vibrations at relatively low frequencies. The displacement amplitude is maximum at the slot with the node located along a line diametrically opposite the slot. The vibratory behavior follows a beam centrally supported along its width.

Unlike the devices shown in Harris' patent, slotted cylinders for sonar applications need to be air backed to attain efficient acoustic outputs and the design challenge is to apply water

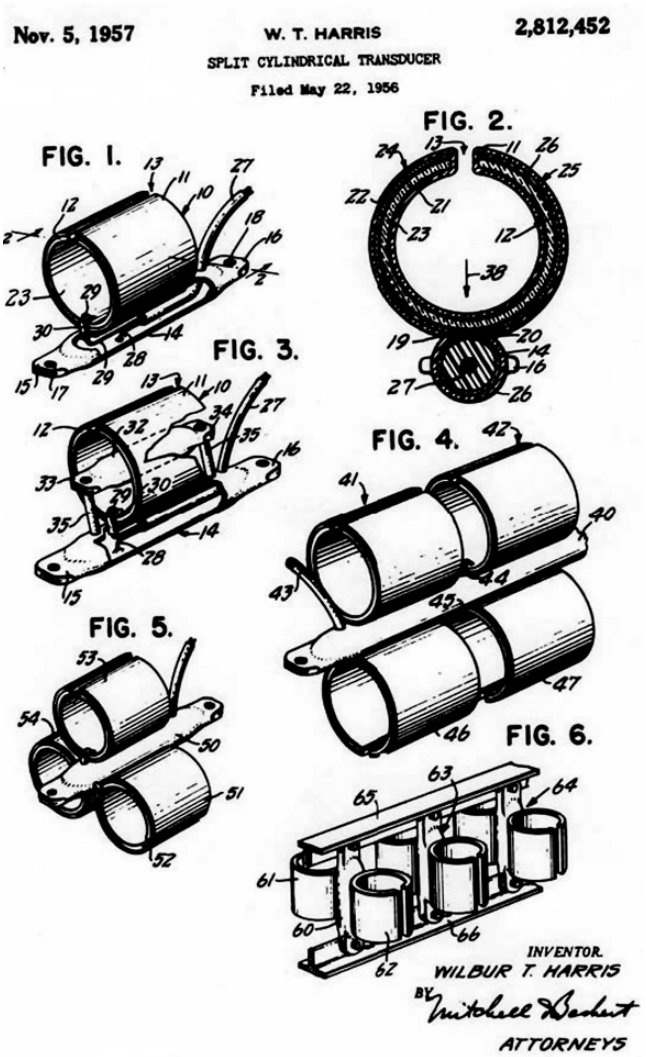


FIGURE 9 Slotted cylinder transducer [52].

proofing or a rubber boot that does not dampen the high Q oscillations. Like most air-backed flexural transducers, slotted cylinders are not useful for deep submergence applications due to the collapse of the slot under hydrostatic pressure.

2.3.4 Barrel Staves

The barrel stave transducer is actually a class III type flextensional consisting of an axially vibrating PZT stack that causes a shell to flex at its fundamental resonance. In some cases, the shell is approximated by a group of segments and a rubber boot is used for waterproofing [53] as shown in Figure 10. More recently folded shells [54] consisting of a single corrugated shell component have been developed. The idea patented by Purcell is illustrated in Figure 11 eliminates the need for a boot reducing the parts count and cost of the barrel stave design.

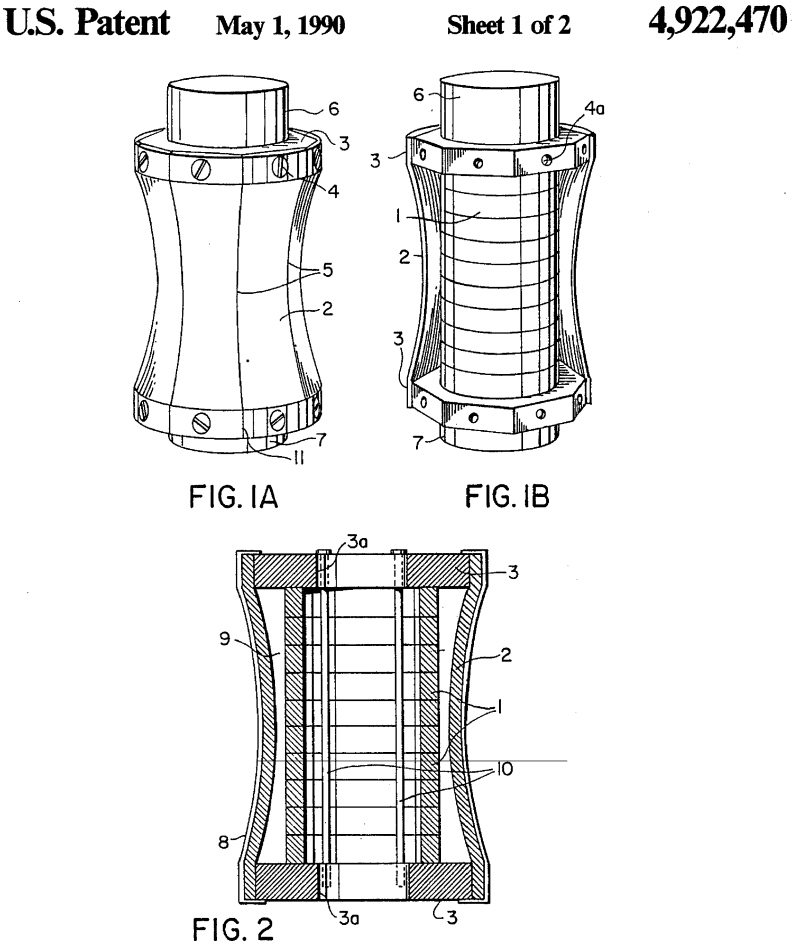


FIGURE 10 Barrel stave flextensional transducer [53].

US005805529A

United States Patent [19]

Purcell

[11] **Patent Number: 5,805,529**

[45] **Date of Patent: Sep. 8, 1998**

[54] **FOLDED SHELL PROJECTOR (FSP)**

[75] Inventor: **Christopher John A. Purcell**, Dartmouth, Canada

[73] Assignee: **Her Majesty the Queen in right of Canada, as represented by the Minister of National Defence of Her Majesty's Canadian Government**, Ottawa, Canada

[21] Appl. No.: **932,581**

[22] Filed: **Sep. 17, 1997**

[51] **Int. Cl.**[6] **H01V 7/00**

[52] **U.S. Cl.** **367/163**; 367/174

[58] **Field of Search** 367/163, 174, 367/142, 158, 159; 310/337

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,922,470	5/1990	McMahon	367/163
5,030,873	7/1991	Owen	367/162
5,136,556	8/1992	Obara	367/163

OTHER PUBLICATIONS

Frank R. Abbot "Compliant Housings to Enhance Low–Frequency Performance of Electrostrictive Transducers" Rept. NEL 574 US Navy Electronics Laboratory, San Diego, California 1955.

Dennis F. Jones, et al. "The Acoustic Perofrmance of a Class III Barrel–Stave Flextensional Projector", Proceeding of the 1996 Undersea Defence Technology Conference and Exhibition Nexus Media, Swanly, U.K. pp. 103–108.

Primary Examiner—Daniel T. Pihulic

Attorney, Agent, or Firm—Larson & Taylor

[57] **ABSTRACT**

An underwater acoustic projector comprising a pair of spaced apart end plates with an acoustic driver positioned between the end plates, the driver having smaller cross-sectional dimensions than the end plates. The end plate's edges are secured to an outer one-piece thin walled shell that provides a waterproof enclosure for the driver. That thin walled shell has a concavely inwardly bent surface between the end plates and a plurality of axially extending corrugations to provide a predetermined axial compliance and radial-to-axial transformation ratio.

22 Claims, 6 Drawing Sheets

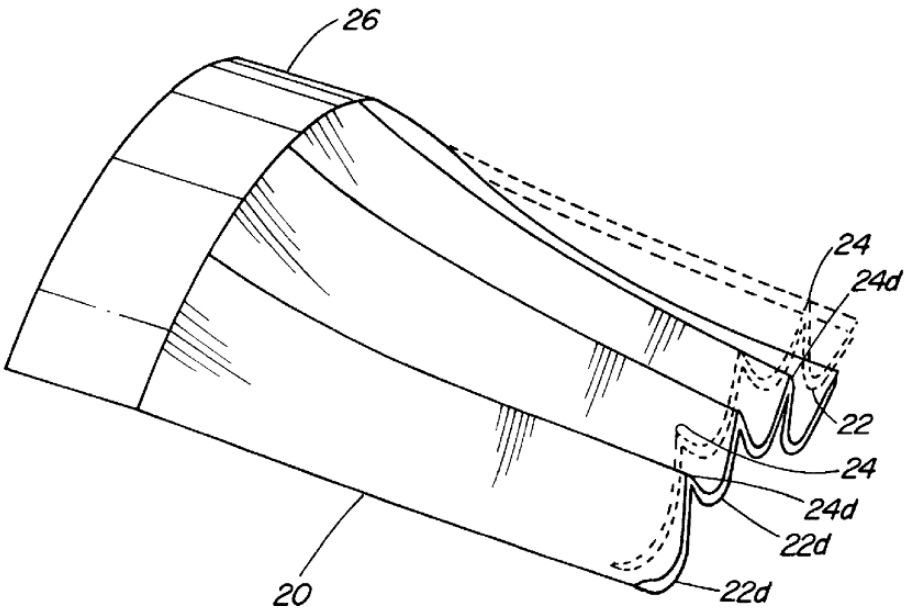


FIGURE 11 Folded shell for use on barrel stave transducer [54].

2.3.5 Design Approaches

Various modeling approaches can be used to model and design transducer elements. The most common are equivalent circuit analogs that are described in references [55–57] and finite element analysis (FEA) [58].

The basic equivalent circuit model design approach is well described by Wilson in reference [59]. The mechanical analogue circuits can contain simple lumped elements (i.e. R, L, and C) that are only valid around the first-order resonance, or can include frequency-dependent transmission line impedances that account for wave propagation effects over very wide frequency ranges.

The FEA models are useful when the component vibrations are not constrained to one-dimension. For example, a slender component such as a head mass or tierod will experience flexural or torsional vibrations that an equivalent circuit model will not directly predict. The flexural transducers, in particular flextensionals, have been modeled using a combination of both equivalent circuits and FEA [60,61].

3 • Broad Bandwidth Transducer Designs

3.1 Piezocomposite Transducers and Arrays

The sonar community has been investigating the use of 1-3 piezocomposite transducers and arrays for broadband applications since the mid-1980s [62,63]. As discussed earlier, the material's intrinsic lateral decoupling and low Q makes it attractive for generating complex wideband signals with a high degree of fidelity. In addition, large apertures, in some cases singly and doubly curved, can easily be realized through the use of piezocomposite technology [64].

To accurately produce wideband coded acoustic pulses, a transducer's impedance phase and amplitude need to be smoothly varying, thus allowing the electric circuitry that applies or receives the frequency-dependent signal to compensate for the response variation. Because large phase and amplitude excursions occur about the mechanical resonance of the transducer, standard sonar calibration hydrophones and projectors are operated below their first mechanical resonance.

Figure 12 shows two early US Navy calibration standards that provide uniformly varying amplitude and phase response over relatively large bandwidths. Interestingly, they look like 1-3 piezocomposite comprised of smaller solid PZT sub-sections (discs and plates). With their rubber acoustic windows in place, they are filled with oil to acoustically couple the active sections to the water.

Figure 13 shows a transducer consisting of significantly fewer parts making it simpler and less time consuming to construct. The

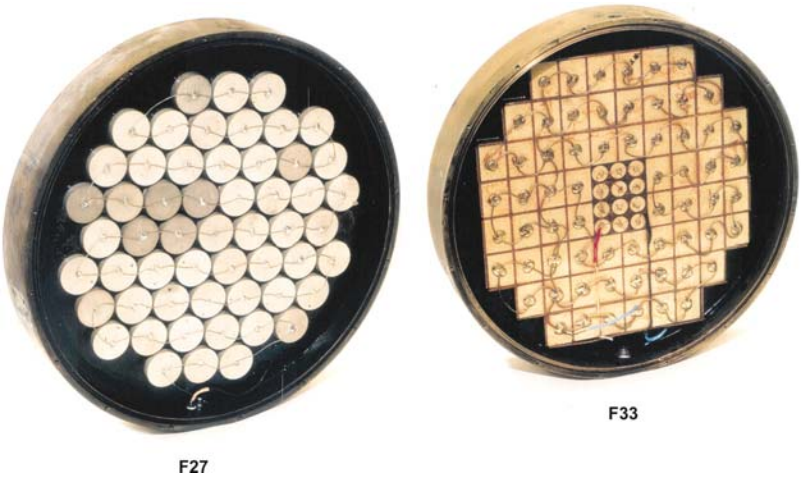


FIGURE 12 Conventional ceramic-based US Navy calibration transducer standards.



FIGURE 13 New 1-3 piezocomposite-based US Navy calibration transducer standard (unassembled); not included is the polyurethane encapsulant used to waterproof the active sub-assembly [64].

active copper-plated piezocomposite layer and passive backing layer are bonded together using a structural epoxy. One of the first demonstrations of 1-3 piezocomposite for wideband reciprocal transduction was the development of a US Navy calibration standard [65]. The measured transmitting voltage response of the new standard at hydrostatic pressures over the range 17.4–6895 kPa (2.5–1000 psi) is shown in Figure 14.

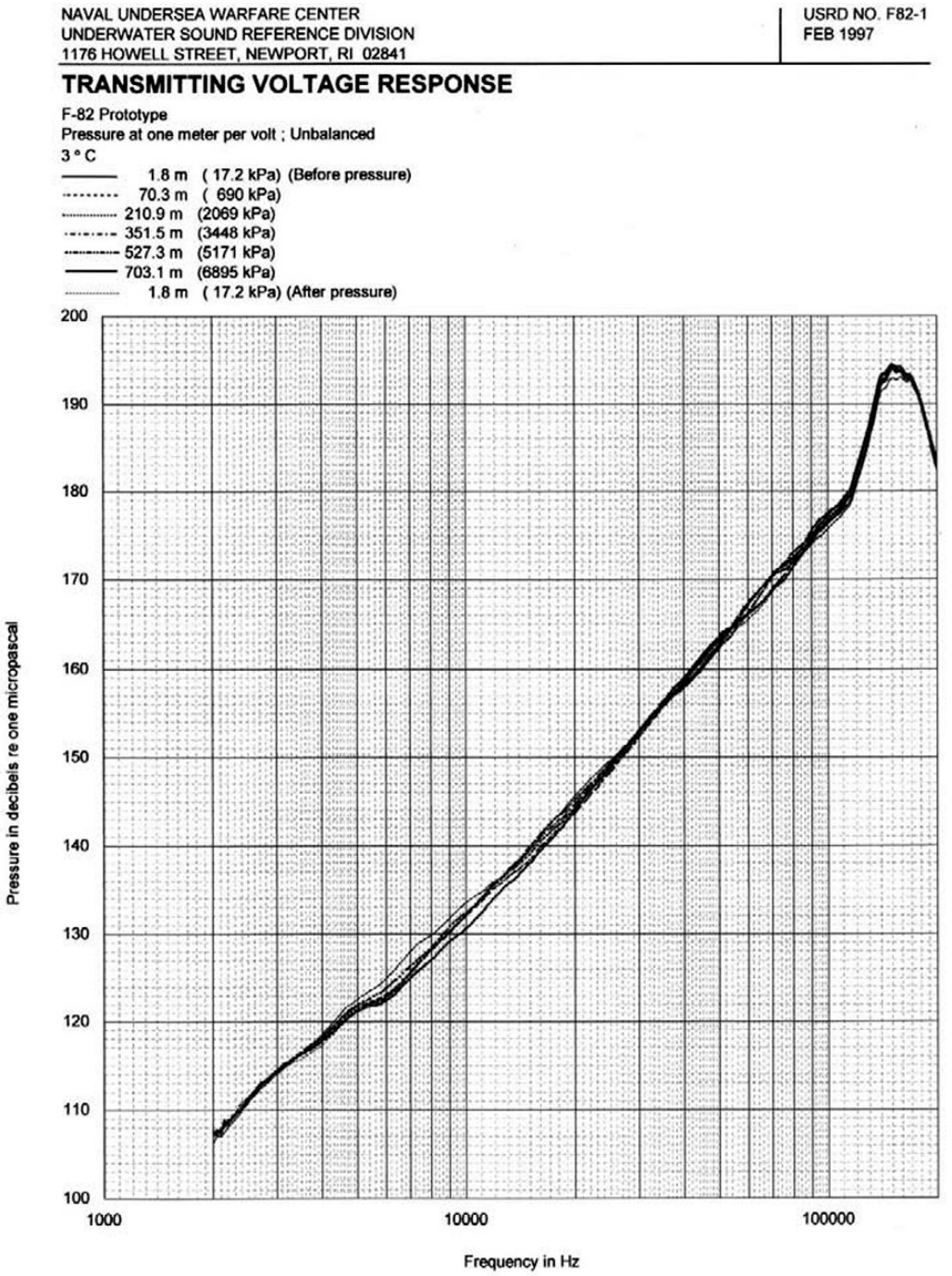


FIGURE 14 Measured transmit voltage response for 1-3 piezocomposite calibration standard.

The response is logarithmic below the region of resonance increasing 12 dB per octave. Over this range the phase angle is relatively flat (around −90°).

3.2 Constant Beamwidth Transducers

The US Navy's Underwater Sound Reference Division (USRD) constructed the first ceramic-based constant beamwidth transducer (CBT) [66], as shown in Figure 15, and continued the development work by demonstrating the applicability of using 1-3 piezocomposite as the active element [67]. The 1-3 piezocomposite was easily electroplated with the required shading pattern and thermo plastically curved to obtain its spherical cap geometry. The basic equations for designing a CBT and a full theoretical treatment of the limits and assumptions related to these equations are presented in reference [68].

Using the approach described in reference [68], two 1-3 piezocomposite-based CBTs were designed and constructed. The first was a 30.5-cm diameter proof-of-concept unit; and the second was 76 cm in diameter weighing over 150 kg. The larger 76-cm unit shown in Figure 16 delivered a constant beamwidth of 14° between 10 and 75 kHz. Both designs had superior



FIGURE 15 Photograph of the first ceramic-based constant beamwidth transducer [66].



FIGURE 16 Photograph of the 76-cm (30-in) piezocomposite-based CBT.

bandwidth characteristics compared to their planar ceramic-based predecessors.

The smaller 30.5-cm proof-of-concept unit, now a USRD calibration standard, has a constant beamwidth between 40 and 230 kHz. Figure 17 shows the normalized beam patterns for this CBT design. The variations in the front to back lobe ratio is related to the frequency-dependent absorption characteristics of the backing material.

3.3 Parametric Arrays

The parametric generation of sound with sound was first put forth by Peter J. Westervelt at Brown University in the early 1960s [69]. When two intense acoustic wave fields having slightly different frequencies are allowed to collimate, say in front of a planar acoustic source, the resulting wave field contains not only the primary frequency components but their sum, difference, and higher harmonics. The degree to which this effect occurs depends on the fluid characteristics and the acoustic pressure amplitude [70–72].

An interesting use of this effect is the formation of a parametric mode array where two high-frequency primary acoustic signals, through the non-linear nature of the fluid, are used to generate sound at their corresponding difference frequency [73,74]. The non-linear interaction region in front of the transducer acts like

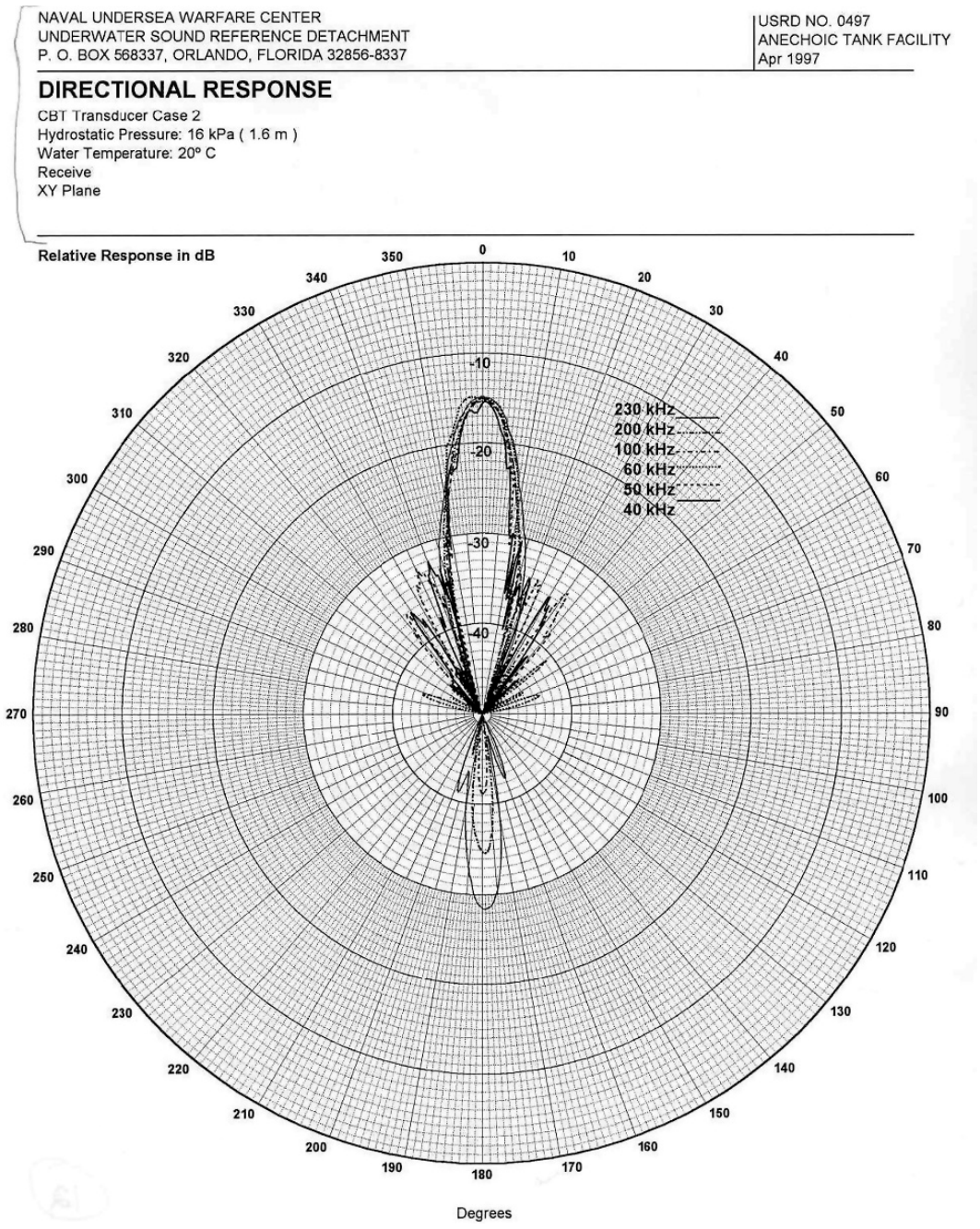


FIGURE 17 Measured beam patterns for the 30.5-cm (12-in) 1–3 piezocomposite-based CBT.

a long end-fire array. The difference frequency radiation retains the directional characteristics of the primaries, which are typically high, and the process produces highly directional beams at low frequency that would be impractical to obtain with linear methods. Difference frequency ranges of over a decade are typical. However, the parametric mode process requires very high sound intensity and is limited by saturation effects [75] associated with the fluid. The sound pressure levels drop as the down shift ratio between the primary and difference frequency increases [76]. The favorable absorption loss at the difference frequency can be offset by the low sound pressure levels attainable. However, in many applications, low frequency is of interest because of other target characteristics such as in sub-bottom profiling.

A parametric mode projector can consist of two arrays of elements each driven with one of the two primary frequencies, or a single array of radiators driven with a composite signal containing both primary frequencies. Care must be taken in the latter case not to subject the transducer to unwanted sub-harmonics generated by the electronics.

The use of 1-3 piezocomposite for parametric mode generation was first demonstrated by the US Navy, at Naval Undersea Warfare Center's Dodge Pond Facility in Niantic, CT. Prior art had exclusively relied on solid piezoceramic blocks similar to the conventional calibration transducer standards shown above.

Figure 18 shows the transmit difference frequency beam patterns for the USRD calibration standard F82 when excited in the parametric mode [65]. The mean primary frequency was approximately 200 kHz corresponding to the 1–3 piezocomposite half-wave thickness resonance. The secondary source levels (L_{ss}) at 5, 10, 30, and 50 kHz are indicated on their respective plots in dB re 1 μ Pa/m.

3.4 Cymbal-based Projectors

One of the most recent innovations in underwater transducers has been the development of the cymbal panel [77]. They consist of an arrangement of cymbal drivers [78] mechanically fastened to stiff, light weight, cover plates. Developed for undersea vehicle applications [79–81], this thin, low frequency, acoustic panel projector technology has shown great promise for not only low frequencies, but broadband operation as well (1–100 kHz). Such low-frequency acoustic source technology is uniquely suited for use on small, unmanned underwater vehicles (UUV) that operate in a shallow water environment. In those cases, cymbal-based technology promises low volume occupation, conformal shapes, substantial low-frequency acoustic output, and low power consumption.

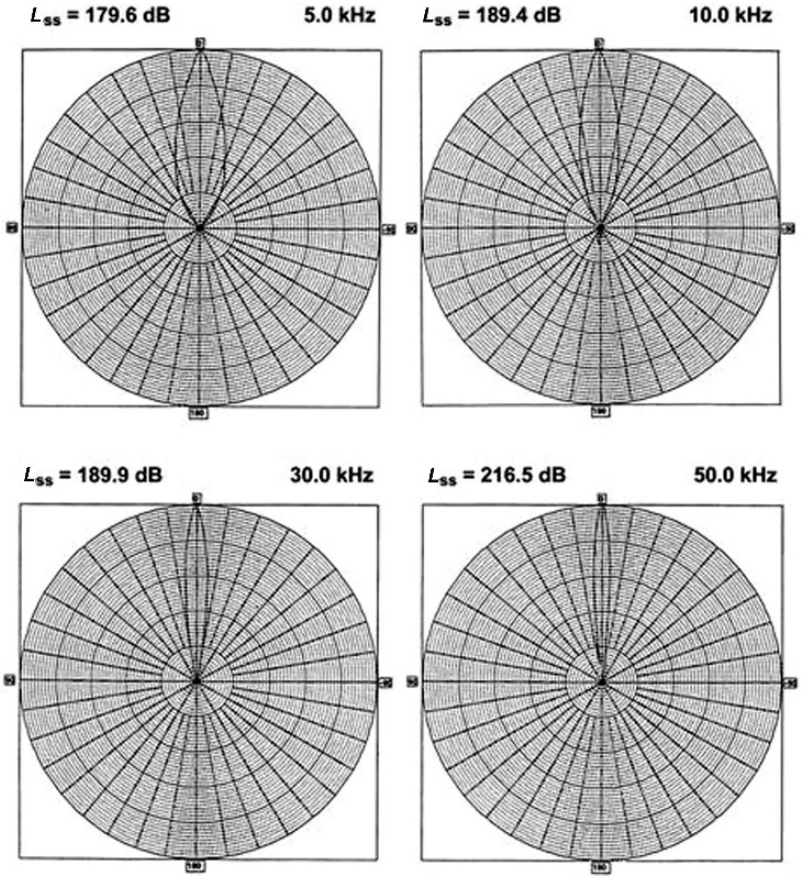


FIGURE 18 Measured difference frequency beam patterns for a 23-cm diameter 1-3 piezocomposite-based parametric mode projector (10 dB/div.) [65].

Figure 19 shows a cymbal panel before encapsulation. Figure 20 compares the low band response of a 10-cm by 10-cm by 1-cm (4″ × 4″ × .5″) cymbal panel with a 25-cm diameter by 6-cm thick (10″ × 2.5″) Class V flextensional transducer. As one can see, the cymbal panel affords comparable low-frequency performance with significantly reduced size and weight. Figure 21 shows the high band response of the prototype cymbal panel.

Lastly in Figure 22 the design curves for the basic cymbal [82] are plotted for material selection and geometry where the range of versatility for this newly emerging technology can be seen.

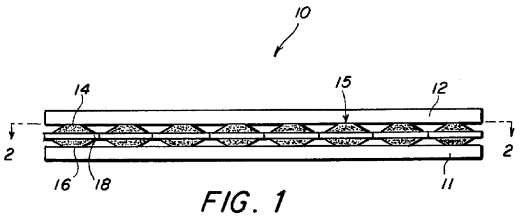


FIGURE 19 Cymbal panel transducer [77].

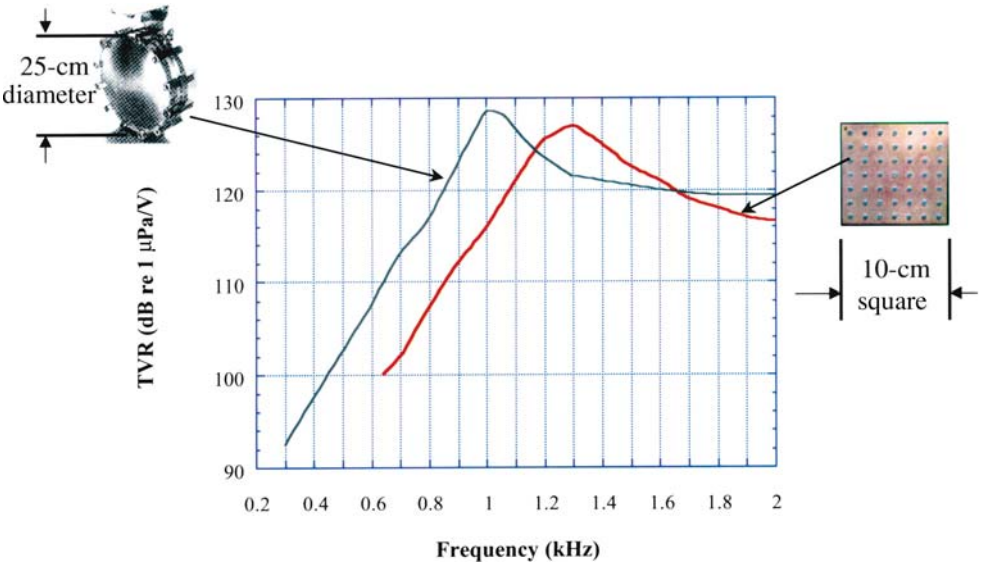


FIGURE 20 Transmit voltage response (TVR) comparison of cymbal panel (*red*); and conventional class V flextenisonal (*black*).

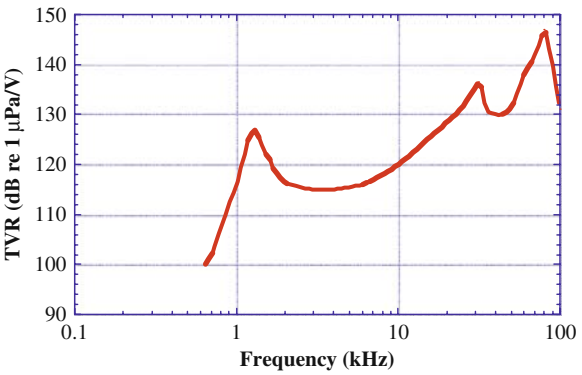


FIGURE 21 Cymbal panel transducer's broadband response.

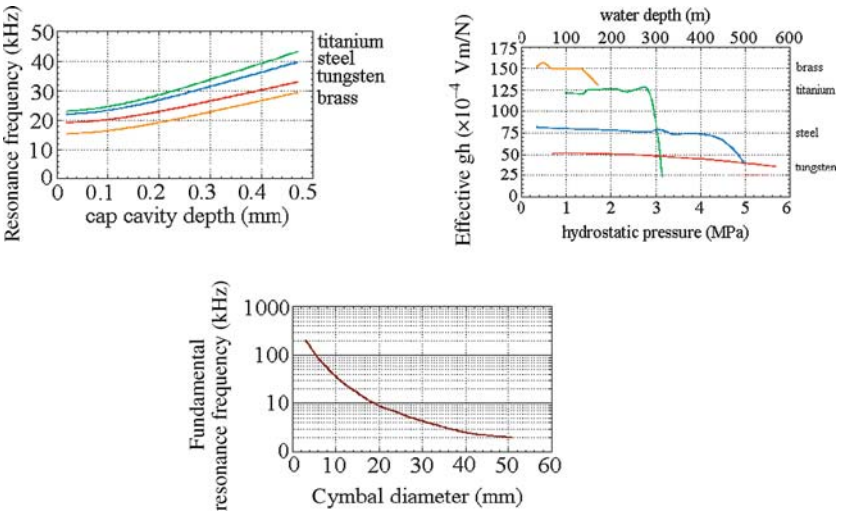


FIGURE 22 Design curves for the basic cymbal driver [82].

4 • Considerations for Remotely Operated Undersea Vehicles

4.1 Environmental Concerns

The highly reverberant and noisy littoral environment precludes the use of high source level, narrow band, near-cw (continuous wave) signals in favor of complex wideband phase coded transients. Thus for most unmanned underwater vehicle (UUV) or autonomous underwater vehicle (AUV) scenarios, wide bandwidth is a more useful attribute than high amplitude. Additionally, the acoustic payloads on vehicles conducting deep water acoustic surveys "fly" close-in to the ocean bottom and therefore do not require high source levels to attain target information. Lastly, it is noted that one of the biggest environmental concerns presently facing the sonar community is the impact of sustained high-intensity sound on marine mammals [83]. For these reasons, and others noted below, piezocomposite-based transducers and arrays are becoming more applicable to undersea vehicle sonar systems.

4.2 Phase and Amplitude Considerations

Much work has been done recently to extend the operational bandwidth of conventional tonpilz transducers. They include multi-layer stacks [84], piezoelectric–magnetostrictive hybrids [85], and the development of single-crystal drive elements [86]. Since all are resonant devices, none of these approaches provide a flat phase response over a wide frequency band (approximately a decade). The reason for this is explained in the figure

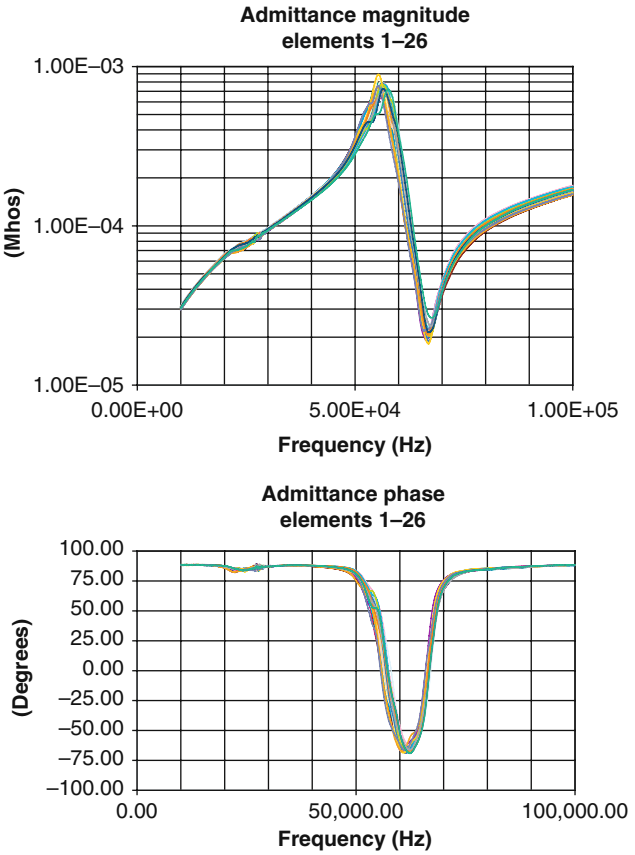


FIGURE 23 Measured impedance magnitude and phase for 26 elements contained in a piezocomposite-based planar array.

below. Figure 23 shows the measured complex admittance for 26 elements contained in a planar piezocomposite array. Note the element to element uniformity in both amplitude and phase is greatest off-resonance. Also, note the respective response levels for all the elements in general are changing rapidly with frequency in the vicinity of the mechanical resonance.

Historically mechanical resonance was used in narrow band systems to achieve high source intensities (viz. high sound pressure levels). This was mainly done to overcome long propagation distances in "blue-water" scenarios. They also made amplifier design and matching circuits simpler with lower applied voltages and narrower operational frequency ranges.

For AUV/UUV sonar systems operating in the littorals, or "flying" at depth, a very wide operational bandwidth can be

realized by placing the transducer's mechanical resonance above the highest frequency of interest and operating the sonar below resonance where the response varies smoothly and the element-to-element variations are relatively small.

4.3 Thermal Runaway

Aside from ultrasonic biomedical transducers, early piezoelectric polymer composites were regarded as "receive only". This was due to thermal runaway effects observed in the substrates that resulted from the ceramic material over heating and softening the polymer backfill. Since then, there have been several advances toward increasing the transmit capability of piezocomposite-based transducers and understanding the limitations associated with high drive [87,88]. These studies experimentally demonstrate that piezocomposite is capable of high drive at resonance if the duty cycle is kept relatively low (for example, $< 2\%$ in parametric mode projectors). For high duty cycle, high drive applications, it has been shown that piezocomposite-driven off-resonance can sustain significant power densities without thermal runaway. The heating effect occurs at resonance when the impedance is low and the drive current is large. Figure 24 further explains this result.

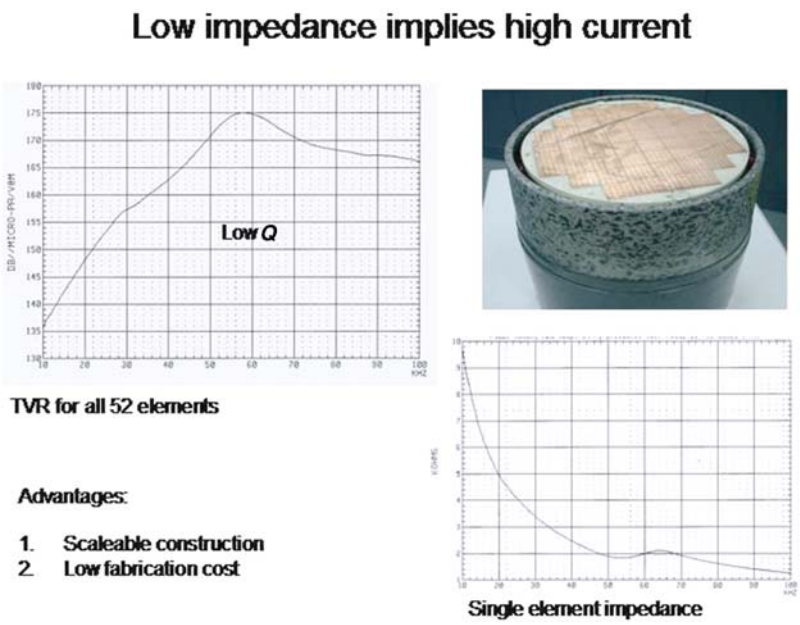


FIGURE 24 Transmit voltage response (TVR) and element impedance for planar piezocomposite array.

For tonpilz elements, the placement of the resonance frequency is critical since the increase in response level is approximately 20 Log (Q_m) and Q_m can be around 10 giving a 20 dB increase in level. Since the mechanical quality factor (Q_m) of the piezocomposite is intrinsically low (~2) moving the resonance to lower frequencies on the transmit response curve (possibly by increasing the layer thickness and/or mass loading) does not provide a significant increase in output. Instead it exacerbates the heating problem along with the other unwanted resonance characteristics mentioned above (i.e. amplitude and phase uniformity).

4.4 Phase Response and Radiating Aperture

One last interesting feature of piezocomposite constructs worth noting is element phase response is *not* a function of area. The phase is governed by the rod dimensions. Shown in Figure 25 are phase plots of elements having differing area. This was done by driving 1, 2, 3, 4, or 5 elements at a time. The admittance magnitude scaled normally; increasing with area, but the phase value versus frequency (shown in the figure) remains constant. This phase consistency would also be observed for identically sized tonpilz elements combined electrically in parallel. However, in the piezocomposite case the array elements neither have to be of equal size or shape. This is a distinct advantage over tonpilz-based arrays where radiating head masses must be both symmetric and identical in order for the elements to resonate in phase.

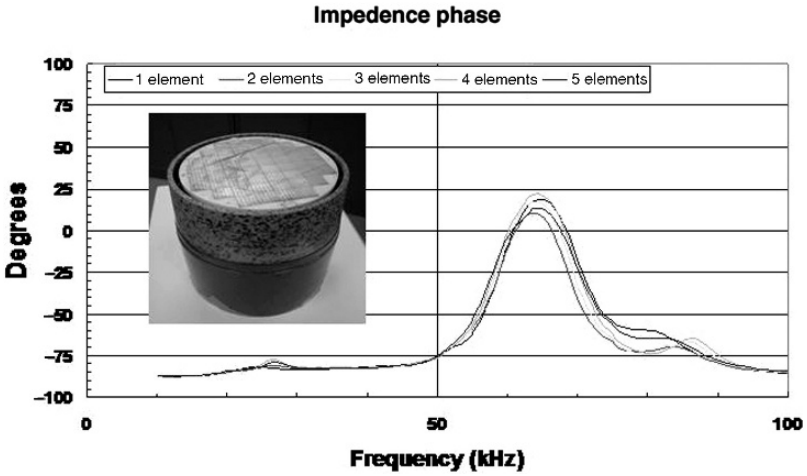


FIGURE 25 Phase response versus area for planar piezocomposite array.

4.5 Fabrication Techniques

This next section describes some of the pertinent fabrication aspects of piezocomposite. The various construction techniques such as thermoforming, cabling, application of electrodes, and machining piezocomposite will be briefly described.

Cabling conventional arrays usually requires attaching two wires to each of the elements. These wires have to be routed outside each individual element. With piezocomposite, arrays of elements are formed over a surface via electrode deposition. At this point wires may be attached to the electrode surfaces in the conventional fashion.

The US Navy has developed fabrication techniques [89,90] that eliminate the need for external cabling by routing the element wiring through the composite interior and subsequently making connection to individual array element electrodes at the outer edges of the panel. This allows large conformal two-dimensional arrays to be fabricated by thermoforming the electroded panels to match doubly curved hull shapes typical of undersea vehicles.

Figure 26 shows two conformal arrays before and after curving. The wire ends for each element exit from the positive surface of the unelectroded piezocomposite panel as seen in the upper left corner of Figure 26. The wire cross section can vary as long as its overall diameter allows it to fit between the rods of the PZT

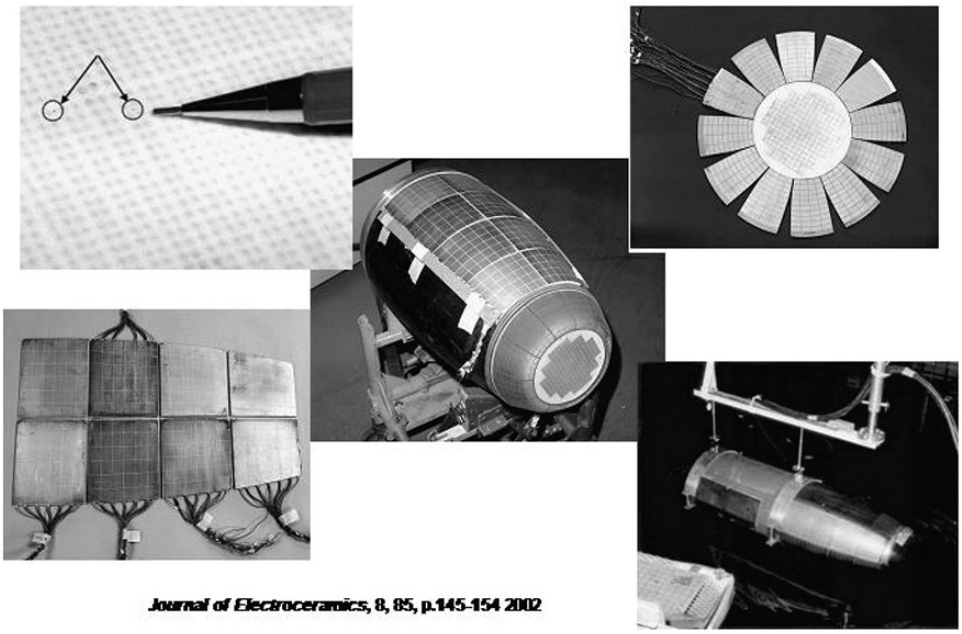


FIGURE 26 Examples of 1-3 piezocomposite conformal arrays.

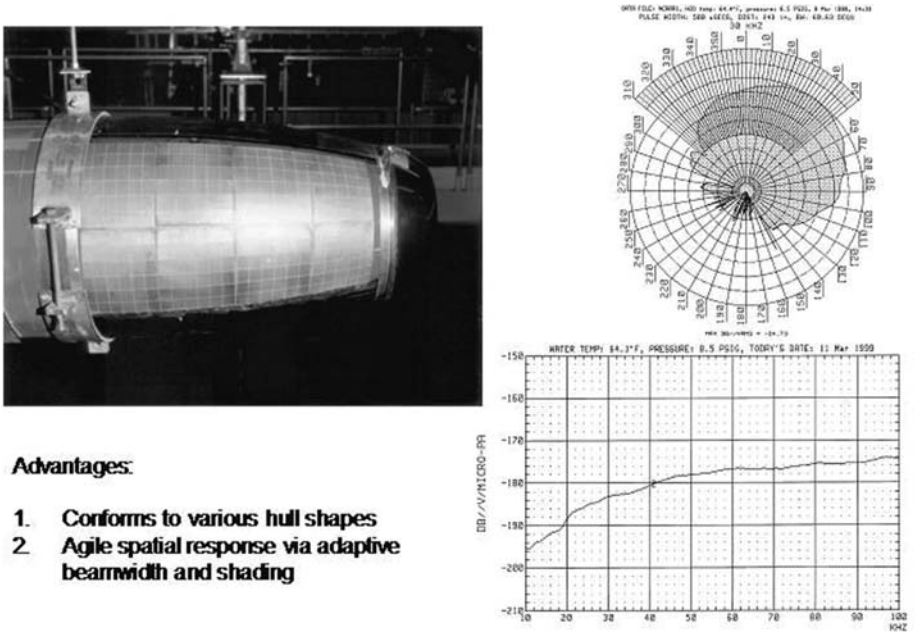


FIGURE 27 Measured results for demonstration array.

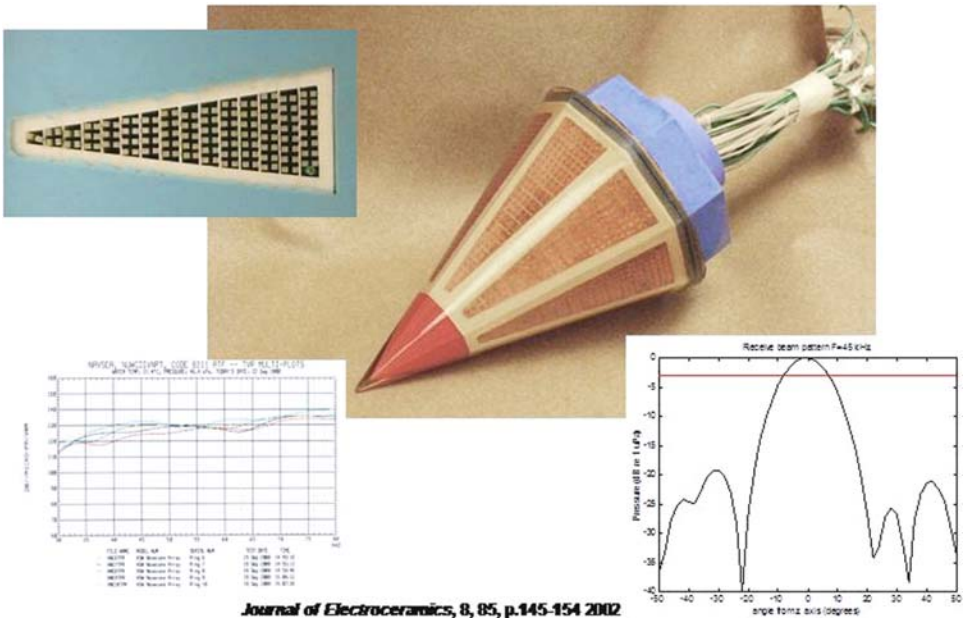


FIGURE 28 Prototype 120-element conical octahedral array.

perform. Copper plating makes electrical contact with these wire ends and defines the individual element apertures. The plating is deposited and photo etched to the desired pattern before curving. The curving of the piezocomposite panels is facilitated by the use of a thermoplastic polymer backfill matrix which softens at about 60°C. For a more detailed description on the fabrication details of this demonstration array, the reader is referred to reference [89]. The measured receive response and roll-plane beam pattern



FIGURE 29 Multi-element cylindrical array module, fabrication tooling, modular array concept, and finished prototype module.

shown in Figure 27 are typical for a 1-3 piezocomposite-based transducer.

Arrays can also be formed by machining and joining panels of piezocomposite. Figure 28 shows a conical array that was formed by adhesively joining eight sub-array sections to a nose (red) and base (blue) section. The nominal dimensions for the array were 18.4-cm long by 12.7-cm diameter base with a 20° half-angle. Measured results for the array include the transmit responses for five of the fifteen rings of elements, and a forward steered beam pattern. Further details are available in reference [89].

The array segments can also be machined to a singly or doubly curved surface using standard grinding techniques [91]. Figure 29 shows a modular array sub-section consisting of radial element sections. In this case, each element consists of two independently controlled piezocomposite layers. The modular array approach is scaleable in both length and diameter and can be configured along with the conical array to provide three-dimensional spatial coverage in both the transmit and the receive mode.

In addition to easing the environmental impact of high-amplitude narrowband sonar on marine mammals, piezocomposite-based transducers operated below resonance are an excellent choice for generating and receiving wideband (approximately one decade) complex acoustic signals. Unlike conventional resonant devices the piezocomposite-based transducer technology provides many interesting advantages for designers of compact, undersea vehicle sonar payloads. These include (1) lightweight, cost-effective, rugged, conformal packaging, (2) phase and amplitude response uniformity, (3) element-to-element phase and amplitude uniformity, and (4) high drive transmit capability.

References

1. L. H. Van Vlack "*Elements of Material Science* 3rd Ed.," Addison-Wesley, Massachusetts, pp. 277, (1975).
2. B. Jaffe, W. R. Cook and H. Jaffe, "*Piezoelectric Ceramics*," Academic Press, New York, (1971).
3. DOD-STD-1376A, "Piezoelectric Ceramic Materials and Measurement Guidelines for Sonar Transducers," (1985).
4. "IEEE Standard on Piezoelectricity," (IEEE Standard 176), (1978).
5. "Modern Piezoelectric Ceramics," Vernitron Piezoelectric Division technical data brochure **PD 9247**, (1988).
6. W. P. Mason, "An electromechanical representation of a piezoelectric crystal used as a transducer," *Proc. I.R.E.*, pp. 1252, (1935).
7. W. P. Mason, "*Electromechanical Transducers and Wave Filters*," D. Van Nostrand, New York (1948).

8. L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders, "*Fundamentals of Acoustics*," Third Edition, John Wiley and Sons, New York, (1982).
9. S. Hanish, "*A Treatise on Acoustic Radiation, Volume II – Acoustic Transducers*," Naval Research Laboratory, Washington, D. C. (1983).
10. H. Kawai, "The piezoelectricity of poly (vinylidene fluoride)," *JPA. J. of Appl. Phys*., **8**, pp. 975, (1969).
11. P. E. Bloomfield, R. A. Ferren, P. F. Radice, H. Stefanou and O. S Sprout, "Piezo- and pyroelectricity in poly (vinylidene fluoride)," *Nav. Res. Rev*., **31**, pp. 1–15, (1978).
12. P. Pantelis, "Properties and applications of piezoelectric polymers," *Physical Technology*, **15**, pp. 239, (1984).
13. Hari Singh Nalwa, "*Ferroelectric polymers: Chemistry, Physics, Applications*," Marcel Dekker, New York, (1995).
14. R. H. Tancrell, D. T. Wilson and D. Ricketts, "Properties of PVDF polymer for sonar," *Proceedings of IEEE Ultras. Symp*., San Francisco, CA, (1985).
15. D. Ricketts, "Recent developments in the U.S.A. in the application of PVF_2 polymer in underwater transducers," *Proc. Inst. Acoustics*, **6** (3), pp. 46, (1984).
16. R. H. Tancrell, D. T. Wilson, "Sheets of stretched and polarized polymer materials and method of manufacturer," U. S. Patent 5,024,872, (1991).
17. R. H. Tancrell, D. T. Wilson, N. T. Dionesotes, and L. C. Kupferberg, "PVDF piezoelectric polymer: processing, properties and hydrophone applications," *Proceedings of Transducers for Sonics and Ultrasonics*, Edited by M. D. Mcollum, B. F. Hamonic, and O. B. Wilson, Technomic Publishing, Pennsylvania, pp. 103, (1992).
18. R. G. Swartz and J. D. Plummer, "On the generation of high frequency acoustic energy with polyvinylidene fluoride," *IEEE Trans. Son. & Ultras*., Vol. SU-27, No. 6, pp. 41, (1980).
19. H. Ohigashi, "Ultrasonic transducers in the megahertz range," *The Applications of Ferro-electric Polymers*, Edited by T. T. Wang, J. M. Herbert and A. M. Glass Routledge, Chapman, and Hall, pp. 237, (1988).
20. J. M. Powers, "Piezoelectric polymer – an emerging hydrophone technology," EASCON 1979 Record (IEEE Publication 79CH1476-1AES), Vol. 3, pp. 517, (1979).
21. D. Ricketts, Model for a piezoelectric polymer flexural plate hydrophone," *J. Acoust. Soc. Am*., **70**, pp. 929, (1981).
22. Q. M. Zhang, V. Bharti, and X.Zhao, "Giant electrostriction and relaxor ferroelectric behavior in electron-irradiated poly(vinylidene fluoride-trifluoroethylene) Copolymer.
23. R. E. Newnham, D. P. Skinner and L. E. Cross, "Connectivity and piezoelectric-pyroelectric composites," *Mat'ls. Res. Bull*., **13**, pp. 525, (1978).
24. W. A. Smith, "The role of piezocomposites in ultrasonic transducers," *Proc. IEEE Ultrason. Symp*., pp. 755, (1989).

25. T. R. Gururaja, W. A. Schulze, L. E. Cross, R. E. Newnham, B. A. Auld, and Y. J. Wang, "Piezoelectric composite materials for transducer applications. Part I: Resonant modes of vibration for PZT rod-polymer composites," *IEEE Trans. Ultrason.*, vol. SU-32, pp. 481, (1985).
26. W. A. Smith, A. Shaulov, and B. A. Auld, "Tailoring the properties of composite piezoelectric materials for medical ultrasound transducers," *Proc. IEEE Ultrason. Symp.*, pp. 642, (1985).
27. L. Bowen and K. French, "Fabrication of piezoelectric/polymer composites by injection molding," *Proceedings of the 8th IEEE International Symposium on Applications of Ferroelectrics*, pp. 160, (1992).
28. J. Kuwata, K. Uchino, and S. Nomura, "Dielectric and piezoelectric properties of PbTiO3 single crystals," *Jpn. J. Appl. Phys.*, **21** (9), pp. 1298, (1982).
29. S. Park and T. R. Shrout, "Ultrahigh strain and piezoelectric behavior in relaxor-based ferroelectric single crystals," *J. Appl. Phys.*, **82**, pp. 1804, (1997).
30. S. Park and T. R. Shrout, "Characteristics of relaxor-based piezoelectric single crystals for ultrasonic transducers," *IEEE Trans. Ultrason., Ferr.*, **44**, pp. 1140, (1997).
31. W. P. Mason, *Electromechanical Transducers and Wave Filters*, 2nd Ed. D. Van Nostrand: New York, (1948).
32. D. Stansfield, *Underwater Electroacoustic Transducers, a handbook for users and designers* (Bath U. P., Bath, I.O.A., 1990).
33. W. Hahnemann and H. Hecht, "Die grundform des mechanischakustischen schwingungskörpers", ("The form of the mechanical oscillator"), *Phys. Z.*, **21**, 187, (1920).
34. D. F. McCammon and W. Thompson Jr., "The design of tonpilz piezoelectric transducers using nonlinear goal programming," *J. Acoust. Soc. Am.*, **68**, pp. 754, (1980).
35. G. E. Martin, "On the theory of segmented electromechanical systems," *J. Acoust. Soc. Am.*, **36**, pp. 1366, (1964).
36. S. C. Butler, "A 2.5 kHz Magnetostrictive tonpilz sonar transducer design," Proceedings SPIE 9th *Symposium on smart structures and materials*, Section 4699–65, Session 11, (2002).
37. J. L. Butler and S. J. Ciosek, "Rare earth iron octagonal transducer," *J. Acoust. Soc. Am.*, **67**, pp. 313, (1982).
38. F. Massa, "Electroacoustic transducer," U. S. Patent No. 3,328,751, (1967).
39. J. L. Butler, J. R. Cippola, and W. D. Brown, "Radiating head flexure and its effect on transducer performance," *J. Acoust. Soc. Am.*, **70**, pp. 500, (1981).
40. S. C. Thompson, M. P. Johnson, E. A. McLaughlin, and J. F. Lindberg, "Performance and recent developments with doubly resonant wide band transducers," in *Transducers for Sonics and Ultrasonics*, Ed. M. D. McCollum *et al.* (TECHNOMIC, U.S.A., Lancaster, Pennsylvania 1993), pp. 17.

41. D. Rajapan, "Performance of a low-frequency, multi-resonant broadband Tonpilz transducer," *J. Acoust. Soc. Am.*, **111**, pp. 1692, (2002).
42. S. C Butler and F. A. Tito, "A broadband hybrid magnetostrictive/piezoelectric transducer array," Proceedings of the MTS/IEEE Oceans 2000 Conference, Providence, RI, (Sept. 2000).
43. A. E. Clark, "Magnetostrictive Rare Earth $-Fe_2$ compounds." Chapter **??** in *Ferromagnetic Materials*. Vol. 1. E. P. Wohlfarth, Ed. North Holland Pub. Co. (1980).
44. G. A. Vincent, "Transducer Housing with flexible supports," U. S. Patent No. 3,341,721, (1964).
45. J. D. Bullock, "Square and rectangular electroacoustic bender bar transducer," U. S. Patent No. 4,140,936, (Feb. 1979).
46. R. S. Woollett, *The Flexural Bar Transducer*, Naval Underwater Systems Center, (1985).
47. W. J. Toulis, "Flexural-extensional electromechanical transducer," U. S. Patent No. 3,277,433, (1966).
48. K. D. Rolt "History of the flextensional electroacoustic transducer," *J. Acoust. Soc. Am.*, **87**, pp. 1340, (1990).
49. L. H. Royster, "Flextensional Underwater Acoustics Transducer," *J. Acoust. Soc. Am.*, **45**, pp. 671, (1969).
50. G. A. Brigham, "Analysis of the class IV flextensional transducer by use of Wave Mechanics," *J. Acoust. Soc. Am.*, **56**, pp. 31, (1974).
51. J. Oswin and J. Dunn, "Frequency, power, and depth performance of Class IV flextensional transducers," *Power Sonic and Ultrasonic Transducers Design*, Edited by B. Hamonic and J. N. Decarpigny, pp. 121, Springer, New York, (1988).
52. W. T. Harris, "Split Cylindrical Transducer," U. S. Patent No. 2,812,452, (1957).
53. G. W. McMahon and D. F. Jones, "Barrel Stave Projector," U. S. Patent No. 4,922,470, (1990).
54. C. J. A. Purcell, "Folded Shell Projector (FSP)," U. S. Patent No. 5,805,529, (1998).
55. D. A. Berlincourt, D. R. Curran, and Hans Jaffe, "Piezoelectric and piezomagnetic materials and their function in transducers," Chapter 3 in *Physical Acoustics* Vol. 1A. W. P. Mason, Ed. Academic Press, New York, (1964).
56. G. E. Martin, "Vibrations of longitudinally polarized Ferroelectric cylindrical tubes," *J. Acoust. Soc. Am.*, **35**, pp. 510, (1963).
57. R. Holland and E. P. EerNisse, *Design of resonant piezoelectric devices*, MIT Press, Cambridge, Massachusetts (1969).
58. O. C. Zienkiewicz, R. L. Taylor, *Finite Element Method: Volumes I–III*, 5th ed., Butterworth-Heinemann (2000).
59. O. B. Wilson, *An introduction to the theory and design of sonar transducers*, Peninsula Publishing, Los Altos, CA, (1988).
60. M. D. Ring, "Two methods for modeling the longitudinal resonance of Class IV flextensional transducers," in *Transducers for Sonics and Ultrasonics*, Ed. M. D. McCollum *et al*. (TECHNOMIC, U.S.A., Lancaster, Pennsylvania 1993), pp. 198.

61. B. F. Hamonic, J. C. Debus, J. N. Decarpigny, D. Boucher, and B. Tocquet, "Analysis of a radiating thin-shell sonar transducer using the finite element method," *J. Acoust. Soc. Am.*, **84**, pp. 1245, (1989).
62. M. J. Huan and R. E. Newnham, "An experimental and theoretical study of 1–3 piezoelectric PZT-polymer composites for hydrophone applications," *Ferroelectrics*, **68**, pp. 123, (1986).
63. A. C. Hladky-Hennion, "Modeling of active and passive periodic structures," in *Transducers for Sonics and Ultrasonics*, Ed. M. D. McCollum *et al*. (TECHNOMIC, U.S.A., Lancaster, Pennsylvania 1993), pp. 6.
64. K. C. Benjamin, "Recent advances in 1–3 piezoelectric polymer composite transducer technology for AUV/UUV acoustic imaging applications," Journal of Electroceramics, 8, pp. 145, (2002).
65. K. C. Benjamin and S. Petrie, "The design, fabrication, and measured acoustic performance of a 1–3 piezocomposite-based Navy calibration standard transducer," *J. Acoust. Soc. Am.*, **109**, No. 5, Pt. 1, pp. 1973, (2001).
66. A. L. Van Buren, L. Dwight Luker, M. D. Jevnager, and A. C. Tims, "Experimental constant beamwidth transducer," *J. Acoust. Soc. Am.*, **73**, pp. 2200, (1983).
67. K. C. Benjamin, A. K. Walden and A. L. Van Buren, "Design and development of a constant beamwidth transducer for sub-bottom acoustic profiling," MTS/IEEE Oceans '97 Conference, Halifax, Nova Scotia, Canada, (1997).
68. P. H. Rogers and A. L. Van Buren, "New approach to a constant beamwidth transducer," *J. Acoust. Soc. Am.*, **64**, pp. 38, (1978).
69. P. J. Westervelt, "Parametric acoustic array," *J. Acoust. Soc. Am.*, **35**, pp. 535, (1963).
70. D. T. Blackstock, "Thermoviscous attenuation of plane, periodic, finite amplitude soundwaves," *J. Acoust. Soc. Am.*, **36**, pp. 534, (1964).
71. R. T. Beyer, *Non-Linear Acoustics*, Naval Ship Systems Command, (1974).
72. T. G. Muir and J. G. Willette, "Parametric acoustic transmitting arrays," *J. Acoust. Soc. Am.*, **52**, pp. 1481, (1972).
73. H. Hoback, "Experimental investigation of an acoustic end fired array," *J. Sound & Vib.*, **6**, pp. 460, (1967).
74. M. B. Moffett and R. H. Mellen, "Nearfield characteristics of parametric acoustic sources," *J. Acoust. Soc. Am.*, **69**, pp. 404, (1981).
75. M. B. Moffett and R. H. Mellen, "Model for parametric acoustic sources," *J. Acoust. Soc. Am.*, **61**, pp. 325, (1977).
76. R. L. Rolleigh, "Difference frequency pressure within the interaction region of a parametric array," *J. Acoust. Soc. Am.*, **58**, pp. 964, (1975).
77. T. R. Howarth, "Acoustic Transducer Panel," U.S. Patent No. 6,438,242, (2002).

78. R. E. Newnham and J. Zhang, "Flextensional metal-ceramic composite transducer," U. S. Patent 6,232,702, (2001).
79. T. R. Howarth and J. F. Tressler, "Development of thin, low frequency electroacoustic projectors for underwater applications," *Proceedings of the 17th International Congress on Acoustics*, AIP Publications, Volume II, 4–5, (2001).
80. J. F. Tressler and T. R. Howarth, "A comparison of the underwater acoustic performance of 1–3 piezocomposite materials to cymbal-based projectors," *J. Electroceram.*, **8**, pp. 175, (2002).
81. J. F. Tressler and T. R. Howarth, "Thin, low frequency, high displacement actuator panels," *J. Mater. Res. Innovat.*, **2**, pp. 270, (1999).
82. J. F. Tressler, Ph.D. Theis, The Pennsylvania State University, University Park, PA (1997).
83. National Research Council, Committee on characterizing biologically significant marine mammal behavior, *Marine Mammal Populations and Ocean Noise*, The National Academies Press, Washington, DC (2005).
84. D. Rajapan, "Performance of a low-frequency, multi-resonant broadband Tonpilz transducer," *J. Acoust. Soc. Am.*, **111**, pp. 1692, (2002).
85. S. C Butler and F. A. Tito, "A broadband hybrid magnetostrictive/piezoelectric transducer array," Proceedings of the MTS/IEEE Oceans 2000 Conference, Providence, RI, (Sept. 2000).
86. J.M. Powers, M.B. Moffett and F. Nussbaum, "Single Crystal Naval Transducer Development," in Proceedings of IEEE International Symposium on the Applications of Ferroelectrics (ISAF2000), pp. 351, (IEEE, New York, NY, 2000).
87. T. R. Howarth, "Investigations of 1–3 piezocomposite drive limits," Proceedings of the 25th Annual International Conference on Advanced Ceramics and Composites, pp. 567, (2001).
88. K. C. Benjamin, "Endurance experiments using pulsed high power drive signals on thick 25.4 mm, (1-inch) injection molded 1–3 Piezoelectric composite in single and multi-layered transducer configurations," Proceedings of the 25th Annual International Conference on Advanced Ceramics and Composites, pp. 555, (2001).
89. K. C. Benjamin, "Recent advances in 1–3 piezoelectric polymer composite transducer technology for AUV/UUV acoustic imaging applications," *J. Electroceram.*, **8**, pp. 145–154, (2002).
90. K. C. Benjamin, "Shaped piezoelectric composite transducer and method of making," U.S. Patent No. 6,255,761, (2001).
91. K. C. Benjamin, "Shaped piezoelectric composite array," U.S. Patent No. 6,711,096, (2004).

99

Flow Noise

Gerald C. Lauchle

State College, PA, USA

1 • Introduction

Undesirable noise in sonar transducers can be due to ambient, electronic, machinery, propulsion, and hydrodynamic sources. This section is concerned with the hydrodynamic sources that cause a spurious output from a sonar transducer known as *flow-induced self noise*. The term *hydrodynamic noise* is used when the dynamics of the turbulence creating the noise is very rarely influenced by the compressibility of the medium. In water, the flows are of extremely low Mach number and thus hydrodynamic. A review of the fundamentals of hydrodynamic noise may be found in Ffowcs Williams (1969).

The degree to which turbulent flow causes transducer self noise depends on the location of the transducer, whether it is within a moving vehicle, being towed, drifting, or fixed, the manner in which it is mounted, and on its directivity characteristics. The time-dependent flows of concern include the turbulent boundary-layer and transitional flows that exist on the exterior hull of moving vehicles, the highly separated flows on drifting and suspended hydrophones exposed to ocean currents, and the thick axisymmetric boundary-layer flows that exist on towed line arrays. Hydrodynamic sources that may be equally important to sensor self noise, but are not considered in this section, are those that occur within auxiliary on-board fluid flow systems and those created by the propellers or propulsors of self-propelled vehicles. Modern sensor systems may not be composed of only pressure hydrophones; some applications use acoustic vector sensors.

The flow-induced noise on vector sensors is an area of considerable contemporary research, and the newest findings have yet to appear in journals (McConnell, 2004); however, the reader may refer to McEachern and Lauchle (1995) and Lauchle et al. (1995) for foundational research in this area.

2 • Flow-Induced Noise in Hull-Mounted Sonar Systems

The vast majority of sonars in self-propelled moving vehicles are located in the bow region because of their intended purpose of locating upstream distance sources and targets. Another reason is that the bow region is farthest away from noisy machinery and propulsion sources. The sonar array directivity is also an advantage because it can discriminate sources in the forward direction from those aft.

On surface ships and submarines, the transducer arrays are mounted within an acoustically transparent dome that serves to protect them from impacts, while at the same time displaces the hydrodynamic sources of self noise away from the transducer face. The radiated noise from the displaced hydrodynamic sources still reaches the transducers however. On smaller vehicles, like torpedoes and unmanned undersea vehicles (UUVs), domes are impractical. Such vehicles may have the transducers mounted flush to the hull exposing them to external flow and structural fluctuations. This occurs with planar as well as conformal arrays. The plane of a planar array is perpendicular to the vehicle flow vector and is located behind a thin rubber window in the most forward stagnation region of the hull. The flow there is intended to be laminar and free of fluctuation, but transition of laminar-to-turbulent boundary layer flow may occur at the edge of a planar array. This phenomenon is also a hydrodynamic source of self noise, Lauchle (1991). For arrays that conform to the hull shape, the elements may be mounted slightly below a compliant rubber blanket, but they are still exposed almost directly to turbulent boundary layer wall pressure fluctuations.

Thus, three possible mechanisms of hull-mounted sonar flow-induced self noise can exist: (1) transition of laminar-to-turbulent boundary layer flow, (2) fully developed turbulent boundary layer flow, and (3) acoustic radiation from sources (1) and (2) into the elements of a domed sonar. A fourth mechanism is the structural waves created by hydrodynamic sources. If the array is poorly isolated from structure-borne noise sources, the flow-induced structural vibration will contribute to sonar self noise. Because this topic is so application specific, it will not be rigorously discussed here.

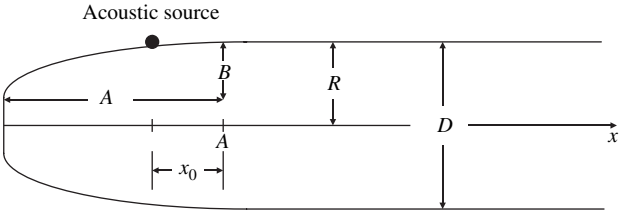


FIGURE 1 Typical geometry of the forward stagnation region of an axisymmetric body with elliptic curvature, where A and B are the major and minor axes of the fairing ellipse.

3 • Transition Zone Noise Scaling

The radiation and diffraction of transition zone noise into a planar array was studied by Arakeri, et al. (1991) and Lauchle (1992). The nose geometry is sketched in Figure 1.

A sonar array is assumed present in the forward flat of radius $R - B$. The curved surface is described by an ellipse of major and minor axes A and B, respectively, and the source is a ring source created by laminar-to-turbulent boundary layer transition. It was shown that the noise received at the center of the planar flat of this nose geometry scales as the 7.5 power of flow speed and the square of the diameter D. The diffraction loss around the nose contour is expressed analytically in Lauchle (1992) and can range from 10 to 60 dB depending on the acoustic wavelength relative to the body diameter. The mechanism described in this work is direct radiation from the transition zone diffracting around the nose contour into a sonar array. It assumes the body is rigid and no structural excitation is occurring. If structure-borne energy were created by the transition of turbulent flow, additional terms would be required in the scaling relationship to account for this. Furthermore, if the array were directional, its directivity index would need to be added to the scaled noise level.

4 • Turbulent Boundary Layer Wall Pressure Fluctuations

We consider now the noise created by turbulent flow over pressure transducers. For modeling purposes, the flow is two dimensional over a rigid planar surface as indicated in Figure 2. The flow velocity right at the wall is zero because of viscosity. The flow far away from the surface is laminar and has velocity, U_0. The thin layer of fluid between the surface and the free stream is a shear layer, or boundary layer. The local velocity $U(x_2)$ within

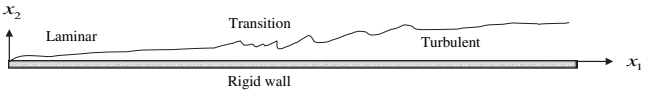


FIGURE 2 Schematic description of boundary layer flow.

this layer goes from zero to U_0 very quickly. Initially, near the leading edge (or forward stagnation point if this were a body) the boundary layer is laminar and steady; it creates no noise. After only a short distance, because of local disturbances in the free-stream flow or surface roughness, the boundary layer transitions to a turbulent one. Once the shear layer is completely filled with turbulent fluctuations, it is called a turbulent boundary layer (TBL).

The TBL is chaotic and creates local pressure fluctuations. Kraichnan (1956) was the first to analyze these pressure fluctuations and he showed that they are solutions of a Poisson equation

$$\frac{\partial^2 P}{\partial x_i^2} = -\rho \frac{\partial^2 (U_i U_j)}{\partial x_i \partial x_j}, \tag{1}$$

where P is total pressure and U_i are the total velocity components in the three principal directions ($i = 1, 2, 3$). Equation (1) shows that the total pressure is caused exclusively by the inertial stresses in the flow; viscosity plays no role.

A Reynolds averaging is now performed, where $U_i = \bar{U}_i + u_i$ and $P = P_0 + p$. Here the upper case symbols represent mean values and the lower case represent unsteady values. The over bar represents a mean. The mean of fluctuating variables is understood to be zero. Kraichnan took the time average of Equation (1) and simplified using the properties of a TBL:

$$\begin{aligned} U_i &= [\bar{U}_1(x_2), 0, 0], \\ u_i &= (u_1, u_2, u_3). \end{aligned} \tag{2}$$

The result is the following equation for the unsteady pressure anywhere within a TBL:

$$\nabla^2 p = -\rho \left[2 \frac{\partial \bar{U}_1}{\partial x_2} \frac{\partial u_2}{\partial x_1} - \frac{\partial^2}{\partial x_i \partial x_j} (u_i u_j - \overline{u_i u_j}) \right], \tag{3}$$

where $\overline{u_i u_j}$ is the correlation function among fluctuating velocity components. The first term of the right-hand side is a mean shear–turbulence interaction term, while the second

is a turbulence–turbulence interaction term. For simplicity, let $\nabla^2 p(\vec{x},t) = -q(\vec{x},t)$. Then the solution is

$$p(\vec{x},t)\big|_{x_2=0} = \frac{1}{2\pi}\int\limits_{x_2>0} \frac{q(\vec{y},t)}{|\vec{x}-\vec{y}|} d^3\vec{y}. \tag{4}$$

The pressure fluctuations at $\vec{x}$ are due to turbulent velocity fluctuations at all $\vec{y}$, but the large $\vec{y}$ contributions are *attenuated* because of the factor $|\vec{x}-\vec{y}|$. This means that the pressure fluctuations are very localized, and intense.

Spectral solutions are derived for $x_2 =$ a constant; (x_1, x_3) form a plane. Assume that the flow is stationary and homogeneous in these planes. The space–time Fourier transform is

$$P(k_1, k_3, x_2, \omega) = \frac{1}{(2\pi)^3}\int\int\limits_{\infty}\int p(\vec{x},t) e^{-i\vec{k}\vec{x}-i\omega t} d\vec{x}dt \tag{5}$$

and the wave vector–frequency power spectrum is

$$\Phi_{pp}(\vec{k},\omega) \equiv \frac{\overline{P^*(\vec{k},\omega)P(\vec{k}',\omega')}}{\delta(\vec{k}'-\vec{k})\delta(\omega'-\omega)}. \tag{6}$$

We note that the wall pressure on a rigid, flat surface $P(\vec{k},\omega)$ is called the *blocked pressure*. It is twice the magnitude of the pressure in the absence of the wall. Modeling of Equation (6) has been performed by many investigators. Of particular note are those of Corcos (1964), Ffowcs Williams (1965) and Chase (1980, 1987). The general features of the wave vector–frequency spectrum of the fluctuating *wall* pressures are shown in Figure 3. The peak is called the convective peak. It indicates that the turbulent flow pattern almost retains its spatial structure as it moves past the flat surface. The eddies are convecting at speed u_c such that $k_c = \omega/u_c$ is the convection wave number. The acoustic wave number is $k_0 = \omega/c_0$. All wave numbers less than or equal to k_0 propagate as true sound. This region shows up as a circle on Figure 3 because the sound propagates in all directions. Other important wave numbers on this plot are related to the TBL thickness δ, and the possible resonant flexural wave number of a flat plate if it were not rigid. The flexural resonance frequency of a flat plate depends on thickness h, density ρ_p, Young's modulus E, and Poisson's ratio μ. The plate bending wave number is

$$k_p = \frac{2\pi}{\lambda_p} = \frac{\sqrt{\omega}}{\alpha} \text{ and } \alpha^4 = \frac{Eh^2}{12(1-\mu^2)\rho_p}.$$

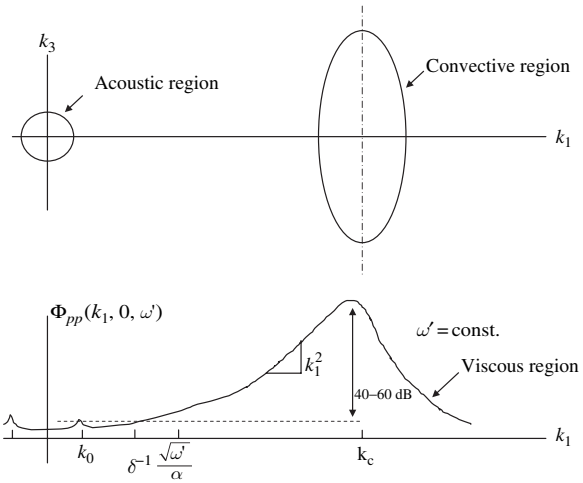


FIGURE 3 General features of the wave vector–frequency spectrum for the fluctuating wall pressure under a TBL.

A $\omega - k_1$ plot of Equation (6) is schematically shown in Figure 4.

The following observations about the TBL wall pressure fluctuation wave vector–frequency spectrum may be made:

1. The wave number spectrum peaks at the convective wave number k_c. The pressure fluctuations within this peak region are propagating subsonically. Because it is not true acoustic noise, it is oftentimes referred to as *pseudo noise*.

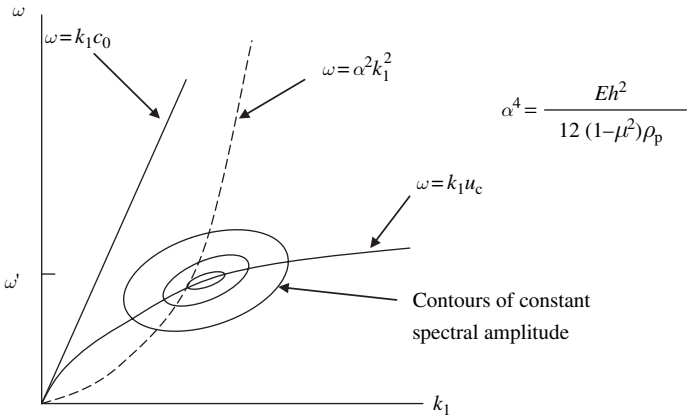


FIGURE 4 Streamwise wave number–frequency plot of the wall pressure fluctuation contours, showing also the behavior of the dispersive and non-dispersive components.

2. The acoustic wavenumber k_0 is different from k_c by a multiplicative factor equal to the convective Mach number $M_c = u_c/c_0$.
3. Any energy at or below the acoustic wavenumber radiates as true sound. The level of radiated sound is some 40–60 dB below the level of the pseudo noise at the convective peak.
4. Under subsonic conditions, energy that lies between k_0 and k_c is called *subconvective*, or low-wave number energy. This energy is capable of exciting real structures into flexural vibration.
5. Low-wave number wall pressure fluctuations excite marine vehicles into vibration because the plating of such vehicles is thick and massive which means that their resonant wavenumbers are low. Because the Mach number is low, energy at the convective peak is inefficient at exciting such plating into vibration.
6. If resonant vibration is induced in real plating by the TBL, the vibration can re-radiate true sound that can be sensed by sonar transducers.

5 • Transducer Size and Array Effects

Sonar pressure transducers placed directly under a TBL will respond to the wave number spectrum indicated in Figure 3 according to

$$\Phi_m(\omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi_{pp}(k_1, k_3, \omega)|H(k_1, k_3)|^2 T(k_1, k_3, \omega)dk_1 dk_3, \quad (7)$$

where $\Phi_m(\omega)$ is the transducer output frequency spectrum, $|H(k_1, k_3)|^2$ is the wave vector response function of the transducer, and $T(k_1, k_3, \omega)$ is the transfer function due to a coating of thickness h between the transducer and the TBL. The reader is referred to Ko (1992) for a complete discussion of the effects of a rubber layer and specific examples for $T(k_1, k_3, \omega)$.

Equation (7) with $T(k_1, k_3, \omega) = 1$ represents the response of a transducer of finite dimensions to TBL wall pressure fluctuations. This case was first studied by Uberoi and Kovasznay (1953) and extended by Corcos (1963) and Ko (1993). Its application to experiments performed in various media may be found in Capone and Lauchle (1995). If the transducer is circular of radius R, then

$$H(KR) = \frac{2J_1(KR)}{KR}, \quad (8)$$

where $K = \sqrt{k_1^2 + k_3^2}$. If R is small, high convective wave numbers are resolved by the transducer, and the pseudo noise dominates the output spectrum. If R is large, the high wave number convective ridge is not very well resolved, but the transducer continues to pass low wave number acoustic information. For this reason, sonar sensors placed under turbulent flow are necessarily large; they are less responsive to the convective hydrodynamic pressure fluctuations. The graphs of Figure 5 illustrate this point.

Further suppression of TBL wall pressure fluctuations can be achieved by using arrays of large transducers. Consider a group of N circular transducers placed in a streamwise line (x_1-direction) flush to the surface of a vehicle under its TBL as sketched in Figure 6. The signals from the array elements are processed to enhance the response of the array to acoustic sources while suppressing the response to TBL wall pressure fluctuations. The measured output spectrum of this array would be a result of convolving the TBL wall pressure wave vector–frequency spectrum with both a transducer wave vector response function and an array wave vector response function.

For the streamwise direction, Equation (7) is solved after replacing $|H(k_1, k_3)|^2$ with $|W(k_1)|^2 = |H(k_1)|^2 |A(k_1)|^2$. Equation (8) is used for $H(k_1)$ with $k_3 = 0$. The array wave number response function is

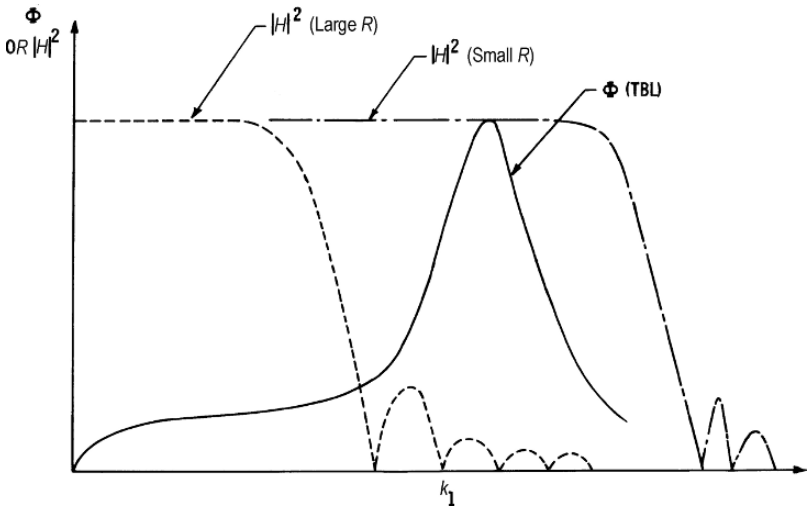


FIGURE 5 Wave number response of a circular transducer of radius R which is convolved with the wave number spectrum of the turbulent boundary layer (TBL) wall pressure fluctuations.

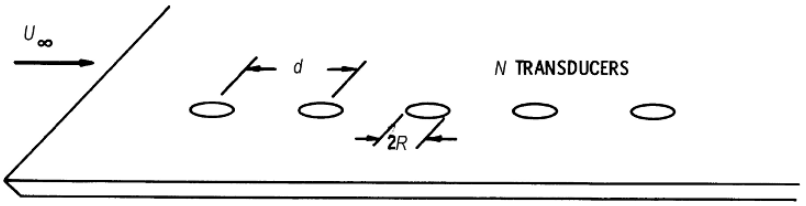


FIGURE 6 Simple line array of flush-mounted transducers oriented in the streamwise direction. This can be called a *wave vector filter*.

$$|A(k_1)|^2 = N^{-2}\left|\sum_{n=0}^{N-1} s_n e^{-ink_1 d}\right|^2, \tag{9}$$

where s_n describes the polarity of the transducers. It is a *shading function*. For an *in-phase* sum of the transducer outputs, $s_n = (-1)^{2n}$, and

$$|A(k_1)|^2 = \frac{1}{N^2}\frac{\sin^2(Nk_1 d/2)}{\sin^2(k_1 d/2)}. \tag{10}$$

Figure 7 shows a sample calculation of Equation (10) for a six-element array.

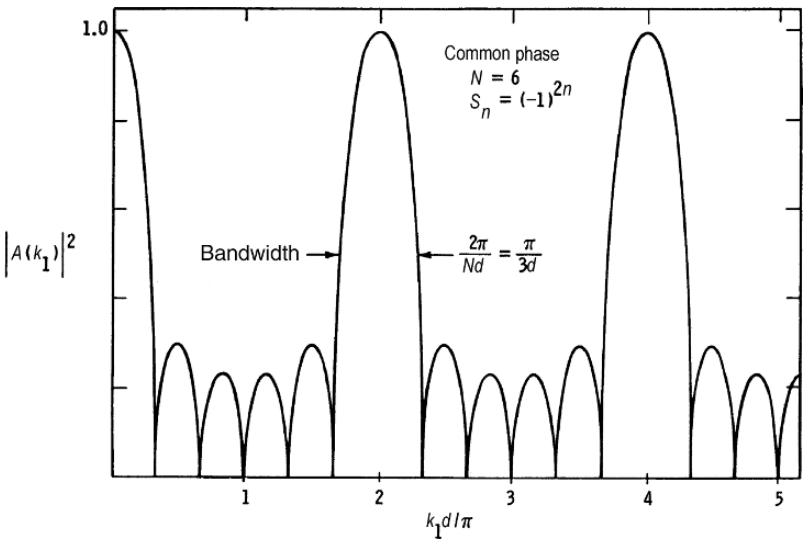


FIGURE 7 Wave number response function for a six-element line array oriented in the x_1-direction. The elements have common phased polarity.

It can be shown that Equation (7) has its first zero at $k_1R = 3.83$. If this zero were to align with the second major lobe of Equation (10), given by $k_1d = 2\pi$, we would have an array filter response that passes low, acoustic wave numbers while suppressing higher ones. For the six-element array, this condition occurs when $d = 1.64R$. Sonar designers would utilize this type of analysis to design arrays that suppress the high wave number TBL wall pressure fluctuations.

References

V. H. Arakeri, S. G. Satyanarayana, K. Mani, and S. D. Sharma, "Studies on scaling flow noise received at the stagnation point of an axisymmetric body," J. Sound Vib. **146**, 449–462 (1991).

D. E. Capone and G. C. Lauchle, "Calculation of turbulent boundary layer wall pressure spectra," J. Acoust. Soc. Am. **98**, 2226–2234 (1995).

D. M. Chase, "Modeling the wavevector-frequency spectrum of turbulent boundary layer wall pressure," J. Sound Vib. **70**, 29–67 (1980).

D. M. Chase, "The character of the turbulent wall pressure spectrum at subconvective wavenumbers and a suggested comprehensive model," J. Sound Vib. **112**, 125–147 (1987).

G. M. Corcos, "Resolution of Pressure in Turbulence," J. Acoust. Soc. Am. **35**, 192–199 (1963).

G. M. Corcos, "The structure of turbulent pressure field in boundary layer flows," J. Fluid Mech. **18**, 353–378 (1964).

J. E. Ffowcs Williams, "Hydrodynamic Noise," Ann. Rev. Fluid Mech. **1**, 197–222 (1969).

J. E. Ffowcs Williams, "Surface-pressure fluctuations induced by boundary-layer flow at finite Mach number," J. Fluid Mech. **22**, 507–519 (1965).

S. H. Ko, "Application of elastomeric material to the reduction of turbulent boundary layer pressure fluctuations (three dimensional analysis)," J. Sound Vib. **159**, 469–481 (1992).

S. H. Ko, "Performance of various hydrophones in the reduction of turbulent flow noise," J. Acoust. Soc. Am. **93**, 1293–1299 (1993).

R. H. Kraichnan, "Pressure fluctuations in turbulent flow over a flat plate," J. Acoust. Soc. Am. **28**, 378–390 (1956).

G. C. Lauchle, "Hydroacoustics of transitional boundary-layer flow," Appl. Mech. Rev. **44**, 517–531 (1991).

G. C. Lauchle, "Flow noise scaling at the stagnation point of an axisymmetric body," J. Sound Vib. **154**, 568–572 (1992).

G. C. Lauchle, J. F. McEachern, A. R. Jones, and J. A. McConnell, "Flow-induced noise on pressure gradient hydrophones," in *Acoustic Particle Velocity Sensors: Design, Performance, and Applications* (M. Berliner and J. Lindberg, eds.). AIP Conference Proceedings **368**, pp. 202–225 (1995).

J. A. McConnell, "Developments and applications of inertial type underwater acoustic intensity probes," Ph.D. thesis in Acoustics, The Pennsylvania State University (2004).

J. F. McEachern and G. C. Lauchle, "Flow-induced noise on a bluff body," J. Acoust. Soc. Am. **97**, 947–953 (1995).

M. S. Uberoi and L. S. G. Kovasznay. "On mapping and measurement of random fields," Qart. Appl. Math., **10**, 375–393 (1953).

PART XVII
ANIMAL BIOACOUSTICS

Hiroshi Riquimaroux

Sensory & Cognitive Neuroscience Research Laboratory, Hiroshi Riquimaroux, Doshisha University, Kyoto, Japan

100 Bat Bioacoustics

Walter Metzner 1835

Introduction • Recording and Monitoring Systems • Call Characteristics • Echolocation Performance and Models • Hearing Performance and Models • Stimulus generation • Archives and Databases

101 Fish Bioacoustics

Richard R. Fay 1851

Recording and Monitoring Systems • Hearing Performance • Vocalization Sounds • Locating and Tracking • Archives and Databases

102 Amphibian bioacoustics

Jakob Christensen-Dalsgaard 1861

Recording and Monitoring Systems • Hearing Performance and Modelling • Characteristics of Anuran Calls • Stimulus Generation – Synthesized Calls • Archives and Databases of Signals

103 Avian Bioacoustics

Kazuo Okanoya 1887

Recording and Monitoring Systems • Hearing Performance • Characteristics of Songs and Calls • Archives and Databases of Signals

104 Cetacean Bioacoustics with Emphasis on Recording and Monitoring

Tomonari Akamatsu 1897

Recording and Monitoring Systems • Locating and Tracking • Echolocation Models

100 Bat Bioacoustics

Walter Metzner

Department of Physiological Science, University of California, California, USA

1 • Introduction

Although bats rarely come to man's attention, they belong to one of the most abundant groups of mammals. More than 1,100 species of bats are distributed over nearly the entire world, except for the most extreme desert and polar regions, amounting to approximately 20% of all living species of mammals. By exploiting the nocturnal three-dimensional space, they gained access to the rich resources of nocturnal aerial life while almost completely avoiding competition from the other main group of flying predators, the birds.

There are two different groups of bats, the megachiropteran bats ("megabats"), also known as flying foxes or fruit bats (some 170 species) and the microchiropteran or bats in the common sense (>900 species). Megabats are frugivorous and predominantly use their well-developed vision (and presumably also their sense of smell) to orient and find food. Microchiropteran bats, on the contrary, do not rely on vision for orientation and to find their prey, which consists mainly of insects. Lazzaro Spallanzani, the Bishop of Padua and an accomplished experimentalist, already described this ability of bats to orient in complete darkness in the late eighteenth century. However, it took more than 150 years until around 1940, independently, Donald Griffin in the USA and Sven Dijkgraaf in the Netherlands first described echolocation in bats.

Microchiropteran bats are not the only group of animals that uses echolocation for orientation purposes. Amongst the

megabats, a very simple form of echolocation is found in the cave-dwelling Egyptian fruit bat, *Rousettus*. Some shrews as well as whales, dolphins, and seals also echolocate. Finally, echolocation independently evolved at least twice among birds: South American oilbirds and some cave swiftlets in Southeast Asia produce tongue clicks to find their way through the darkness of the caves where they roost.

2 • Recording and Monitoring Systems

Most bats produce vocalizations for two major purposes: echolocation calls allow the bats to navigate and detect prey, and communication calls enable them to exchange signals in the context of social interactions. For both call types, the major spectral components are above the range of human hearing, and thus the recording equipment has to be able to either sample the sound signals at rates that are sufficiently high to conform to the sample theorem (at least twice the call's upper spectral frequency limit). Most commonly, however, especially in field studies, which often necessitate the use of numerous small and versatile monitoring devices, special acoustic devices have been developed that electronically transfer ultrasonic sounds into frequencies that are audible to humans. Different design principles are available for such "bat detectors", of which the three main types are based on frequency division, heterodyning, and time expansion.

Frequency division employs an electronic technique in which only a certain fraction of the original waveform is sampled, e.g. every tenth zero-crossing. Hence an original call that sweeps from 70 kHz down to 30 kHz will be transformed into a signal that now ranges from 7 to 3 kHz. While this rather simple and straightforward technique allows for the design of inexpensive and easy-to-use bat detectors, it also has some disadvantages when a detailed call analysis is required. The most serious drawbacks are that due to the "under-sampling" the original waveform and any information about the calls' harmonic composition are lost.

Similarly, devices that are based on heterodyning, while relatively simple and inexpensive, do not retain such detailed temporal or spectral information. Heterodyning refers to a technique in which the original signal is electronically mixed (i.e. multiplied) with an electronically generated signal of one particular frequency that can be determined by the user. The output of the heterodyning device is the result of the electronic multiplication and corresponds to the frequency difference between incoming sound signal (i.e. bat call) and electronic signal (dialed in by the user). If a bat calls at 80 kHz, for instance, and the electronic signal is set at 82 kHz, the heterodyning device

generates an output signal at 2 kHz. Such heterodyning bat detectors are very useful for determining the frequency value of narrowband calls, such as of horseshoe bats (see below), but allow for detecting wide band signals only if the frequency of the electronic signals is centered within or very close to the spectral bandwidth of the bat's call.

Finally, time-expansion bat detectors avoid the disadvantages of the previous two types of bat detectors, however, at significant costs. These digitally operating bat detectors are based on a circular buffer that can, depending on the sample rate, store up to several minutes of recordings and play the signals back at a reduced rate, while maintaining their original waveform at an expanded time scale. Most time-expansion bat detectors also include frequency division and heterodyning options.

3 • Call Characteristics

Not all echolocating bats use the same type of call. There is a wide variety of sounds. They share some features, however, and most of them contain broadband or narrowband components or combinations of them (Figure 1). Broadband calls normally are short, downward frequency-modulated (FM) signals. Typically, call durations are under1 ms and the bandwidth covers at least half an octave. FM calls often contain one to several harmonic components which stretch the frequency band covered by this signal even wider.

There are two types of narrowband calls: quasi-constant frequency (QCF) calls, in which the frequency sweeps, usually downward, only over a few kilohertz between call onset and offset, and long constant frequency (CF) signals with even smaller frequency changes of only a few hundred hertz. Long CF calls,

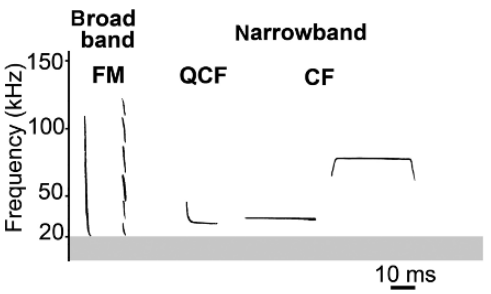


FIGURE 1 Schematic spectrograms of basic types of echolocation calls used by echolocating bats. For details, see text. The range of human hearing is indicated by the gray area below 20 kHz. FM: frequency modulated; QFM: quasi-frequency modulated; CF: constant frequency.

which normally have durations of several tens of milliseconds, can also contain brief FM components at their beginning and end (FM–CF–FM calls). Some bats that emit long CF calls possess a very sophisticated echolocation system (see below), such as the old world horseshoe and leaf-nosed bats (*Rhinolophus* spec. and *Hipposideros* spec.), and the new world Parnell's mustache bat (*Pteronotus parnellii*).

Since most bats usually do not forage opportunistically in all habitat types, but instead occupy a particular ecological niche, the different types of echolocation calls used by different bat species often correlate with a certain type of habitat (Neuweiler, 2000; Schnitzler and Kalko, 2001). Even when bats do change their foraging habitat, they often change the structure of their echolocation signals as well. When they fly in open space at high speeds (10–15 m/s), for example above canopy, most species use shallow FM signals at rather low frequencies (below 50 kHz). Some of them are even audible to the human ear. The lower the frequency of echolocation sounds, the longer is their travel distance before their intensity is substantially reduced by atmospheric attenuation. By concentrating most of the sound energy into a rather narrow band at low frequencies, these bats are able to listen to echoes that return from targets that are up to 50 m away. Thus, they can take full advantage from the wide open space by covering a wide area at high speed and by using a far-reaching sonar system.

Many bats that hunt for insects close to surfaces, for example within foliage or close to the ground, shorten their echolocation calls considerably to avoid an overlap between the echo and the emitted sound. When bats, for example, move from their foraging areas above widely open grasslands to their hunting grounds between canopies they switch from long and narrowly band QFM or even CF calls to brief broadband FM signals.

4 • Echolocation Performance and Models

Echolocating bats navigate and detect prey by comparing each outgoing call with the returning echo(es). Generally, target information is encoded in amplitude and frequency changes as well as in the arrival time of the returning echo relative to the emitted call.

4.1 Range Discrimination

The major strategies for determining target distance are to measure the time lapse and/or the intensity difference between outgoing call and returning echo. Bats can resolve temporal differences in echo delays between 50 and 90 μs, which translates into a difference in target distance of only 8–15 mm (Moss

and Schnitzler, 1995). To determine echo delays accurately, the bats rely on the precise timing of call emission and echo arrival. Since long CF calls often temporally overlap with their echoes, they do not provide an appropriate basis for range determination. Instead, bats use brief broadband FM signals to measure object distance. Thus, mustache and horseshoe bats listen to their FM components that follow the long CF parts to evaluate target range.

The time delays of returning echo signals are extracted in the central auditory system by a precise "stopwatch mechanism" that is thought to be based on coincidence detection, i.e., a synchronous excitation by an emitted sound and a subsequent echo signal. Such a coincidence detector may involve delay lines, inhibitory activity, and/or paradoxical latency shift (Suga, 1997; Galazyuk et al., 2005; Nataraj and Wenstrup, 2005). An array of delay lines could delay the arrival of excitation from the first signal, i.e., the emitted call, at the coincidence detector so that in at least one of the delay line channels, the first excitation matches the arrival of excitation from echoes returning over a wide range of different delays. While there still is no firm experimental evidence for the existence of such delay lines in the bat auditory system, the latter two mechanisms clearly appear to be involved. Different types of inhibitory auditory activity, for example, contribute to generating the delay-sensitive responses observed in many auditory neurons, and neurons exhibiting paradoxical latency shift, i.e., longer (instead of shorter) response latencies for intense sounds than for weaker sounds, are well suited for construction of delay-tuned neuronal responses.

Since echolocation is mainly used in flight and bats regularly fly at speeds between 3 and 8 m/s, Doppler effects will result in a compression of the time between call emission and echo return (as well as increase echo frequencies by up to 5%). This leads to a distortion of the perceived target range. While some bats emit Doppler-tolerant calls (that follow a hyperbolic frequency–time course), thus avoiding this ranging problem, other bats might adjust their call design to the target distance (Boonman et al., 2003).

4.2 Object Discrimination

The use of high frequencies is essential for resolving acoustic images of small objects, such as insect prey. Similar to optical resolution, in acoustic imaging, the wavelength of the echo signal has to be no more than half the size of the object to be resolved.

Bats can also discriminate between objects of different form and size. Obstacle avoidance tests showed that bats can detect wires with diameters of less than 100 μm, i.e., less than 1/100 of the wavelength of the returning echo signal. In object

discrimination tasks, bats can detect size differences of approximately 17%, which corresponds to sound level differences in the returning echo of 1–3 dB (Neuweiler, 2000), suggesting that the primary cue for discriminating different target sizes is echo intensity.

Echo signals presumably also carry information about the form and texture of objects: bats can discriminate between same-sized objects that differ in shape, surface structure, or even the material they are made of (e.g., Plexiglass vs. plastic). The bat's ability to discern an object's texture using a broadband sonar signal could be based on two different acoustic clues: analogous to the phenomenon of refraction of white light into its spectral components, different target textures cause echo signals with different interference patterns, resulting in "echo colors" that are texture specific (Neuweiler, 2000). In contrast to such spectral cues, a bat might also be able to use arrival time differences in the echo wave fronts returning from the different planes or scattering points of the textured object. It has been demonstrated that bats are indeed able to detect changes in echo delay ("echo jitter") of a little as 10–40 ns using strictly temporal cues (Simmons et al., 2003). Although the physiological basis for such remarkable temporal hyperacuity is unknown, recent modeling studies have provided suggestions as to how such delay accuracy might be achieved by the peripheral auditory system (Sanderson et al., 2003). Inner-ear transduction was modeled with a bank of parallel band-pass filters followed by low-pass smoothing. This model achieves a delay acuity of 17 ns when tested with a signal-to-noise ratio of 50 dB. The modeling results show that smoothing by low-pass filtering in the bat's inner ear strongly affects the accuracy of such a filterbank receiver. They suggest that only a modest rise in the frequency response of smoothing in the bat's inner ear can confer full phase sensitivity on subsequent processing and account for the bat's extraordinary temporal sensitivity.

A highly specialized case of object detection is found in those bats that produce long CF calls (see call characteristics above), such as horseshoe bats (*Rhinolophus* spec.) and the mustache bat (*Pteronotus parnellii*) (Neuweiler, 2000). These bats are highly specialized insect hunters. They forage close to or even within dense vegetation and thus have to deal with heavy echo cluttering caused by reflections from the background vegetation. The long, narrowband signals enable them to detect intensity and frequency modulations in the returning echo signal, so-called acoustic glints, that are caused by the wig beats of flying insect prey. Fluttering insects can cause frequency modulations of up to ± 2 kHz around the echo's carrier frequency and the amplitude

can be modulated by more than 75%. Observations in the natural habitat of these bats as well as experiments in the laboratory have demonstrated that they, indeed, forage only on wing-beating insects. Even a single wing movement will elicit prey catching behavior while non-fluttering targets are ignored.

The acoustic glints in the echoes that return from flying insects vary between different insect species. Therefore, the bat could use them not only to detect but also to identify its prey. However, the structure of these echo modulations changes with the reflection angle between bat and prey. As a result, the bat has to choose among an immense variation of echo signals. Nevertheless, behavioral experiments show that horseshoe bats are able to distinguish between different insect species, regardless of the echo's reflection angle. They can even recognize echoes of particular insect species that are reflected at angles never presented before. Thus, the echo modulations that are caused by the prey's wing movements obviously contain all the information necessary for the bat to detect and even recognize its prey.

The basis for this remarkable echolocation ability is provided by specializations within the bat's auditory system. A filter mechanism that is narrowly tuned to the echo's narrowband frequency component rejects background clutter while it helps to detect acoustic glints. This auditory filter is found in the cochlea as well as in every subsequent auditory structure of the central nervous system, and results in an increased number of receptor cells and neurons, respectively, that are very sensitive to the echo pure tone. Since in the mammalian eye and within the visual central nervous structures, the area around the optic fovea is overrepresented in a very similar manner, this auditory filter in horseshoe and mustache bats is similarly called an auditory fovea.

4.3 Doppler Shift Compensation Behavior

Horseshoe and mustache bats inevitably face one particular problem by using such narrowband echolocation signals. During flight, the movement of the bat relative to the background, such as dense vegetation, induces Doppler shifts in the echo frequency. In contrast to those Doppler effects that are caused by fluttering insects, the bat's flight causes shifts of the whole echo away from the frequency that is emitted in resting bats and that matches the frequency of the bat's auditory fovea. As a consequence, the carrier frequency of the faint echo modulations that indicate prey drops out of the auditory fovea and the bat is no longer able to detect essential echo information.

Therefore, these bats continuously adjust the frequency of their echolocation calls to compensate for Doppler effects in the

returning echo. Interestingly, this Doppler shift compensation (DSC) behavior (Schnitzler, 1968) has independently evolved in horseshoe and in mustache bats. DSC behavior has its equivalent in visual foveation, where moving the eyes keeps images of interest fixated on the fovea. DSC behavior allows the bats to accurately compensate for changes in echo frequency up to 5 kHz, but they do so through a sequence of small, temporally independent, step changes in call frequency. The relationship between an echo's frequency and its subsequent impact on the frequency of the very next call is fundamental to how DSC behavior works. Such important aspects of DSC control are (1) DSC is only elicited by changes in echo frequency that exceed a threshold, (2) during DSC single changes in call frequency vary in magnitude and polarity, and (3) there seems to be an upper limit to the amount of change in frequency that can occur between successive calls.

A detailed quantitative analysis of DSC allows one to develop a simple mathematical representation of this behavior which can well predict DSC performance under different experimental conditions, and also offers insight into the neurophysiology of DSC in horseshoe bats (Smotherman and Metzner, 2003a). The model, illustrated in Figure 2, is based on similar three assumptions as listed above. It describes how changes in call frequency in response to varying echo frequencies are mediated by a pair of (echo)frequency-specific sigmoidal functions that are characterized by (1) a threshold, (2) a slope representing the increasing magnitude of changes in call frequency relative to increasing changes in echo frequency ($\Delta F_{\text{voc}}/\Delta F_{(\text{echo})}$), and (3) a maximum possible single change in call frequency $\Delta F_{\text{voc}_{\text{max}}}$. Taken collectively, these values provide the framework for a simple three-parameter sigmoidal curve summarized as

$$\Delta F_{\text{voc}} = \Delta F_{\text{voc}_{\text{max}}}/(1 + e^{-(F_{\text{echo}} - F_{\text{echo}_{\text{mid}}})/b)}$$

where $\Delta F_{\text{voc}_{\text{max}}}$ is the maximum single change in call frequency, F_{echo} represents the relative echo frequency, $F_{\text{echo}_{\text{mid (midpoint)}}}$ is the echo frequency midpoint of the slope, and b is the maximum slope between the minimum and maximum single changes in call frequency (representing the gain of the system). In this equation, the threshold is most closely represented by the midpoint of the slope, $F_{\text{echo}_{\text{mid}}}$.

This curve fit analysis can provide a meaningful reflection of how DSC occurs, and it can be useful in the analysis of experimental manipulations of DSC. For example, it has been demonstrated that DSC threshold is highly sensitive to the intensity of the echo (Smotherman and Metzner, 2003b). The effects of

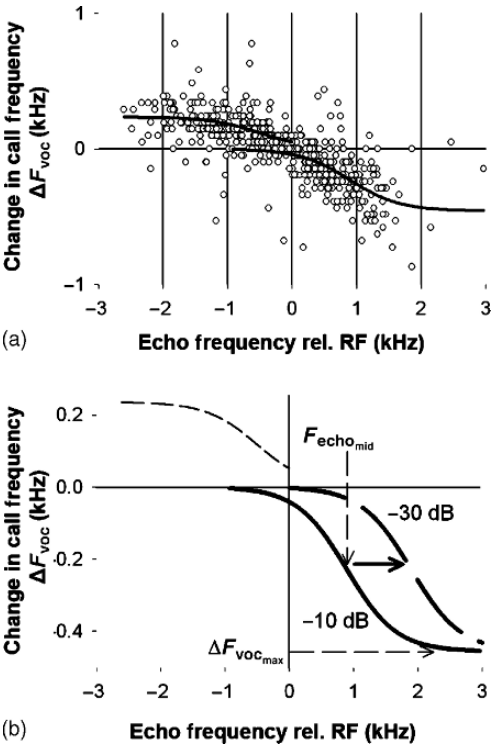


FIGURE 2 Simple mathematical interpretation of the relationship between echo frequency and subsequent changes in call frequency for variations in echo intensity (Smotherman and Metzner, 2003a). (**a**) Open circles indicate changes in call frequency in response changing echo frequencies (given rel. bat's "resting frequency" RF, i.e., call frequency when not exhibiting DSC) presented at intensities attenuated by 10 dB rel. emitted call intensity. Solid lines: Two sigmoidal curves of the form $\Delta F_{voc} = \Delta F_{voc_{max}}/(1+e)(F_{echo} - F_{echo_{mid}})/b)$ (for term definitions see text) were fit to data above and below RF, separately. For the response to echoes above RF, the change in F_{voc} was optimally fit using the following parameters: $\Delta F_{voc_{max}} = 0.456$ kHz, $b = 0.387$, and $F_{echo_{mid}} = 0.895$ kHz rel. RF ($R^2 = 0.28$). For echoes below RF, the best curve fit required values of $\Delta F_{voc_{max}} = 0.339$ kHz, $b = 0.332$, and $F_{echo_{mid}} = 0.972$ ($R^2 = 0.13$). (**b**) Effects of changes in echo intensity. The curve shown in (a) for echoes returning above RF at −10 dB is superimposed with a curve that approximates the data for the response to echoes above RF at −30 dB. Lower echo intensities (−30 dB vs. −10 dB) require higher echo frequencies to yield the same change in call frequency. This lateral shift to higher echo frequencies can be summarized by a single parameter change, the addition of approximately 1 kHz to the $F_{echo_{mid}}$ value (indicated by the horizontal arrow).

echo intensity on the relationship between echo frequency and subsequent changes in call frequency can be represented mathematically by a imposing a change in the value of single parameter, $F_{\text{echo}_{\text{mid}}}$.

The exact values of these model parameters necessarily reflect properties of the underlying neural circuitry of DSC and the motor control of vocalization, and provide insight into how neural feedback can accommodate the need for speed without sacrificing stability (Smotherman and Metzner, 2003a).

5 • Hearing Performance and Models

The auditory system of bats from the cochlea to the auditory cortex is made up of the same elements as that of other mammals, although the structures are generally larger in echolocating bats and some frequencies may be overrepresented, as for example in horseshoe and mustache bats (see above).

The input stage of the auditory system of most bats (except for bats emitting long CF signals, such as horseshoe and mustache bats) can be described as a neuronal filter bank with similar Q values (frequency at the lowest response threshold divided by bandwidth), and is therefore similar to that of any mammalian hearing generalist. The information that can be extracted from the echoes of various signal elements depends on their physical structure and on the bat's auditory performance. Depending on the signal's absolute frequency, bandwidth, and sound level, a signal element evokes neuronal activity in one or more of the frequency-selective filters, and the longer a signal remains within the frequency range of a filter channel, the higher is the neuronal activity in this channel.

Bats have been shown to localize brief sounds with high accuracy. The neural mechanisms for detecting the direction from where an echo returns from are thought to be similar to those employed by other mammals to locate sound sources in space. Azimuthal direction (i.e., the location in the horizontal plane) can be extracted from arrival time differences between the two ears and/or by interaural sound level differences. In addition, monaural spectral cues caused by interference patterns generated in the pinna are available as well. Considering the small size of the bats' head, the interaural time difference is very small (on the order of 50 μs). In addition, the ultrasonic frequencies do not allow them to analyze phase information for determining arrival time differences, as is the case in many birds. It is therefore generally thought that interaural level differences and monaural spectral cues, especially in the high-frequency range, represent

the primary cues for bats when localizing sound in azimuth (Wotton et al., 1995; Aytekin et al., 2004). Modeling studies have produced an electronic model simulating many aspects of neuronal responses observed in neurons encoding for different sound locations in azimuth (Horiuchi and Hynna, 2001).

Interaural level differences and monaural spectral cues are also the chief clues for sound localization in elevation. A bat-specific feature of the pinna, the tragus, which is a skin flap that protrudes upward from the lower end of the external ear canal, has been implicated in aiding particularly sound localization in elevation by providing monaural high-frequency spectral cues due to acoustic interference (Wotton et al., 1995).

Many bats (like other mammals) have extremely mobile pinnae, which conceivably increases sound localization ability.

6 • Stimulus generation

Echolocation and communication signals in microchiropteran bats are generated, like vocalizations in most other mammals, in the larynx. (The simple echolocation signals of megabats are, in contrast, produced by clicking the tongue.) The general architecture of the larynx of bats is identical to that of other mammals. The stream of expired air that passes through the larynx provides the sound energy. During expiration, air pressure builds up beneath the closed glottis, and when the glottis opens, sound is produced. Bat echolocation calls are extremely loud, reaching up to 120 dB SPL (sound pressure level). Accordingly, the subglottic pressures are more than five times that measured in humans, and necessitate that the lips of the vocal cords as well as the cricothyroid membranous wall are much thicker than in other mammals, enabling them to withstand and maintain these high subglottic pressures even during the emission of several calls during one expiratory cycle (Neuweiler, 2000).

Another specialization unique to bats are the vocal membranes, which are positioned along the rims of the vocal folds. They are only a few microns thick and contain no muscles or fibers. These vocal membranes are believed to allow for the generation of the high frequencies typical for bat echolocation signals, especially since echolocating bats lack the vocalis muscle that is responsible for generating sound frequencies in other mammals. The precise mechanism, however, is still unclear.

The absence of a vocalis muscle and the presence of a non-muscular vocal membrane also has consequences for how the frequency of vocalizations is generated. Instead of using the direct action of the vocalis muscle to change the tension on the vocal

cords (high tension resulting in high call frequencies), echolocating bats have to rely on an indirect mechanism and change the tension on the vocal folds by tilting the thyroid cartilage through the action of the cricothyroid muscle. When the cricothyroid muscle contracts, the thyroid tilts downward tautening the vocal cords, which results in higher call frequencies.

Therefore, precise timing of the inner laryngeal muscles and the cricothyroid muscle determine the opening and closing of the glottis, yielding call beginning and end, and the call's frequency, respectively. This coordinated activity of the different muscle sets results in the various types of echolocation calls emitted by bats (see Figure 1). This simple model of sound production is illustrated in Figure 3 (Pye, 1980). When the glottis opens during relaxation of the cricothyroid muscle from a previous contraction, downward FM calls are generated. The FM–CF–FM call type that is characteristic for horseshoe bats, for example, is a result of opening of the glottis while the muscle is contracting and then remaining open during the contracted state until finally closing during muscle relaxation.

Resonating cavities of the supraglottic air passages are also involved in sound production by introducing harmonic components of the fundamental frequencies generated in the larynx. Especially in bats that emit echolocation calls through the nostrils, such as horseshoe and leaf-nosed bats, these resonance cavities can shift the dominant energy component away from the fundamental frequency component to higher integer multiples of it, such as the second harmonic in horseshoe bats.

It appears that that such a sophisticated laryngeal echolocation behavior evolved only once in the common ancestor of all bats. Interestingly, it was subsequently lost in megabats (Teeling et al., 2005), in which then one genus, the Egyptian fruit bat,

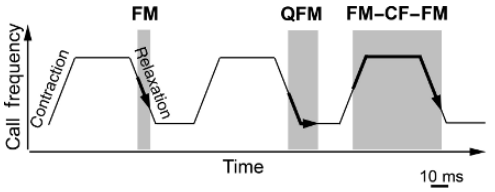


FIGURE 3 Model of sound production, modified after Pye (1980). The timing and the duration of glottis openings and closings (gray areas) with contraction and relaxation of the cricothyroid muscle (indicated by the trapezoidal contraction and relaxation pattern), which controls call frequency, creates calls with different spectrotemporal patterns (thick arrows). Compare with Figure 1 for call types produced by echolocating bats.

Rousettus, "re-invented" a simpler form of echolocation by using tongue clicks as sonar signals. This enables these megabats to exploit dark caves as roost sites, which are inaccessible for all other, strictly visually guided fruit bats.

7 • Archives and Databases

Although bat detectors are now readily accessible and used by many bat researchers and amateurs to identify bat species and to study their activity profiles, libraries of bat echolocation calls are very sparse. While this lack of data management has already been recognized, there is still no general, globally accessible database (Korine and Kalko, 2001). The authors suggest establishing such a database that describes call characteristics using standardized recording and analysis techniques. The database should also link the sequences of sonar calls to behavioral observations, which would allow one, for example, to assess the degree of signal plasticity caused by different environmental conditions. The database should be structured into three portions: (1) recording methods and equipment; (2) general description of the collected data, such as the time and location of the observation/recording, weather conditions, and descriptions of the recording situation; and (3) general description of call structure and echolocation patterns, e.g., behavioral context of call emissions, analysis methods used, and detailed call characteristics. Such a database would be an invaluable tool not only for basic bat research but it also for global issues in bat conservation by providing information on species identification, community comparisons, and geographic distributional patterns of bats.

Only a few local bat databases exist to date. However, they do not conform to the standards outlined above (Korine and Kalko, 2001) and some of them only accept data sets obtained by one particular recording technique or equipment. The most comprehensive database so far is the "Bat Call Library", which is managed by the Division of Mammals in the Museum of Southwestern Biology. It is part of the Department of Biology, University of New Mexico. As of August 2005, it contains files of 7,855 calls. It intends to serve as a central distribution site for reference, research, and exchange of calls. However, it accepts only calls recorded using the Anabat hardware and software (http://www.msb.unm.edu/mammals/batcall/html/calllibrary.html).

Another local database is the "South-eastern Australian bat call library", maintained and edited by A. Herr and N. Klomp

of the Johnstone Centre of Parks, Recreation and Heritage at Charles Sturt University, New South Wales, Australia (http://batcall.csu.edu.au/batcall1.html).

Finally, a small database is "Batcalls.org", which was developed to facilitate sharing of time-expanded echolocation calls. It is maintained by Mary Wenzel, Goleta, California, USA (http://www.batcalls.org/).

Acknowledgments Supported by NIDCD and the Whitaker Foundation.

References

Aytekin M, Grassi E, Sahota M, Moss CF (2004) The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation. J Acoust Soc Am 116:3594–3605

Boonman AM, Parsons S, Jones G (2003) The influence of flight speed on the ranging performance of bats using frequency modulated echolocation pulses. J Acoust Soc Am 113:617–628

Galazyuk AV, Lin W, Llano D, Feng AS (2005) Leading inhibition to neural oscillation is important for time-domain processing in the auditory midbrain. J Neurophysiol 94:314–326

Horiuchi T, Hynna K (2001) Spike-based VLSI modeling of the ILD system in the echolocating bat. Neural Networks 14:755–762

Korine C, Kalko EK (2001) Toward a global bat-signal database. IEEE Eng Med Biol Mag 20:81–85

Moss CF, Schnitzler HU (1995) Behavioral studies of hearing in bats. In: Hearing in bats (Popper AN, Fay RR, eds), pp 87–145. New York: Springer

Nataraj K, Wenstrup JJ (2005) Roles of inhibition in creating complex auditory responses in the inferior colliculus: facilitated combination-sensitive neurons. J Neurophysiol 93:3294–3312

Neuweiler G (2000) The biology of bats. New York: Oxford University Press

Pye JD (1980) Echolocation signals and echoes in air. In: Animal sonar systems (Busnel RG, Fish JF, eds), pp 309–333. New York: Plenum Press

Sanderson MI, Neretti N, Intrator N, Simmons JA (2003) Evaluation of an auditory model for echo delay accuracy in wideband biosonar. J Acoust Soc Am 114:1648–1659

Schnitzler H-U, Kalko EKV (2001) Echolocation by insect-eating bats. Bioscience 51:557–569

Schnitzler HU (1968) Die Ultraschallortungslaute der Hufeisennasen-Fledermäuse (Chiroptera, Rhinolophidae) in verschiedenen Orientierungssituationen. Z Vergl Physiol 57:376–408

Simmons JA, Ferragamo MJ, Sanderson MI (2003) Echo delay versus spectral cues for temporal hyperacuity in the big brown bat, *Eptesicus fuscus*. J Comp Physiol A 189:693–702

Smotherman M, Metzner W (2003a) Fine control of call frequency by horseshoe bats. J Comp Physiol Neuroethol Sens Neural Behav Physiol 189:435–446

Smotherman M, Metzner W (2003b) Effects of echo intensity on Doppler-shift compensation behavior in horseshoe bats. J Neurophysiol 89:814–821

Suga N (1997) Parallel-hierarchical processing of complex sounds for specialized auditory function. In: Encyclopedia of acoustics (MCrocker MJ, ed), pp 1409–1418. New York: John Wiley

Teeling EC, Springer MS, Madsen O, Bates P, O'Brien S J, Murphy WJ (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. Science 307:580–584

Wotton JM, Haresign T, Simmons JA (1995) Spatially dependent acoustic cues generated by the external ear of the big brown bat, *Eptesicus fuscus*. J Acoust Soc Am 98:1423–1445

101

Fish Bioacoustics

Richard R. Fay

Parmly Hearing Institute, Loyola University Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, USA, e-mail: rfay@luc.edu

1 • Recording and Monitoring Systems

There is no standard set of equipment and techniques for the study of hearing and sound production in fishes. The most important issues for both types of study is the use of hydrophone with calibration traceable to a standard, and a linear recording device (i.e., without automatic gain control) with sufficient bandwidth. For hearing studies, acoustic particle motion should be manipulated and measured. Measurement can be accomplished by monitoring pressure gradients with a hydrophone array. A stimulation system for simulating acoustic particle motion in the study of hearing of generalist species is described in Fay and Edds-Walton (1997).

2 • Hearing Performance

Bioacoustic behaviors of fishes depend on mechanisms for sound production and sound detection. These two capacities are likely to be correlated, but not "co-evolved". In the only study specifically aimed at uncovering evolutionary the relationships between sound production and sound reception, Ladich (1999, 2000) reached the conclusion that the development of sound reception characteristics probably predated sound production and communication among fishes, and presumably developed in response to general requirements that had little to do with processing

communication sounds. Fay (1998, 2000) has presented empirical and theoretical evidence that an important and universal capacity of the auditory system is for determining and segregating simultaneous sound sources for auditory scene analysis (Bregman, 1990). Lewis and Fay (2004) argue that this capacity is both computationally difficult and universally present in auditory systems, and as such probably played a role in shaping early brain evolution. However, once evolved, this capacity probably made acoustic communication possible as sound production mechanisms and signals developed to take advantage of this great sensory capacity to process and individuate sound sources. Thus, sound production and detection mechanisms are likely correlated.

Sound detection in fishes begins at otolith organs (saccule, lagena, and utricle) in which a dense otolith overlays a sensory epithelium (macula), made up of hair cells. The primitive or shared mode of hearing among fishes depends on the relative motion between the otolith and hair cells that takes place across the hair cell cilia by virtue of the otolith's inertia. Hair cell excitation occurs when the cell's cilia are deflected along a line between the center of the hair bundle and the eccentrically placed kinocilium. The stimulus for this primitive mode of hearing are hydrodynamic particle motions in the near field, and acoustic particle motion (the kinetic component of propagated sound waves) in the far field. Otolith organs have been shown to be sensitive to particle displacements as small as 0.1 nm at about 100 Hz (Chapman and Sand, 1974; Fay, 1984; Fay and Edds-Walton, 1997). The available evidence shows that the frequency range for this sort of direct displacement sensitivity ranges from below 100 Hz to several hundred hertz. All unspecialized fishes would be expected to hear somewhere in this range. Some fishes have specializations for hearing that consist of the swim bladder (or other gas bubble), and some mechanical pathway between the swim bladder and the ear that transmits motions of the swim bladder walls (caused by sound pressure fluctuations) to the otolith organs. This makes these species sound pressure sensitive: ears respond in proportion to the sound pressure level. These specializations for detecting sound pressure include rostral extensions of the swim bladder (at least some species of Holocentridae, Gadidae, Cichlidae, Notopteridae, Elopomorpha, some Percomorpha, Haemulidae, Nematistiidae, Sparidae, Centropomidae, Gerridae, Kuhlidae, Lactaridae, Moronidae, Percicthydae, Polyprionidae, Priacanthidae, Sciaenidae, and Scombridae), auditory bullae in Clupeomorpha (some with ultrasonic sensitivity – Mann and Lu, 1998; Mann et al. 2001), suprabranchial

chambers in anabantoidei, and the Weberian ossicle linkage between the swim bladder and ear (carps, suckers, minnows, knifefishes, catfishes).

Figure 1 shows selected behavioral frequency threshold curves (sound pressure thresholds) published for fishes. Figure 1 shows sound pressure thresholds for two groups of fish, those that are probably sensitive to sound pressure (thick solid lines), and those that are not, and probably respond only to acoustic particle motion (light lines). Note that these latter frequency threshold curves are not interpretable; the animals respond to particle motion, but thresholds are expressed in sound pressure units. In typical lab tanks where behavioral thresholds have been determined, the relation between sound pressure level and acoustic particle motion is indeterminate. Figure 2a shows behavioral displacement audiograms for *Gadus morhua* – Atlantic cod, *Pleuronectes platessa* – plaice, *Limanda limanda* – dab) that are probably determined by acoustic particle motion. These are probably valid audiograms. Figure 2b plots these same audiograms in terms of particle acceleration. From these data, we can say that specialized species generally hear in a frequency range between 100 Hz and 3 kHz, with thresholds between 50 and 90 dB re: 1 μPa. Furthermore, unspecialized species detect

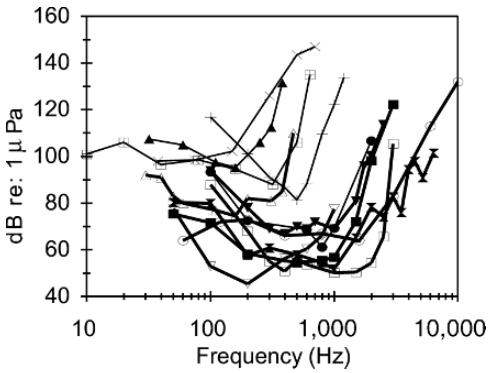


FIGURE 1 Audiograms for selected fish species illustrating specialists (thick lines) and non-specialists (thinner lines). The specialists are *Carassius auratus* (goldfish), *Amiurus nebulosus* (catfish), *Arius felis* (marine catfish), *Astyanax jordani* (Mexican blind cave fish), *Myripristus kuntee* (soldierfish), *Gnathonemus petersii* (elephant nose) and *Gadus morhua* (Atlantic cod). Non-specialists are *Opsanus tau* (oyster toadfish), *Salmo salar* (Atlantic salmon), *Eupomacentrus dorsopunicans* (a damselfish), *Negaprion brevirostris* (lemon shark), and *Equetus acuminatus* (chubbyu). Taken from a review by Fay (1988).

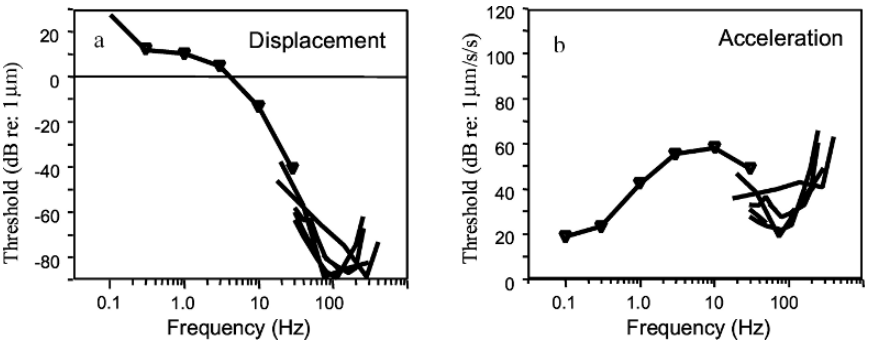


FIGURE 2 Behavioral particle motion audiograms for cod (Buerkle, 1967; Chapman and Hawkins, 1973; Offutt, 1973; Sand and Karlsen, 1986 (▼); plaice (Chapman and Sand, 1974); and dab (Chapman and Sand, 1974). Panels **a** and **b** plot the same thresholds in terms of displacement and acceleration, respectively. Taken from a review by Fay (1988).

acoustic particle motion between 50 and 300 Hz at levels between −60 and −85 dB re: 1 μPa (between 1 and 0.056 nm).

Recently, there have been several frequency threshold curves presented based on averaged evoked brain potentials, rather than on behavioral measures. These curves are uninterpretable for the most part because, in addition to acoustical uncertainties, the relations between behavioral and electrophysiological thresholds are not known. Behavioral thresholds are considered the best definition of sensitivity.

Under normal circumstances, signal detection by fishes is likely masked, meaning that thresholds are determined by the levels of ambient noise within their hearing bandwidths. Therefore, what an animal hears is determined jointly by the ambient noise levels and the audiogram. Masking is relatively well understood using the power spectrum model. In this model, the receiving auditory system is assumed to be composed of band-pass filters. When noise power passing through the filter is equal to signal power passing through the filter, the signal will be masked (i.e., will have to be raised in level in order to be detected). For broadband white noise, the power of masking noise passing through the filter is determined by the receiving filter bandwidth. The wider the filter bandwidth, the more noise power there is to cause masking. For detecting a pure tone signal masked by broadband white noise through a filter of effective width W, the signal will be at threshold (or just masked or just detectable) if the spectrum level (level per hertz) of the noise is $10^* \log(W)$ above the signal level. Masking has been studied in behavioral experiments (e.g., Fay, 1974)

where critical ratio is measured. This is the decibel difference between the level of the signal (pure tone) at threshold and the level per hertz of the broadband, flat spectrum noise that masks the tone. This is essentially a signal-to-noise ratio at threshold. The critical ratio has been measured for goldfish (*Carassius auratus*) (Fay, 1974), marine catfish (*Arius felis*), Afrrican mouthbreeder (*Tilapia macrocephala*), pinfish (*Lagodon rhomboides*) (Tavolga, 1974), and Atlantic cod (*Gadus morhua*) (Chapman and Hawkins, 1973) to be between 13 and 24 dB, between 100 and 1,200 Hz. This means that signals must be this level or higher with respect to the masking noise level to be detected. Once the thresholds for tones are masked by noise, the thresholds are a linear function of noise level, and the detectability of signals falling within the detection bandwidth may be estimated simply by measuring the ambient noise levels.

3 • Vocalization Sounds

Most fish vocalizations are pulsatile, broadband signals that are repeated at different intervals and rates, but some well-known fish sounds are longer duration multiharmonic calls (e.g., toadfishes such as *Opsanus tau* and *Porichthys notatus*). Voluntary sound production is often associated with defense of resources, and reproduction. A few species are known to produce more than one call type (e.g., the mormyrid "grunt", "growl", and "moan") (Crawford, 1997 – Figure 3) that are distinguished by duration, pulse repetition rate, and spectral envelope. Thus, some species may be able to determine the message through

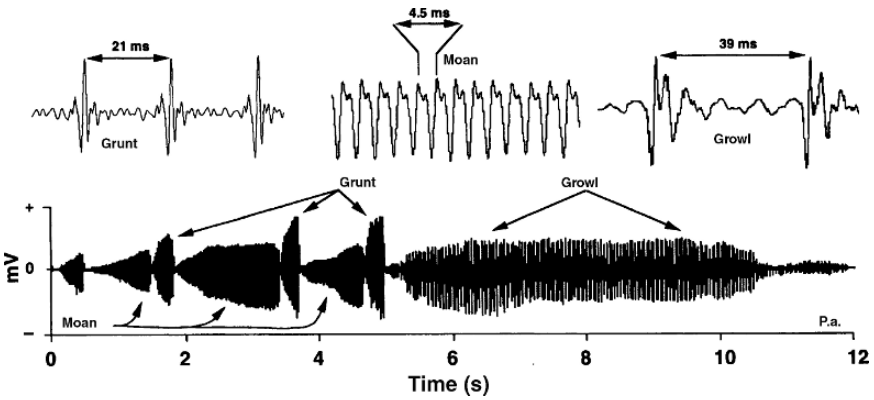


FIGURE 3 Mormyrid vocalizations recorded in the field on two time scales, illustrating the grunt, moan, and growl (from Crawford, 1997).

control of pulse repetition rate and burst duration. Damselfishes can recognize species through a combination of pulse number (duration) and pulse interval.

Table 1 lists the orders having species known to vocalize. After each order, the number of families having vocalizing species is given (Ladich, 2002). There are 53 families containing soniferous species. This number is conservative and likely to grow as research advances; it is exceedingly difficult to visually observe and identify a species while at the same time recording its vocalizations.

Fishes are known to produce sounds using five mechanisms: (1) swim bladder pulsations, (2) stridulation, (3) hydrodynamic movements, (4) substrate, body and tendon vibration, and (5) air release (Kaatz, 2002). Swim bladder and stridulation mechanisms are most common. Swim bladders produce low-frequency sounds of variable duration from pulses to continuous signals, which may be frequency modulated, amplitude modulated, and contain multiharmonic components with low fundamental frequency (no more than several hundred hertz). The best-known examples of swim bladder sounds come from toadfishes. *Opsanus tau* (oyster toadfish) produces the "boatwhistle", a brief, loud sound with a fundamental frequency of 150–250 Hz, that could be described as a "honk", that functions as an advertisement call

TABLE 1 List of the orders within which are found soniferous fishes. The column labeled "#" lists the number of families within the order having species that are soniferous

Order	#
Polypteroformes	1
Osteoglossiformes	2
Cypriniformes	4
Siluriformes	15
Ophidiiformes	1
Gadiformes	3
Batrachoidiformes	1
Beryciformes	1
Zeiformes	1
Gasdterosteiformes	1
Scorpaeniformes	5
Perciformes	17
Tetraodontiformes	1

to attract females during spring mating season (Winn, 1972). *Porichthys notatus* (midshipman) males make a 100-Hz fundamental "hum", a long duration (up to 1/2 hour) advertisement call (Brantley and Bass, 1994). Stridulation sounds consist of a series of broadband acoustic pulses. Both signal sources can produce harmonics and other frequency components that are considerably higher in frequency than the hearing bandwidth. So, the effective vocalizations of fishes are the acoustic recordings as filtered through a particular species' audiogram (frequency–threshold curve – i.e. Figure 1). Vocalization components above 3 kHz are not detected by any species (except by a few related species of shad which themselves do not vocalize).

4 • Locating and Tracking

Using fish vocalizations to assess populations, monitor spawning, locate spawning grounds or other critical habitat, and monitor the general health and abundance of a species are some of the motivations for recording vocalizations. Two of the world's most valuable species, cod and haddock, produce characteristic courtship and spawning sounds (Nordeide and Kjellsby, 1999; Hawkins and Amorim, 2000; Hawkins et al., 2002) that can be used to locate spawning populations. Lobel (2002) has reported on the possibility of monitoring reef health by listening for spawning sounds of *Ostracion meleagris, Dascyllus albisella, Hypoplectrus nigricans*, and *Scarus iserti*. Passive hydrophone census has also recently been reported for the croaker (*Sciaena umbra*) (Bonacito et al., 2002), weakfish (*Cynoscion regalis*) (Connaughton et al., 2002; Luczkovich and Sprague, 2002), red drum (*Sciaenops ocellatus*), spotted sea trout (*Cynoscion nebulosus*), and silver perch (*Bairdiella chrysoura*) (Luczkovich et al., 1999; Luczkovich and Sprague, 2002). All this work must be considered as at the concept demonstration stage at this time, and yearly surveys aimed at systematically tracking populations have not yet been established.

Acoustic telemetry has also been used to study movement patterns and aggregation sites for the dusky grouper (*Epinephelus marginatus*), a threatened and vulnerable species in the Mediterranean Sea (Lembo et al., 2002). These authors describe a system that consists of "cells" covered by wireless hydrophones spaced 200 m apart that monitor the position of fishes tagged with acoustic transmitters. This technology has the advantage that non-vocalizing individuals can be automatically tracked for a long time over a wide range.

5 • Archives and Databases

There have been few systematic attempts to create sound archives or databases of fish sounds. One of the earliest and most comprehensive attempts at a database was that of Fish and Mowbray (1970) who recorded the sounds of 153 species of Western North Atlantic fishes. This body of work is now archived as digital recordings from the original tapes available as a CD from the Graduate School of Oceanography, University of Rhode Island (Rountree et al., 2002). This is the first step toward a proposed National Archive of Fish Sounds that is being assembled at the University of Rhode Island.

References

Bonacito, C., Costantini, M., Picciulin, M., Ferrero, E.A. and Hawkins, A.D. (2002) Passive hydrophone censusof *Sciaena umbra* (Sciaenidae) in the Gulf of Trieste (Northern Adriatic Sea, Italy). Bioacoustics, 12, 292–294.

Brantley, R.K. and Bass, A.H. (1994) Alternative male spawning tactics and acoustic signalling in the plainfin midshipman fish, *Porichthys notatus*. Ethology, 96, 213–232.

Bregman, A.S. (1990) Auditory Scene Analysis: The perceptual organization of sound. MIT Press: Cambridge.

Buerkle, U. (1967) An audiogram for the Atlantic cod, *Gadus morhua*. J. Fish. Res. Bd. Canada, 24, 2309–2319.

Chapman, C.J. and Hawkins A.D. (1973) A field study of hearing in the cod, *Gadus morhua* L. J. Comp. Physiol. 85, 147–167.

Chapman, C.J. and Sand, O. (1974) Field studies of hearing in two species of Flatfish *Pleuronectes platessa*(L.) and *Limanda limanda*(L.) (family pleuronectidae). Comp. Biochem. Physiol. 47A, 371–385.

Connaughton, M.A., Fine, M.L. and Taylor, M.H. (2002). Use of sound for localization of spawning weakfish in Delaware Bay (USA) and effects of fish size, temperature and season on sound parameters. Bioacoustics, 12, 294–296.

Crawford, J.D. (1997) Feature-detecting auditory neurons in the brain of a sound-producing fish. J. Comp. Physiol. A, 180, 439–450.

Fay, R.R. (1974) Masking of tones by noise for the goldfish (*Carassius auratus*). *J. Comp. Physiol. Psychol*. 87, 708–716.

Fay, R.R. (1984) The goldfish ear codes the axis of acoustic particle motion in three dimensions. *Science*, 225, 951–954.

Fay, R.R. (1988) *Hearing in Vertebrates: A Psychophysics Databook*. Hill-Fay Associates, Winnetka, IL,.

Fay, R.R. (1998) Auditory stream segregation in goldfish (*Carassius auratus*). *Hearing Research*, 120, 69–76.

Fay, R.R. (2000) Frequency contrasts underlying auditory stream segregation in goldfish. *J. Assoc. Res. Otolaryngol*. 01, 120–128.

Fay, R.R. and Edds-Walton, P.L. (1997) Directional response properties of saccular afferents of the toadfish, *Opsanus tau*. Hearing Research, 111, 1–21.

Fish, M.P. and Mowbray, W.H. (1970) Sounds of Western North Atlantic Fishes. The Johns Hopkins Press, Baltimore, MD.

Hawkins, A.D. and Amorim, M.C.P. (2000). Spawning sounds of the male haddock, *Melanogrammus aegelfinus*. Environ. Biol. Fishes, 59, 29–41.

Hawkins, A.D., Casaretto, L., Picciulin, M. and Olsen, K. (2002). Locating spawning haddock by means of sound. Bioacoustics, 12, 284–286.

Kaatz, I.M. (2002). Multiple sound-producing mechanisms in teleost fishes and hypotheses regarding their behavioural significance. Bioacoustics, 12, 230–233.

Ladich, F. (1999) Did auditory sensitivity and vocalization evolve independently in otophysan Fishes? Brain Behav. Evol. 53, 288–304.

Ladich, F. (2000) Acoustic communication and the evolution of hearing in fishes. Phil. Trans. Roy. Soc. Lond. B, 355, 1285–1288.

Ladich, F. (2002). Did auditory sensitivity and sound production evolve independently in fishes? Bioacoustics, 12, 176–180.

Lembo, G., Spedicato, M.T., Okland, F., Carbonara, P. Fleming, I.A., McKinley, R.S., Thorstad. E.B., Sisak, M. and Ragonese, S. (2002) A wireless communication system for determining site fidelity of juvenile dusky groupers Epinephelus marginatus (Lowe, 1834) Using coded acoustic transmitters. Hydrobiologia, 483, 249–257.

Lewis, E.R. and Fay, R.R. (2004) Environmental variables and the fundamental nature of hearing. In: Manley, G.A., Popper, A.N. and Fay, R.R. (Eds.) Springer Handbook of Auditory Research: The Evolution of the Vertebrate Auditory System. Springer-Verlag: New York.

Lobel, P.S. (2002). Diversity of fish spawning sounds and the application of passive acoustic monitoring. Bioacoustics, 12, 286–289.

Luczkovich, J.J. and Sprague, M.W. (2002). Using passive acoustics to monitor estuarine fish populations. Bioacoustics, 12, 289–291.

Luczkovich, J.J., Sprague, M.W., Johnson, S.E. and Pullinger, R.C. (1999). Delimiting spawning areas of weakfish, *Cynoscion regalis* (family Sciaenidae), in Pamlico Sound, North Carolina using passive hydroacoustic surveys. Bioacoustics, 10, 143–160.

Mann, D.A. and Lu, Z. (1998) Detection of ultrasonic tones and simulated dolphin echolocation clicks by a teleost fish, the American shad (*Alosa sapidissima*). J. Acoust. Soc. Amer. 104, 562–568.

Mann, D.A., Higgs, D.M., Tavolga, W.N., Popper, A.N. and Souza, M.J. (2001) Ultrasound detection by clupeiform fishes. J. Acoust. Soc. Amer. 109, 3048– 3054.

Nordeide, J.T. and Kjellsby, E. (1999). Sound from spawning cod at their spawning grounds. J. Mar. Sci. 56, 326–332.

Offutt, G.C. (1973) Structures for the detection of acoustic stimuli in the atlantic codfish, *Gadus morhua*. J. Acoust. Soc. Amer. 56, 665–671.

Rountree, R.A., Perkins, P.J., Kenney, R.D. and Hingi, K.R. (2002) Sounds of western North Atlantic fishes – data rescue. Bioacoustics, 12, 242–244.

Sand, O. and Karlsen, H.E. (1986) Detection of infrasound by the Atlantic cod. J. Exp. Biol. 125, 197–204.

Tavolga, W.N. (1974) Signal/Noise ratio and the critical band in fishes. J. Acoust. Soc. Amer. 55, 1323–1333.

Winn, H.E. (1972) Acoustic discrimination by the toadfish with comments on signal systems. In: Winn, H.E. and Olla, B.L. (Eds.) Behavior of Marine Animals. New York: Plenum, pp. 361–385.

102

Amphibian Bioacoustics

Jakob Christensen-Dalsgaard

Institute of Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, e-mail: jcd@biology.sdu.dk

Abbreviations

A	Anterior Thalamic Nucleus
AP	Amphibian Papilla
BP	Basilar Papilla
C	Columella
Cb	Cerebellar Nucleus
Cer	Cerebellum
CF	Characteristic Frequency
DMN	Dorsal Medullary Nucleus
DP	Dorsal Pallium
EC	Extracolumella
ET	Eustachian Tube
IL	Ipsilateral
LA	Lateral Amygdala
LC	Locus Coeruleus
LPv	Lateral Pallium, Ventral Portion
MA	Medial Amygdala
MEC	Middle Ear Cavity
MP	Medial Pallium
MS	Medial Septum
NA	Nucleus Angularis
NI	Isthmal Nucleus
Nis	Secondary Isthmal Nucleus
NL	Nucleus Laminaris
NM	Nucleus Magnocellularis
NVIII	Viiith Nerve
O	Operculum
OT	Optic Tectum
P	Posterior Thalamic Nucleus
PA	Preoptic Area
PT	Posterior Tuberculum
RW	Round Window
SON	Superior Olivary Nucleus
Stv	Ventral Striatum
Tec	Optic Tectum
Teg	Tegmentum
TM	Tympanic Membrane
TS	Torus Semicircularis
VH	Ventral Hypothalamic Nucleus
VM	Ventromedial Thalamic Nucleus

Anuran amphibians (frogs and toads) of most of the 3,500 species that exist today are highly vocal animals. In most frogs, males will spend considerable energy on calling and incur sizeable predation risks and the females' detection and localization of the calls of

conspecific males is often a prerequisite for successful mating. Therefore, acoustic communication is evidently evolutionarily important in the anurans, and their auditory system is probably shaped by the selective pressures associated with production, detection and localization of the communication calls.

Because of the conspicuous, robust calling and sound localization behavior, frogs have been favorite bioacoustic research organisms for almost a century. However, much information is still needed. Even though it is realized that the call may be one of the main defining species characteristics, description and accessible databases of the calls are lacking for most of the species.

1 • Recording and Monitoring Systems

One advantage of anuran bioacoustics is that the calls only rarely contain frequencies above 5 kHz (but see below). Therefore, most mainstream audio-recording equipment is useable. Typically, recording equipment consists of three parts: the microphone (or hydrophone for underwater calls), a signal conditioner (optional, for frequency filtering and amplification) and the sound recorder. See Geissmann (2003) and Gerhardt (1998) for good reviews of bioacoustic recording techniques.

1.1 Microphones

A variety of microphones are available. The two most common types are condenser microphones (in which deflection of the membrane produces a change in capacitance) and electrodynamic microphones (in which deflection of the membrane induces current by moving an electromagnet). The main functional difference is that condenser microphones need a polarization voltage across the membrane. Most high-end microphones used for laboratory studies or recording are condenser microphones. Some of the field microphones are strongly directional, which can be a help in isolating the response of the desired caller. However, for other purposes (general assessment of call activity, investigations of the calling activity of a group of animals), an omnidirectional microphone would be preferable. Laboratory microphones (like the Brüel and Kjaer series) are easy to calibrate, also in the field, using a portable pistonphone delivering 1 Pa, with the advantage that the whole recording setup (microphone, signal conditioner, and sound recorder) is calibrated by recording the 1-Pa stimulus. For field microphones, calibration is more difficult. One method is to compare the response of the microphone to that of a calibrated microphone in the laboratory. Alternatively, the microphone could be calibrated

in the field by recording the sound of a pitch pipe or other kind of reference sound played at a fixed distance. Note, however, that the sound level recorded on the microphone will depend on its immediate surroundings.

Underwater mating calls should be recorded on a hydrophone. Most commercially available hydrophones are piezoceramic and respond to sound pressure. Hydrophones can be calibrated (in air) using a closed sound delivery system (pistonphone). Note, however, that most frogs call in shallow water (typical water depth below 1 m), a medium with special (complex) acoustical properties. Briefly, the medium is high-pass filtering the signal with a cut-off frequency dependent on water depth and bottom material, and the temporal characteristics are distorted due to the multiple reflections from bottom and surface. Therefore, the characteristics of the recorded call may change with the location of the hydrophone (especially with distance, but also with direction if the environment is non-homogenous), and recordings from frogs calling in small tanks may have very different characteristics from those of calls in the natural habitat. See Rogers and Cox (1988) for an introduction to the properties of underwater sound.

1.2 Signal Conditioning (Preamplifier)

It is often useful to filter the microphone signal before recording, for example to reduce wind or traffic noise (high-pass filter), or insect noise, especially from cicadas (low-pass filter). Also, even though the recorder has adjustable recording gain (automatic gain should not be used, see below), it may be advantageous to have fixed gain settings of, for example, 1, 10,100, and 1000 (making it easier to calibrate the recordings). The filter settings should also be variable with a separate low- and high-pass filter, for example off, 10 kHz, 2.5 kHz (low-pass section) and off, 1 Hz, 100 Hz, 200 Hz (high-pass section). Note that if you reduce the sample rate in digital recordings, you will need to low-pass filter the microphone signal to avoid aliasing by under-sampling frequencies above the Nyquist frequency (see below). The steepness of the antialiasing filter determines the minimal useable sampling rate (see Stoddard 1998). Suitable preamplifier units are not readily commercially available but can be built by any electronics workshop.

1.3 The Sound Recorder

Previously, high-quality analog tape recorders such as the Nagra were used for field recordings (Geissmann 2003). However, they have now largely been replaced by high-quality cassette recorders or digital audio tape recorders (DAT) and recently by digital recorders that record directly on hard disks or on

memory cards. The minidisk recorders or MP3-recorders that are used for consumer audio recording are not advisable to use for scientific purposes, because they use non-linear compression algorithms that is incorporating elements of human psychoacoustic processing, making the effects of the compression inaudible (to humans), but nevertheless distorting the recorded signal irreversibly. However, high-definition minidisk recorders (MD-HD) that record on 1 GB disks and where the compression can be switched off have recently been released.

The advantage of a digital over an analog recording is that the recording quality depends less on tape quality and the mechanical performance of the recorder (for example, the fluctuations in tape speed that will generate wow and flutter in analog recordings, Gerhardt 1998). sixteen-bit digital recording also has a larger dynamic range than analog recording, and the signal can be entered directly and noise-free in the computer, either through optical links (DAT) or through card readers, in both cases using the USB interface. In contrast, analog signals on reel-to-reel recorders or cassette recorders must be sampled, and the AD-conversion may introduce additional noise. On the other hand, digital recording depends crucially on avoiding undersampling of the signal. If the signal contains frequency components that exceed one-half of the sampling frequency (the so-called Nyquist frequency) they will be *aliased*, i.e., erroneously represented by frequencies below the Nyquist frequency. For example, if the sample rate is 22 kHz, it is vital that there are no frequency components above 11 kHz (see Stoddard 1998 for an example of aliasing). Thus, for the fidelity of the recording, it is necessary that the frequencies are limited before sampling, i.e. by filtering the analog signal (Stoddard 1998). Digital recorders have built-in antialiasing filters, and it is advisable to check the characteristics (i.e. steepness and f_{3dB} corner) of these filters when selecting a recorder (or a computer sound card). A common source of high frequencies in recordings is clipping caused by overloading of the input amplifier (Gerhardt 1998); because in addition to distortion produced by clipping the clipped signals may also be aliased, it is even more important to avoid overload in digital than in analog recorders by close monitoring of the recorded signal. Many recorders have automatic gain control (AGC), where the recording level is increased automatically when the input signal is low. Because AGC distorts the signal, this option should be avoided (choose recorders where AGC can be switched off), as should the Dolby noise reduction options on analog cassette recorders. On digital recorders, the MP3 or MP2 compression

TABLE 1 Checklist of recording details

Checklist of annotations	Comments
File name (or tape counter), tape/disk ID	
Name of experimenter	
Time and date	Digital recorders usually time stamp the recording – make sure that the clock is set
Location	Geographical location as well as location in habitat and description of habitat
Species name (and other animal identification)	If known – otherwise collect the voucher specimen for later identification and refer to the voucher specimen number or refer to good photographs of the animal
Distance to calling animal	Optimally, the calling animal can be observed
Temperature	Call characteristics are very dependent on temperature
Humidity	Calling activity is usually very dependent on humidity
Microphone/signal conditioner type	
Recorder type	
Recorder gain	
Amplifier gain	
Sampling rate	For digital recording only
Filter settings	(if a signal conditioner is used)
Calibration details	Location of calibration file
Comments	Observations relating to the recording or to the frog's behavior

likewise produces a non-linear distortion of the signal, and these options should be avoided when recording for scientific purposes.

Because the DAT recorders likely will disappear from the market soon, the best choice today probably is digital recorders or high-definition (non-compression) minidisk recorders. A very interesting alternative, however, is recording with a notebook or palmtop computer containing a professional quality sound card, ideally digital signal processor (DSP)-based. Such a system has nearly unlimited potential, for example for on-the-spot analysis and calibration of the recorded signal, and for interactive playback based on features in the recorded call (see below and review in Schwartz 2001).

A main issue in field recordings of frogs is humidity, because most frog habitats are wet, and humidity can approach 100%, leading to condensation in the recording gear. Most vulnerable to condensation are (condenser) microphones and tape recorders. DAT tape recorders, for example, will stop when humidity levels are too high. When microphone sensitivity drops due to humidity, sensitivity may be restored by drying the microphones overnight in bags with dried silica gel. Another practical issue is the power requirement of the recording setup. DAT recorders (e.g. the SONY TCD-D8) run for 2 h on 4 AA cells, a digital recorder (Marantz PMD670) 6 h on 8 AA cells. Many recorders have special, rechargeable batteries, but these can be difficult to recharge in the field, so it is probably optimal that standard disposable batteries can also be used. Most portable computers, however, will need to be recharged. A solution can be to bring a motorcycle or car battery that can be recharged at any gas station, or solar panels.

Finally, the most important aspect of recording is to record, either on the medium or in writing, the details of the recording that is necessary to make it useful, such as details of the equipment (microphone, distance to animal, amplification, filter settings), location, time, weather conditions etc. A suggested checklist (based on Heyer 1994) is summarized in Table 1.

2 • Hearing Performance and Modelling

2.1 Structure of the Ear

Anurans have no external ear structures and external ear canals. The majority of anurans have a tympanic ear with one middle ear bone, the columella that is homologous to the mammalian stapes. The tympanum of anurans, however, is probably derived independently from the tympanum of other vertebrates, so the anuran tympanic ear is not an intermediate stage of the avian or mammalian tympanic ear (Clack and Allin, 2004). The tympanum of anurans is relatively large (diameter typically 0.3–0.8 cm), and in most species relatively undifferentiated from normal skin, although some species have thin, transparent eardrums, in some cases associated with ultrasonic sensitivity (Feng et al. 2006, Shen et al. 2008). In the aquatic pipid frogs (*Xenopus*) the eardrum is a cartilaginous disk that together with an air-filled middle ear provides an efficient reception of underwater sound (Christensen-Dalsgaard and Elepfandt 1995).

Figure 1 shows a schematic diagram of the frog ear [see Lewis and Narins (1999) for a review of amphibian ear structure]. One major entrance for sound to the ear is through the tympanum. However, sound can also reach the internal side of

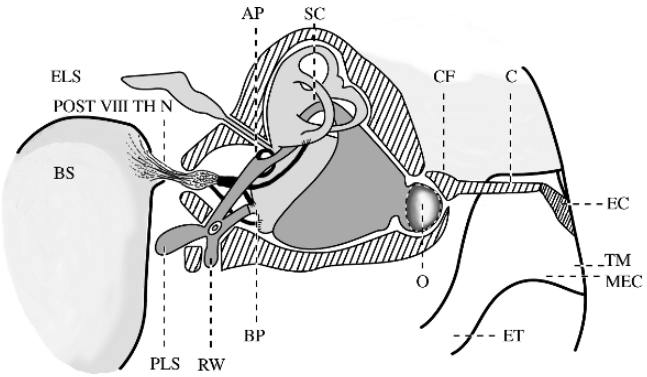


FIGURE 1 A.diagram of the frog ear seen in transverse section at the level of the eardrum. Redrawn and altered from a diagram by Frishkopf and Goldstein (1963). Note that the operculum (O) is drawn in simulated 3-D (extending out of the plane of the figure). TM, tympanic membrane; ET, Eustachian tube; EC, extracolumella; C, columella; CF, columellar footplate; SC, semicircular canals; RW, round window; PLS, perilymphatic sac; AP, amphibian papilla; BP, basilar papilla; ELS, endolymphatic sac; Post VIIIth n, Posterior branch of the VIIIth nerve, BS, brain stem. (© 1963 American Institute of Physics).

the tympanum through the contralateral tympanum and mouth cavity, through the lungs and through the acoustically transparent head region via the mouth cavity and large Eustachian tubes (Narins et al. 1988; Jørgensen 1991). Tympanic vibrations are coupled to the inner ear through the cartilaginous extracolumella and the columella. The columellar footplate transmits vibration through the oval window to the inner ear. The inner ear is contained in a bony otic capsule, and the vibrations exit through the pressure release window, the round window. Amphibians (uniquely) also have another movable element in the oval window, the operculum. The operculum is cartilaginous and connected to the scapula by the opercularis muscle. The function of the operculum is unclear, but it probably transduces vibrations from the ground through the legs and shoulder girdle to the inner ear (Hetherington 1985). In addition, low-frequency sound can reach the inner ear by an unknown, extratympanic pathway that accounts for most of the sensitivity and directionality to sound below 400 Hz (Christensen-Dalsgaard et al. 1997).

Three inner-ear organs respond to air-borne sound. The otolith organ sacculus primarily responds to low-frequency vibrations (BF 40–80 Hz) but can also be stimulated by intense low-frequency sound (Christensen-Dalsgaard and Narins, 1993). The amphibian papilla responds to low-frequency sound (from below

100 Hz up to 1200 Hz) and vibrations (below 500 Hz), whereas the basilar papilla responds to high-frequency sounds (above approximately 1500 Hz). In the amphibian papilla, the sensory hair cells are collected in a large macula and covered by a tectorial membrane. The best frequencies of the sensory hair cells are organized tonotopically, and the amphibian papilla shows non-linearities such as difference tones (Capranica and Moffat, 1980) and active processes (demonstrated by otoacoustic emissions, Palmer and Wilson, 1982). In contrast, the basilar papilla is probably a simple resonator that is tuned to a single best frequency, which is usually the higher frequency component of the advertisement call. Axons from all inner-ear organs are collected in the VIIIth cranial nerve. The main stations in the central auditory pathway consists of the dorsomedullary nucleus (DMN, analog to the mammalian cochlear nucleus), the superior olivary nucleus (SON), torus semicircularis (TS) in the midbrain and the thalamus and preoptic nucleus in the forebrain (Figure 2). Each of these centers is processing the auditory input. In the auditory nerve, auditory information is represented in separate frequency bands, and the temporal modulation of the acoustic signal (amplitude modulation, for example) is encoded in the modulation of the neural discharge (i.e., the fibers are envelope detectors, Feng et al. 1990). Also, the auditory fibers phase lock to low-frequency sound (below 700 Hz, Hillery and Narins, 1984) and thus encode the fine-scale temporal features. As this signal ascends the auditory pathway the modulation of neural discharges is largely degraded (Walkowiak 1980a; review

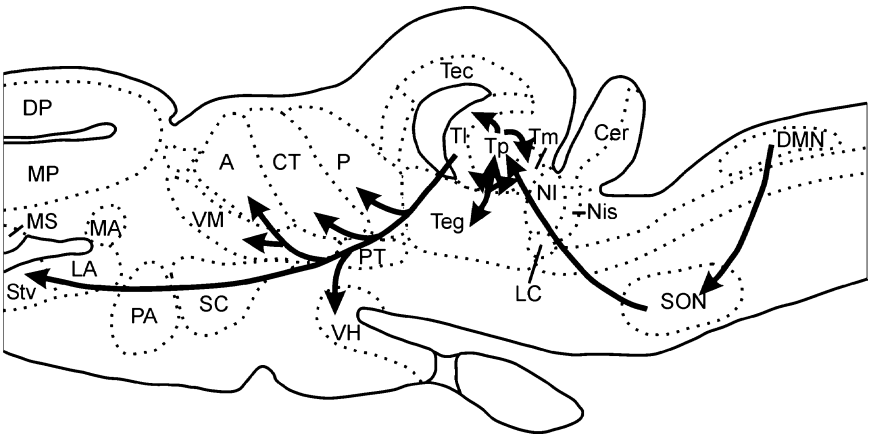


FIGURE 2 Diagram of the ascending auditory connections in the anuran brain, redrawn from Endepols et al. 2000 (© Elsevier Science B.V).

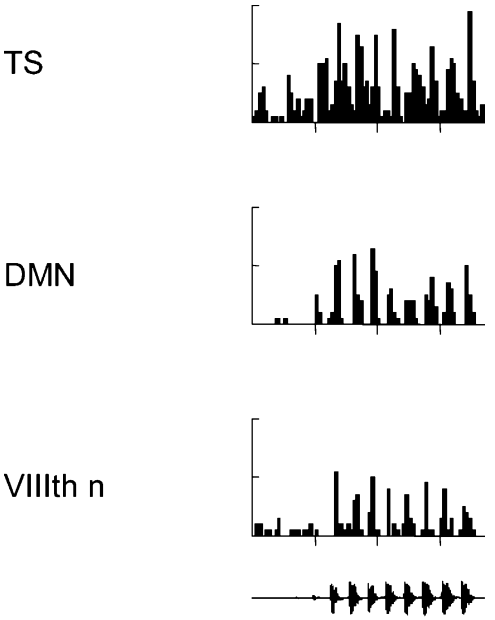


FIGURE 3 PST-histograms showing the response of three cells in the VIIIth nerve, DMN and TS of Xenopus laevis to tympanic stimulation (vibration of the tympanic disk) with a 370-ms chunk of the advertisement call. Taffeta Elliott, unpublished data.

in Feng and Schellart, 1999). This degradation is illustrated in Figure 3, where the response to the mating call of *Xenopus laevis* is shown for a cell in the VIIIth nerve, the DMN and the TS (Elliot, 2007). Instead, important features such as amplitude modulation (Rose et al. 1985), the number of call pulses (Edwards et al. 2002) and probably also rise–fall time are encoded by the discharge rate of specialized cells in the SON (Condon et al. 1991) and TS.

Because anurans are mostly unresponsive to non-communication sounds, it has been supposed that their auditory system is specialized in processing communication sounds. In the pioneering studies of auditory processing in anurans (reviewed in Capranica and Moffat, 1983), it was proposed that the auditory periphery was a matched filter that suppressed non-communication sounds (by non-linearities such as two-tone suppression in the inner ear) and processed call features (such as for example frequency bands and AM-rate) after a hierarchical fashion, leading to hypercomplex mating-call detector cells. It is still debatable, however, to what extent the anuran auditory system is specialized. Auditory systems across the vertebrate phylum show comparable sensitivities suggesting that a general sense of hearing predate the specializations for sound communication (compare, for example, the excellent hearing in many non-vocal fish and lizards), probably because the processing of information about acoustic events has been important throughout the course

of vertebrate evolution. To some extent, however, the frog calls are matched to the auditory sensitivity, but a likely explanation is that there is (sexual) selection for calls that are maximally audible and localizable by the female frogs (the strategy of sensory exploitation, see Wilczynski et al. 2001). Also, the search for hypercomplex mating-call selective cells in the frog CNS has not yielded any results so far, apart from multi-unit responses in the thalamus showing non-linear facilitation (Mudry and Capranica, 1987; Feng et al. 1990). However, it is evident that important acoustical features such as the temporal characteristics described above, behaviorally important frequency bands and sound direction are encoded by specialized cells or small networks and that the complexity increases ascending the auditory pathway. However, it has been shown that the latency for some behavioral responses (antiphonal calling, see below) is too short to involve forebrain structures (Walkowiak 1992), so all processing of this type of calls takes place at or below the TS. Because most cells in the TS have relatively broad specificities, it is most likely that sound stimuli are encoded in the distributed activity of many nerve cells. This hypothesis has not been yet corroborated by simultaneous recordings from many nerve cells in the TS. However, a recent study of *egr-1* early gene expression (correlated with recent synaptic activity) in the TS of *Physalaemus pustulosus* (Hoke et al. 2004) show a highly distributed activity pattern in response to stimulation with different conspecific call types. The activity cannot be explained by any simple acoustic parameters, and the collected activity of four toral subnuclei is the best predictor of the call type.

2.2 Performance of the Anuran Ear

2.2.1 Thresholds

The lowest thresholds of anurans measured by neural and behavioral methods range from 20 to 40 dB SPL (0.2–2 mPa) and the lowest thresholds are usually found from 400 to 900 Hz, i.e. within the amphibian papilla range (see Fay 1988 and examples in Table 2). An example of a neural audiogram based on VIIIth nerve recordings is shown in Figure 4. The lowest thresholds are found in frogs with normal tympanic ears. Earless frogs such as *Bombina bombina* in Table 2 (that have a reduced and non-functional tympanic middle ear, but a largely normal inner ear, Jaslow et al. 1988) have reduced sensitivities, especially at high frequencies. The aquatic frog *Xenopus laevis* has the highest

TABLE 2 Best frequencies and lowest thresholds in six anuran species representing six different families

Species	Best frequency (BF)	Threshold (dB SPL)	Method	Stimulation	Reference
Bombina bombina	300	52	TS, multiunit	Free field	Walkowiak (1980b)
Bufo americanus	370	35	VIIIth nerve	Closed field	Capranica and Moffat (1980)
Eleuthero-dactylus coqui	700	35	VIIIth nerve	Closed field	Zelick and Narins (1985)
Hyla cinerea	900	20	Reflex modifi-cation	Free field	Megela-Simmons et al. (1985)
Limnodynastes dumerilii	700	30	TS	Free field	Loftus-Hills (1973)
Rana temporaria	600	20	VIIIth nerve	Free field	Christensen-Dalsgaard et al. (1998)

sensitivity at 1. Comparable sound pressure sensitivity to that of terrestrial frogs (Elliott et al. 2007).

2.2.2 Frequency Discrimination

The tuning curves of anuran auditory fibers are broad in comparison with tuning curves of other vertebrates, $Q_{10\,dB}$ usually ranging from 1 to 2 (review in Zakon and Wilczynski, 1988). Accordingly, behavioral measures of frequency selectivity (review in Fay and Simmons, 1998) show discrimination limens of 9–33%, about an order of magnitude larger than in mammals and birds at comparable frequencies. Note, however, that these and most other behavioral experiments on frogs use the response to communication sounds, which limits the signals that can be tested (to attractive signals). The frog ear is usually most sensitive to frequencies below 8 kHz, but recently, ultrasonic sensitivity (above 20 kHz) has been reported (Feng et al. 2006).

2.2.3 Temporal

The temporal structure of signals is well represented in the anuran auditory system. In behavioral experiments, male and female frogs (*Hyla versicolor*) will be selective for deviations of less than 25% from the species-specific pulse rate when offered a choice, although calls that deviate by more than 50% from the natural rate can also be attractive (the preference can be offset by other parameters such as call duration and level, Gerhardt

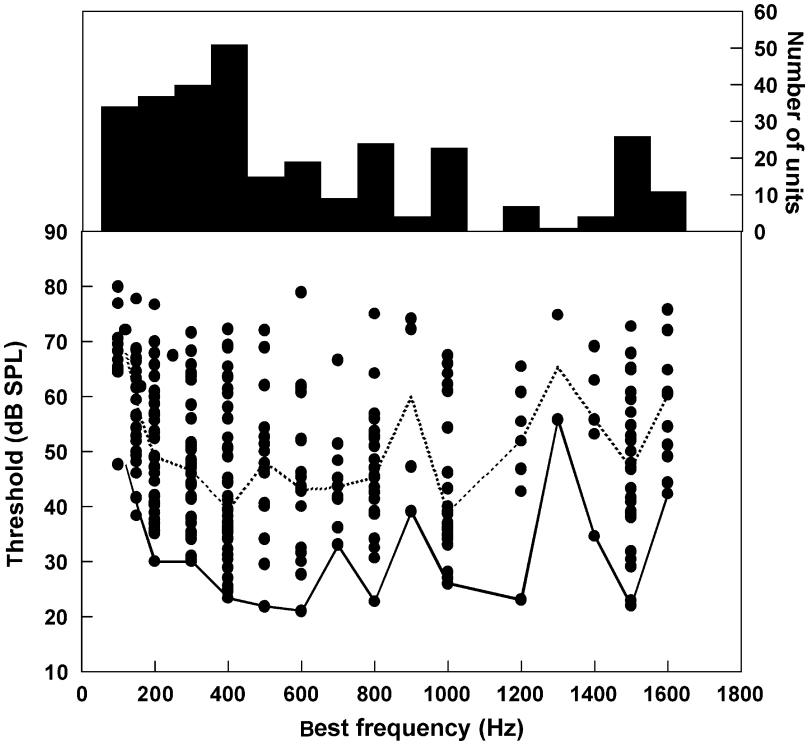


FIGURE 4 Distribution of best frequencies and threshold in single units in the VIIIth nerve in *Rana temporaria* (free-field stimulation). The solid curve connects the units with lowest thresholds. The dashed curve connects the median values at each frequency band. From Christensen-Dalsgaard et al. (1998). © 1998 Elsevier B.V.

and Doherty, 1988; Gerhardt 1988, 1994). Another important temporal parameter in behavioral experiments is the pulse shape. Both pulse rate and shape are important for species identification (Gerhardt and Doherty, 1988; review in Gerhardt 1988).

2.2.4 Directional Hearing

Behavioral studies of directional hearing in frogs have exploited the female phonotaxis toward the mating call. A study of the sound localization behavior in two treefrog species, *Hyla cinerea* and *H. gratiosa* (Feng et al. 1976) showed that gravid females could accurately locate the advertisement call, but were unable to locate the sound source when one eardrum was occluded, indicating that interaural comparison is necessary for a normal phonotactic response (Feng et al. 1976). Localization accuracy has later been quantified in a number of species, and generally,

the accuracy is remarkably similar in all species and approximately 20° in azimuth (see Rheinlaender and Klump, 1988; Klump 1995; Christensen-Dalsgaard 2005). These studies were closed-loop experiments (in which sound is emitted continuously). However, open-loop experiments (in which the sound is switched off after the orienting response) in the barking treefrog, *H. gratiosa* show very similar true angle discrimination accuracy of 21–24° (Klump and Gerhardt, 1989). The localization ability in elevation is poorer (see Christensen-Dalsgaard 2005). The directionality is based on the acoustical coupling of the eardrums. Sound can travel with little attenuation to the internal surface of the eardrum, and the eardrum vibrations depend on the phase difference between internal and external sound, which in turn depend on direction. Measurements of eardrum directionality show ovoidal characteristics with a maximal difference of 6–10 dB between ipsi- and contralateral sound incidence (Figure 5a), and decreasing eardrum vibration amplitude as well as directionality at low and high frequencies. (Jørgensen 1991; Jørgensen et al. 1991; Christensen-Dalsgaard 2005). However, when the directional responses are measured at the level of the auditory nerve fibers, they generally show much larger directionality (15 dB equivalent dB, i.e. calculated from the fibers' rate-level curve) and sensitivity at low frequencies than can be inferred from the eardrum vibrations (Figure 5b and c). Also, the directional pattern is figure-of-eight shaped, with a pronounced frontal null (Jørgensen and Christensen-Dalsgaard, 1997). The origin of this extratympanic sensitivity is presently unknown (discussed in Christensen-Dalsgaard 2005).

Binaural cells are found already at the first auditory nucleus in the CNS, the dorsal medullary nucleus (Feng and Capranica, 1976; Christensen-Dalsgaard and Manley, 2005; Kanneworff and Christensen-Dalsgaard, in prep.). The binaural cells are usually excited by stimuli from one ear and inhibited from the other ear. However, when stimuli are presented monaurally, both ears are excitatory (Kanneworff and Christensen-Dalsgaard in prep.). There is no evidence that time and intensity pathways are separated; rather, the binaural cells respond to specific combinations of ITD and ILD. Also, in response to free-field sound the cells have a sharpened directional response compared with the VIIIth nerve fibers, and this sharpening may be their main function. No clear spatial maps have been found in the midbrain Torus semircicularis, but the directional response is sharpened by lateral inhibition (Zhang et al. 1999) and the ensuing difference in activity between the left and right TS may suffice in steering the frog in the right direction.

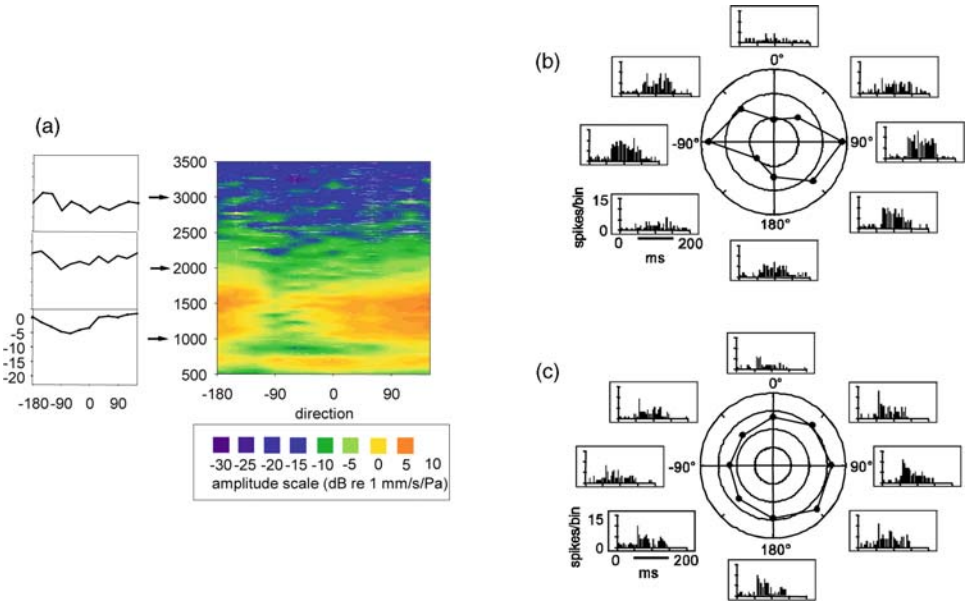


FIGURE 5 (a) Cylinder Surface Plot of eardrum directionality in *Rana temporaria*. The normalized velocities (colour scale, in dB re 1 mm/s*Pa) are plotted as a function of direction (*x*-axis, contralateral angles on the left and ipsilateral angles on the right) and frequency (*y*-axis). The right row of figures show line plots of velocity (*y*-axis, dB re 1 mm/s*Pa) as a function of direction (*x*-axis) at three frequencies (1000, 2000, and 3000 Hz, arrows) corresponding to three horizontal lines in the cylinder surface plot. (b,c) Response of VIIIth nerve fibers in *Rana temporaria* to directional stimulation. (b) shows the response of a low-frequency fiber stimulated at BF (300 Hz). Note the "figure-of-eight" response with low sensitivity at frontal directions. (c) is the response of a BP fiber stimulated at its BF (1500 Hz). Here, the response is ovoidal. (a) from Christensen-Dalsgaard and Manley 2005, © 2005 Company of Biologists ltd., (b) and (c) from Jørgensen and Christensen-Dalsgaard 1997. © 1997 Springer-Verlag.

2.3 Modelling

The anuran periphery has been the subject of several modelling studies based on electrical analog circuits. The response of the eardrum in a closed coupler could be modelled accurately (Palmer and Pinder 1984; Aertsen et al. 1986), but the models were not able to produce a good fit to experimental free-field data (Jørgensen 1991), either predicting largest directionality at very low frequencies and small directionality in the frequency range of maximal sensitivity (Aertsen et al. 1986) or a figure-of-eight characteristic at low frequencies (Palmer and Pinder 1984). Because the response of the isolated eardrum in a coupler can be modelled the reason for the discrepancy does not lie in the

estimation of parameters but likely is that the acoustics of the frog auditory system are very complicated. For example, it is suggested that the entire head region is transparent to sound (Aertsen et al. 1986) and it has been shown that another major sound input is through the lungs, an area that covers most of the dorsum (Jørgensen et al. 1991; Narins et al. 1988; Christensen-Dalsgaard 2005). These inputs probably cannot be modelled by simple lumped elements.

The peripheral auditory system (van Stokkum and Gielen, 1989; Simmons et al. 1996) and the auditory midbrain (van Stokkum 1990) have been modelled. The peripheral models are very similar "sandwich" models consisting of a band pass filter, a non-linear rectifier and a low-pass filter (van Stokkum and Gielen include a spike generator after the low-pass filter). With appropriate adjustment of the parameters the model can fit the responses of AP, especially with regard to periodicity extraction (Simmons et al. 1996), but it is evident that the frog ear contains more than one compressive non-linearity: for example, two-tone suppression and active processes are not included in the model. The midbrain model by van Stokkum (1990) assumes convergence of the peripheral input on central neurons. Again, the model is very general and do not take the complicated connections of the central neurons into account.

A simulation experiment on visual orientation in salamanders showed that the localization of moving prey can be simulated quite accurately by a relatively small network consisting of 300 neurons, (100 optic tectum neurons with "coarse coding" (large receptive fields), 100 interneurons, 100 motoneurons, and four muscles) and incorporating neuroanatomical and -physiological features of the salamander visual brain (Eurich et al. 1995). The model does not incorporate a motorneuron map. Rather, all neurons participate in coding of prey location, and the distributed coding of the tectal neurons is transformed directly into a distributed activation of the muscles. It would be interesting to apply a similar kind of model to the anuran auditory system.

3 • Characteristics of Anuran Calls

3.1 Call Generation Mechanism

Frog calls are usually stereotyped and consist of few different call types. In almost all species, predominantly the males are calling, although females often have a release call uttered when a non-receptive female is clasped by a male, and a female response call has been described in few species (Tobias et al. 1998). Two different call mechanisms have been described. In most terrestrial frogs, sound is produced when air from the lungs is

forced through the larynx and the vocal cords vibrate (passively) in the air stream. This calling by expiration is the most common form, but a few species (*Bombina)* call by inspiration (producing sound when air is forced from the buccal cavity to the lungs), and some species call by expiration and inspiration (*Discoglossus*). The vocal cords extend from the inner surface of the bowl-shaped arytenoid cartilages. The muscles of the larynx are attached to the external surface of the arytenoid cartilages (Schneider 1988). In other words, the tension of the vocal cords is not under muscular control, and the pitch of the sound they emit is chiefly changed by increasing or decreasing the air flow from the lungs. Furthermore, the fundamental frequency emitted by the vocal cords is correlated with their mass, and therefore usually with the size of the animal. However, the arytenoid cartilages are controlled by the laryngeal muscles, not only starting and stopping the sound emission, but also imposing amplitude modulation on the sound generated by the vocal cords (Martin and Gans, 1972). It should be noted, however, that a detailed analysis of the calling mechanism has only been undertaken in a few species of anurans, and some of the species that have very sophisticated calling behavior showing a large diversity of call types (Narins et al. 2000; Christensen-Dalsgaard et al. 2002), ultrasonic harmonics (Narins et al. 2004, Figure 6d) or aharmonic calls (Figure 6c) might warrant further investigations of the mechanism. Aquatic pipid frogs such as *Xenopus* produce sound by another mechanism than the terrestrial anurans (Yager 1992). These frogs have no vocal cords. Instead, when muscular contractions force two disk-like extensions of the arytenoid cartilages apart they generate a click. Probably, this calling mechanism is an adaptation to underwater calling without moving air through the larynx.

The most prominent accessory structure in sound production is the vocal sac. The vocal sac is not found in all anuran species that call, but in those that do, extension of the vocal sac usually is correlated with an increase in the output of sound power, although the vocal sac is probably not a resonator (Rand and Dudley, 1993). Furthermore, in *R. catesbeiana* it has been shown that a major part of the sound energy is emitted through the eardrum (Purgue 1997).

3.2 Neural and Hormonal Control of Calling

The immediate control of calling is through the innervation of the laryngeal and respiratory muscles by motoneurons in the vagal (X) and the hypoglossal (XII) cranial nerves. The laryngeal muscles are innervated by motoneurons in the vagal nerve. The motor nuclei are controlled by the pre-V or pretrigeminal nucleus.

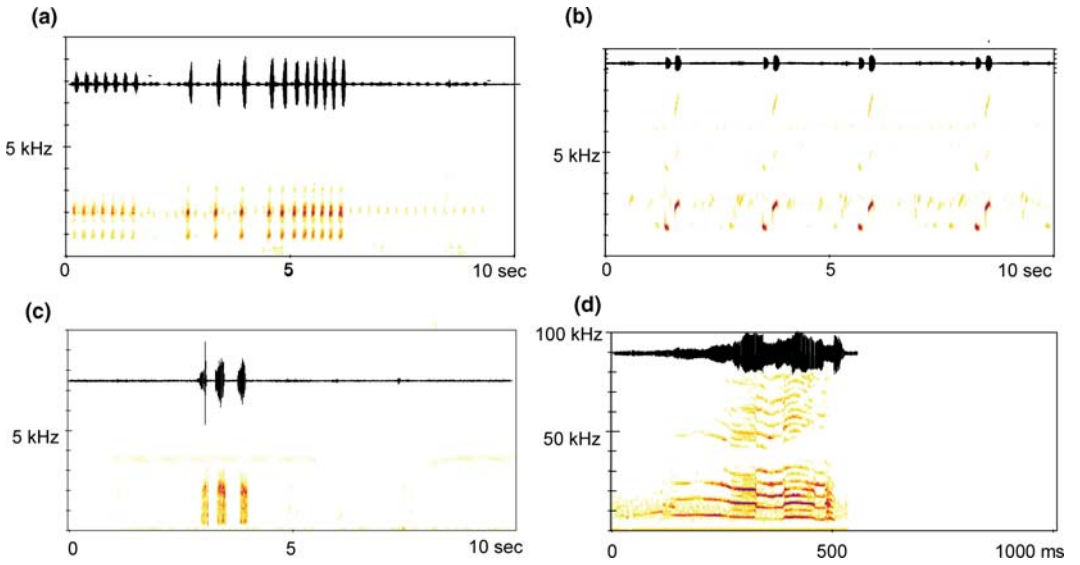


FIGURE 6 Calls of four frog species. Both oscillograms (inserts) and spectrograms (256-point FFT, Hann window) are shown. (a) *Hyla arborea*, Denmark, (b) *Eleutherodactylus coqui*, Puerto Rico, (c) *Rana lini*, Thailand, (d) *Amolops tormotus*, China (courtesy of Albert Feng). Note the change of scale in D and the ultrasonic components in the call, described by Narins et al. (2004).

The direct hormonal control of calling is mediated by androgens such as testosterone in most species, although additional hormones such as the neuropeptide arginine vasotocin (AVT) may be required in some species. In many species, AVT injections can induce calling (review in Wilczynski and Chu, 2001; Yamaguchi and Kelley, 2003). It has been proposed that AVT may induce fast modifications of calling behavior (within minutes), whereas androgens effect changes in behavior within days or weeks (Wilczynski and Chu, 2001). The targets of androgens and AVT are the laryngeal motor nucleus and the pretrigeminal nucleus (Yamaguchi and Kelley, 2003). Release of androgens is controlled by gonadotrophine releasing hormone from the forebrain, especially in the preoptic area, and cells in this area are responsive to acoustic stimuli (review in Wilczynski et al. 1993). Environmental stimuli that can induce calling behaviour are increased humidity, changes in temperature, changes in water composition (in aquatic frogs) and changes in photoperiod.

3.3 Examples of Call Types

Figure 6 shows examples of anuran calls (see also examples in Gerhardt 1998). Two call types are common: calls consisting of one or several pulses of a tonal carrier shaped by a gamma-tone

envelope (i.e. with a polynomial rise and an exponential decay) (Figure 6a) and a whistle-like call type (Figure 6b). Some species, however, produce calls that are much less stereotyped or with an almost bird-like characteristic. Also, some species produce many different call types (a record of 28 in the Madagascan *Boophis*, Narins et al. 2000) that may be used in different behavioral contexts (Christensen-Dalsgaard et al. 2002). Also, it has been shown recently that the Chinese frog *Amolops tormotus* has prominent ultrasonic components in its call (Figure 6d). As described above it has been shown that frogs can hear 5d. The function of the ultrasonic components (Feng et al. 2006), but their function is presently unknown (Narins et al. 2004).

3.4 Calling Behavior

Usually, male anurans call frequently during the mating season, with highest activity from dusk till midnight depending on temperature and humidity. Some species, however, like the aposematically colored poison arrow frogs (Dendrobatidae) will call during the day. The calls may either be alternated, minimizing the overlap between calls of individuals (antiphonal calling) or be more or less synchronized (chorusing) (reviews in Wells 1977; Gerhardt 1994; Zelick et al. 1999). One function of the call is obviously mate attraction, and it has been shown that females of many species are attracted to the call and highly selective to features (for example pulse rate) that are distinctive for the species (Gerhardt 1994), indicative of the caller's fitness (Welch et al. 1998) or that make the call conspicuous and easy to localize (Wilczynski et al. 2001). In other species, however, the call functions primarily in intersexual selection, i.e. male – male interactions (territorial behavior, for example), and the females will be relatively unselective (Arak 1983). Also, some species may have specialized aggression calls used in male – male interactions (Grafe 1995), and these interactions may even be an acoustic display with many different call types used only in male aggressive behavior (Christensen-Dalsgaard et al. 2002). The usual outcome of male interactions is the retreat of one of the males. Alternatively, a "losing" male may choose a "satellite" strategy by staying silently in the vicinity of a calling male and attempting to mate with females attracted to the caller.

The costs of calling are partly the increase in metabolism and partly the increased predation risk incurred by the caller. The maximum oxygen consumption of calling frogs increase by more than ten times over the resting metabolic rate in laboratory measurements (review in Wells 2001; note that all measurements so far have been undertaken on small frog species with high

TABLE 3 Archives and databases of anuran calls signals

Name	Location	URL	Number of anuran recordings	Comments
Tierstimmenarchiv, Humboldt University	Berlin, FRG	http://www.tierstimmenarchiv.de/	Approx. 20	Database searchable from the internet. Small sample of African species
Sound Archive, British Library	London, UK	http://www.bl.uk/collections/sound-archive/wildex.html#amphibians	Approx. 800	Large collection of species from North America, South Africa, Europe, and Australia and collections from Sarawak, Venezuela. Catalogue searchable from the internet
Borror Library	Ohio State University, USA	http://blb.biosci.ohio-state.edu/	74 species	Only North and South American species. Catalogue searchable from the internet
Macaulay Library	Cornell University, USA	http://www.birds.cornell.edu/MacaulayLibrary/Search/	Approx. 100	Most recordings from North and South America. Catalogue searchable from the internet

call rates), and prolonged calling leads to depletion of glycogen stores and even to a reduction of body mass (Murphy 1994), so males may be energy-limited. The more indirect cost of calling by increased predation risk is more difficult to assess. However, it has been shown that frog-eating bats will detect and capture calling frogs (Tuttle and Ryan, 1982), and many other nocturnal mammals and birds are likely to be attracted to calling males.

4 • Stimulus Generation – Synthesized Calls

Synthesized calls offer the possibility of manipulating call rate and carrier frequency independently. Anuran calls are generally easy to synthesize, because they often consist of a tonal carrier (with few harmonics)and a relatively smooth envelope and are emitted with a constant rate (Gerhardt 1998). Previous studies have used analog call synthesizers (Narins and Capranica, 1978), but digital processing is probably a better choice today owing to the unlimited flexibility. The most straightforward approach would be to synthesize the envelope and the carrier separately and multiply them. See Gerhardt (1992) for an introduction to playback using synthetized signals.

5 • Archives and Databases of Signals

Compared to the collections of bird calls, anuran calls are scantily represented in the major sound archives, with entire regions – such as Southeast Asia – hardly covered at all. The major archives of animal sound are listed in Table 3. Presently, most of the archive material is analog recordings and not available from the internet, but the recordings will hopefully be digitized in the future. In addition to these established archives many sound recordings can be found by searches on the internet. For this material, however, the documentation is rarely sufficient for scientific purposes.

References

Aertsen AMHJ, Vlaming MSMG, Eggermont JJ, Johannesma PIM (1986) Directional hearing in the grassfrog (*Rana temporaria* L.). II. Acoustics and modelling of the auditory periphery. Hear Res 21: 17–40

Arak A (1983) Sexual selection by male-male competition in natterjack toad choruses. Nature 306: 261–262

Capranica RR, Moffat AJM (1980) Nonlinear properties of the peripheral auditory system of anurans. In: Popper AN, Fay RR (eds) Comparative studies of hearing in vertebrates. New York: Springer-Verlag, pp. 139–165

Capranica RR, Moffat AJM (1983) Neurobiological correlates of sound communication in anurans. In: Ewert J-P, Capranica RR, Ingle DJ (eds) Advances in vertebrate neuroethology. New York: Plenum Press, pp. 701–730

Christensen-Dalsgaard J (2005) Directional hearing in non-mammalian tetrapods. In: Popper AN, Fay RR (eds) Sound source localization, New York: Springer-Verlag, pp. 67–123

Christensen-Dalsgaard J, Elepfandt A (1995) Biophysics of underwater hearing in the clawed frog, Xenopus laevis. J Comp Physiol A 176: 317–324

Christensen-Dalsgaard J, Jørgensen MB, Kanneworff M (1998). Basic response characteristics of auditory nerve fibers in the grassfrog (*Rana temporaria*). Hear Res 119: 155–163.

Christensen-Dalsgaard J, Kanneworff M, Jørgensen MB (1997) Extratympanic sound sensitivity of frog auditory fibers. In: Lewis ER, Long GR, Lyon RF, Narins PM, Steele CR, Hecht-Poinar E (eds) Diversity in auditory mechanics, World Scientific, Singapore New Jersey London Hong Kong, pp. 64–68

Christensen-Dalsgaard J, Ludwig TA, Narins PM (2002) Call diversity in an old world treefrog: Level dependency and latency of acoustic responses. Bioacoustics 13: 21–35

Christensen-Dalsgaard J, Manley GA (2005) Directionality of the lizard ear. J Exp Biol 208:1209–1217

Christensen-Dalsgaard J, Narins PM (1993) Sound and vibration sensitivity in the frogs *Leptodactylus albilabris* and *Rana pipiens pipiens*. J Comp Physiol A 172: 653–662

Clack, JA, Allin E (2004) The evolution of single- and multiple- ossicle ears in fishes and tetrapods. In: Manley GA, Popper AN, Fay RR (eds) Evolution of the Vertebrate Auditory System. New York, Springer-Verlag, pp. 128–163

Condon C, Huang S-H, Feng AS (1991) Processing of behaviorally relevant temporal parameters of acoustic stimuli by single neurons in the superior olivary nucleus of the leopard frog. J Comp Physiol A 168: 709–725

Edwards CJ, Alder TB, Rose GJ (2002) Auditory midbrain neurons that count. Nat Neurosci 5: 934–936

Elliott TM (2007) The neural basis of click rate coding in the auditory system. Ph.d. thesis. New York: Columbia University

Elliott TM, Christensen-Dalsgaard J, Kelley DB (2007) Tone and call responses in the auditory nerve and dorsal medullary nucleus of Xenopus laevis. J Comp Physiol A 193: 1243–1257

Endepols H, Walkowiak W, Luksch H (2000) Chemoarchitecture of the anuran auditory midbrain. Brain Res Rev 33: 179–198

Eurich C, Roth G, Schwegler H, Wiggers W (1995) Simulander: A neural network model for the orientation movement of salamanders. J Comp Physiol A 176: 379–389

Fay RR (1988) Hearing in vertebrates: A psychophysics databook. Winnetka, IL: Hill-Fay Associates

Fay RR, Simmons AM (1998) The sense of hearing in fish and amphibians. In: Fay RR and Popper AN (eds) Comparative Hearing: Fish and Amphibians. New York: Springer-Verlag, pp.269–318

Feng AS, Capranica RR (1976) Sound localization in anurans. I. Evidence of binaural interaction in the dorsal medullary nucleus of the bullfrog (*Rana catesbeiana*). J Neurophysiol 39: 871–881

Feng AS, Schellart NAM (1999) Central auditory processing in fish and amphibians. In: Fay RR, Popper AN (eds.) Comparative hearing: Fish and amphibians. New York: Springer-Verlag, pp. 218–268

Feng AS, Gerhardt HC, Capranica RR (1976) Sound localization behavior of the green treefrog (*Hyla cinerea*) and the barking treefrog (*H. gratiosa*). J Comp Physiol 107: 241–252

Feng AS, Hall JC, Gooler DM (1990) Neural basis of sound pattern recognition in anurans. Prog Neurobiol 34: 313–329

Feng AS, Narins PM, Xu C-H, Lin W-Y, Yu Z-L, Qiu Q, Xu Z-M, Shen J-X (2006) Ultrasonic communication in frogs. Nature 440: 333–336

Frishkopf LS, Goldstein MH (1963) Responses to acoustic stimuli from single units in the eighth nerve of the bullfrog. J Acoust Soc Am 35:1219–1228

Geissmann T (2003) Tape-recording primate vocalisations. In: Setchell, JM, Curtis, DJ (eds) Field and Laboratory Methods in Primatology: A Practical Guide. Cambridge: Cambridge University Press, pp. 228–238

Gerhardt HC (1988) Acoustic properties used in call recognition by frogs and toads. In: In Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W (eds) The evolution of the amphibian auditory system. New York Chichester Brisbane: Wiley, pp. 537–558

Gerhardt HC (1992) Conducting playback experiments and interpreting their results. In: McGregor PK (ed) Playback and studied of animal communication. New York: Plenum Press, pp. 59–77

Gerhardt HC (1994) The evolution of vocalization in frogs and toads. Ann Rev Ecol Syst 25: 293–324

Gerhardt HC (1998) Acoustic signals of animals: Recording, field measurement, analysis and description. In: Hopp SL, Owren MJ, Evans CS (eds) Animal acoustic communication. New York: Springer-Verlag, pp. 1–25

Gerhardt HC, Doherty JA (1988) Acoustic communication in the gray treefrog, *Hyla versicolor*: evolutionary and neurobiological implications. J Comp Physiol A 162: 261–278

Grafe TU (1995) Graded aggressive calls in the African reed frog, *Hyperolius marmoratus* (Hyperoliidae). Ethology 101: 67–81

Hetherington TE (1985) Role of the opercularis muscle in seismic sensitivity in the bullfrog, *Rana catesbeiana*. J Exp Zool 235: 27–34

Heyer WR (1994) Recording frog calls. In: Heyer WR, Donelly MA, McDiarmid RW, Hayek LC, Foster MS (eds) Measuring and monitoring biological diversity. Standard methods for amphibians. Washington: Smithsonian Institution Press, pp. 285–287

Hillery CM, Narins PM (1984) Neurophysiological evidence for a traveling wave in the amphibian inner ear. Science 225: 1037–1039

Hoke KL, Burmeister SS, Fernald RD, Rand AS, Ryan MJ, Wilczynski W (2004) Functional mapping of the auditory midbrain during mate call reception. J Neurosci 24: 11264–11272

Jaslow AP, Hetherington TE, Lombard RE (1988) Structure and function of the amphibian middle ear. In: Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W (eds) The Evolution of the Amphibian Auditory System. New York: John Wiley, pp. 69–91

Jørgensen MB (1991) Comparative studies of the biophysics of directional hearing in anurans. J Comp Physiol A 169: 591–598

Jørgensen MB, Christensen-Dalsgaard J (1997) Directionality of auditory nerve fiber responses to pure tone stimuli in the grassfrog, *Rana temporaria*. I. Spike rate responses. J Comp Physiol A 180: 493–502

Jørgensen MB, Schmitz B, Christensen-Dalsgaard J (1991) Biophysics of directional hearing in the frog *Eleutherodactylus coqui*. J Comp Physiol A 168: 223–232

Klump GM (1995) Studying sound localization in frogs with behavioral methods. In Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psycho-acoustics. Basel: Birkhäuser Verlag, pp. 221–233

Klump GM, Gerhardt HC (1989) Sound localization in the barking treefrog. Naturwissenschaften 76: 35–37

Lewis ER, Narins PM (1999) The acoustic periphery of amphibians: Anatomy and physiology. In: Fay RR, Popper AN (eds.) Comparative Hearing: Fish and Amphibians. New York: Springer-Verlag, pp. 101–154

Loftus-Hills JJ (1973) Comparative aspects of auditory function in Australian anurans. Aust J Zool 21: 353–367

Martin WF, Gans C (1972) Muscular control of the vocal tract during release signalling in the toad *Bufo valliceps*. J Morph 137: 1–28

Megela-Simmons A, Moss CF, Daniel KM (1985) Behavioral audiograms of the bullfrog (*Rana catesbeiana*) and the green treefrog (*Hyla cinerea*). J Acoust Soc Am 78: 1236–1244

Mudry KM, Capranica RR (1987) Correlation between auditory evoked reponses in the thalamus and species-specific call characteristics. I. *Rana catesbeiana*. J Comp Physiol 160: 477–489

Murphy CG (1994) Chorus tenure of male barking treefrogs, *Hyla gratiosa*. Anim Behav 48: 763–777

Narins PM, Capranica RR (1978) Communicative significance of the two-note call of the treefrog *Eleutherodactylus coqui*. J Comp Physiol 127: 1–9

Narins PM, Ehret G, Tautz J (1988) Accessory pathway for sound transfer in a neotropical frog. Proc Natl Acad Sci USA 85: 1255–1265

Narins PM, Feng AS, Lin W, Schnitzler H-U, Denzinger A, Suthers RA, Xu C (2004) Old World frog and bird vocalizations contain prominent ultrasonic harmonics. J Acoust Soc Am 115:910–913

Narins PM, Lewis ER, McClelland BE (2000) Hyperextended call note repertoire of the endemic Madagascar treefrog *Boophis madagascariensis* (Rhacophoridae) J Zool Lond 250: 283–298

Palmer AR, Pinder AC (1984) The directionality of the frog ear described by a mechanical model. J Theor Biol 110: 205–215

Palmer AR, Wilson JP (1982) Spontaneous and evoked acoustic emissions in the frog *Rana esculenta*. J Physiol Lond 324: 66P

Purgue AP (1997) Tympanic sound radiation in the bullfrog *Rana catesbeiana*. J Comp Physiol A 181: 438–445

Rand AS, Dudley R (1993) Frogs in helium: the anuran vocal sac is not a cavity resonator. Physiol Zool 66: 793–806

Rheinlaender J, Klump G (1988) Behavioral aspects of sound localization. In: Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W (eds) The Evolution of the Amphibian Auditory System. New York: John Wiley, pp. 297–305

Rogers PH, Cox M (1988) Underwater sound as a biological stimulus. In: Atema J, Fay RR, Popper AN, Tavolga WN (eds) Sensory biology of aquatic animals. New York: Springer-Verlag, pp. 131–149

Rose GJ, Brenowitz EA, Capranica RR (1985) Species specificity and temperature dependency of temporal processing by the auditory midbrain of two species of treefrogs. J Comp Physiol A 157: 763–769

Schneider H (1988) Peripheral and central mechanisms of vocalization. In Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W (eds) The evolution of the amphibian auditory system. Wiley, New York Chichester Brisbane, pp. 537–558

Schwartz JJ (2001) Call monitoring and interactive playback systems. In: Ryan MJ (ed) Anuran communication. Washington: Smithsonian Institution Press, pp. 183–204

Shen JX, Feng AS, Xu ZM, Yu ZL, Arch VS, Yu XJ, Narins PM (2008) Ultrasonic frogs show hyperacute phonotaxis to the female's courtship calls. Nature, In Press

Simmons AM, Shen Y, Sanderson MI (1996) Neural and computational basis for periodicity extraction in frog peripheral auditory system. Aud Neurosci 2: 109–133

Stoddard PK (1998) Application of filters in bioacoustics. In: Hopp SL, Owren MJ, Evans CS (eds) Animal acoustic communication. New York: Springer-Verlag, pp. 105–127

van Stokkum IHM (1990) Modelling the response of auditory midbrain neurons in the grassfrog to temorally structured monaural stimuli. Hear Res 43: 231–250

van Stokkum IHM, Gielen CCAM (1989) A model for the peripheral auditory system of the grassfrog. Hear Res 41: 71–86

Tobias ML, Viswanathan S, Kelley DB (1998). Rapping, a female receptive call, initiates male/female duets in the South African clawed frog. Proc Natl Acad Sci 95:1870–1875

Tuttle MD, Ryan MJ (1982) The role of synchronized calling, ambient light, and ambient noise, in anti-bat-predator behavior of a treefrog. Behav Ecol Sociobiol 11: 125–131

Walkowiak W (1980a) The coding of auditory signals in the torus semicircularis of the fire-bellied toad and the grass frog: Responses to simple stimuli and to conspecific calls. J Comp Physiol 138: 131–148

Walkowiak W (1980b) Sensitivity, range and temperature dependence of hearing in the grass frog and fire-bellied toad. Behav Proc 5: 363–373

Walkowiak W (1992) Acoustic communication in the fire-bellied toad: an integrative neurobiological approach. Ethol Ecol Evol 4: 63–74

Welch AM, Semlitsch RD, Gerhardt HC (1998) Call duration as an indicator of genetic quality in male gray tree frogs. Science 280: 1928–1930

Wells KD (1977) The social behavior of anuran amphibians. Anim Behav 25: 666–693

Wells KD (2001) The energetics of calling in frogs. In: Ryan MJ (ed) Anuran communication. Washington: Smithsonian Institution Press, pp. 45–60

Wilczynski W, Chu J (2001) Acoustic communication, endocrine control, and the neurochemical systems of the brain. In: Ryan MJ (ed) Anuran Communication. Washington: Smithsonian Institution Press, pp. 23–35

Wilczynski W, Allison JD, Marler CA (1993) Sensory pathways linking social and environmental cues to endocrine control regions of the amphibian forebrain. Brain Behav Evol, 42: 252-264

Wilczynski W, Rand AS, Ryan MJ (2001) Evolution of calls and auditory tuning in the *Physalaemus pustulosus* species group. Brain Behav Evol 58: 137–151

Yager DD (1992) A unique sound production mechanism in the pipid anuran *Xenopus borealis*. Zool J Linn Soc 104: 351–375

Yamaguchi A, Kelley DB (2003) Hormonal mechanisms in acoustic communication. In: Simmons AM, Popper AN, Fay RR (eds) Acoustic communication. New York: Springer-Verlag, pp. 275–323

Zakon HH, Wilczynski W (1988) The physiology of the anuran eighth nerve. In: Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W, (eds) The evolution of the amphibian auditory system. New York: John Wiley & Sons, pp. 125–155

Zelick R, Mann DA, Popper AN (1999) Acoustic communication in fishes and frogs. In: Fay RR, Popper AN (eds): Comparative Hearing: Fish and amphibians. New York: Springer-Verlag, pp. 363–411

Zelick R, Narins PM (1985) Temporary threshold shift, adaptation and recovery characteristics of frog auditory nerve fibers. Hear Res 17: 161–176

Zhang H, Xu J, Feng AS (1999) Effects of GABA-mediated inhibition on direction-dependent frequency tuning in the frog inferior colliculus. J Comp Physiol A 184: 85–98

103

Avian Bioacoustics

Kazuo Okanoya

Laboratory for Biolinguistics, Brain Science Institute, Riken,
2-1 Hirosawa, Saitama 351-0198, Japan, e-mail: okanoya@brain.riken.jp

1 • Recording and Monitoring Systems

Modern avian bioacoustics enjoy multitude of recording equipment and analyses software. Because most avian sounds fall within the range of normal human hearing, equipments from consumer electronics can easily be applied for avian bioacoustics research. Similarly, most analyses software for human speech will also cover the range of avian sounds that rarely exceeds 20 kHz. Nevertheless, cautions should be required since some of these equipment and software use special compression or enhancement techniques that are optimized for human speech recording. When using consumer electronics, it is advertised to find a "linear" or "straight" pulse code modulation (PCM) setting to avoid signal conditioning optimized for human speech.

Historically, avian bioacoustics advanced rapidly when the sound spectrograph of Key Elemetrics became available for scientific research (Thorpe 1961). This equipment was typically called as "sonogram". Early equipment utilized analog filter banks for frequency analyses and electro-sensitive papers for recording that turned black when electric discharges were applied. At mid-1980s specialized sonograms used digital array processors and at mid-1990s personal computers were fast enough to allow software implementation of sonogram (Specht). Today, most avian bioacousticians use software tools for recording, monitoring, and analyzing avian sounds.

1.1 Recording Equipment

Physical and electric considerations when making field or laboratory recordings of avian sound have been described in detail by Wickstorm (1982). Also theoretical issues in digital signal acquisition and filter utilization have been discussed by Clements (1998) and by Stoddard (1998) in depth. I therefore deal with only practical issues here.

1.1.1 Microphone

Since most bird sounds are within the range of human hearing, microphones available for consumer use can be utilized for professional researchers. When recording in noisy environment some degrees of directionality is preferred. In such cases, use of parabolic or gun-type microphones are recommended. They come with several characteristic directional sensitivity and readers are advised to consult homepages of dealers.

1.1.2 Recorders

Digital audio tape recorders (DAT) dominated for a period of avian bioacoustics and they are still in use. They have a 44.1 kHz sampling rate and this is enough for recording most bird sounds up to 20 kHz with stereo. However, equipment that records directly into digital memory has become available. Cautions should be paid in setting these equipment, since their default setting compresses sound data. Linear PCM setting or wave setting should be selected. More recent progress is direct hard disk recording with a personal computer. Some analyses software is adopted for such purpose. When a laptop or a notebook-type computer is used with such system, the entire system can conveniently be carried over the field.

1.2 Analyses Systems

There are at least four sound analyses system specialized for avian bioacoustics. One of them is free software and another three are commercially available.

Sound Analysis Pro (SA+) is an integrated system designed to manage the acquisition and analysis of animal vocalization. This software is specialized for zebra finch (*Taeniopygia guttata*) songs and suitable for laboratory studies. Some modifications may be necessary to use this for sounds in other species. SA+ is open source software developed by Ofer Tchernichovski and others, and is provided under GNU general public license for free. This system includes an automated operant training tool, a smart data acquisition system, a wealth of sound analysis toolboxes, and a fast database with exporting of data to EXCEL and to Matlab. In particular, SA+ is powerful tool for long-term vocal learning experiments; it automatically calculates acoustic features such as amplitude and a mean frequency from millions of sound data, and records them to a syllable table in the

database. It is easy to access a syllable table and perform its visualization and clustering. Furthermore, SA+ has a useful tool called dynamic vocal development (DVD) maps, showing vocal changes as movies. This software runs on Windows Platform (2000 or XP).

Raven is a software application for the acquisition, visualization, measurement, and analysis of acoustic signals, developed by the Cornell Bioacoustics Research Program (Charif et al. 2004). Raven has inherited the basic features of *Canary*, which is the former version, and gained more usability. The advantage of the software is the operating system independence. Raven runs on both Windows (98, NT, 2000, and XP) and Mac OS X (v10.2 or newer) while *Canary* runs only on Mac (OS 9.2 or older). Raven is also appealing in terms of user-friendliness and reasonable price. Currently two versions of Raven are available: Raven 1.2 and Raven Lite. The former targets biologists working with acoustic signals and the latter is for the learners. Raven 1.2 can record multiple channel audio into memory, read files in the formats of WAVE, AIFF, MP3, and AIFC, and calculate parameters varying parameters of the spectrogram. Raven is widely used in various researches in animal acoustic communication, ranging from elephants, whales to birds.

SASLab Pro is a powerful sound analysis and synthesis software provided by Avisoft Bioacoustics. It has been used by researchers investigating animal sound communication. Similar to other sound analysis software, it can record sounds through a sound card in the computer, filter to remove noise, make spectrograms, and edit arbitrary part of sound. It also has various advanced analysis tools such as auto- and cross-correlation, frequency shift using fast fourier transfrom (FFT) technique, root mean square, sound similarity matrix for comparison of spectrograms, and automated classification of syllables by means of spectrogram cross-correlation with templates. In addition, it acts as a synthesizer for generating artificial sound with setting parameters such as fundamental frequency, envelope, harmonics, frequency, and amplitude modulation. Combining these techniques, we can execute automated playback experiments based on real-time classification and re-synthesis.

SIGNAL is a signal processing graphical platform provided by Engineering Design. This software/hardware package has a history of more than 20 years and thus optimized for many applications in bioacoustics. SIGNAL is more like a computer language than application software. Since it is a programming environment, any requirement for signal processing can be satisfied when one got used to with this environment. Also, users

can either write a script program file to automate analyses procedures in SIGNAL or they can examine, step by step, effects of various manipulations. SINGAL also has an additional real-time spectrograph and an event detector module with which adaptive online experiment with sound and event I/O is possible. SINGAL is available both for Windows and McIntosh platforms.

2 • Hearing Performance

Contrary to a general belief, birds usually do not hear frequencies above 10 kHz and their range of hearing is very narrow. Avian basilar membrane is not coiled but straight and this is one of the reasons for their narrow range of hearing. The length of papilla is approximately 1.6 mm in the canary and zebra finch, 2.2 mm in budgerigar and European starling, 3.1 mm in pigeon, 3.7 mm in chicken, 4.2 mm in emu, and 9 mm in the barn owl (Gleich & Manley 2000). Most birds hear well between 500 Hz and 6 kHz (about 30 dB SPL (sound pressure level)) but outside of this range hearing deteriorates rapidly (Dooling 2004). However, within this range hearing, performance of birds rivals to that of humans and outperforms humans in some specialized domains. These behavioral data are summarized in Fay (1988).

Figure 1 compares audiograms from several avian species including the canary, song sparrow, budgerigar, pigeon, Starling and the zebra finch and average of these six species (Fay 1988). In these species, best frequency occur at between 2 and 4 kHz. Below this, hearing decreases at about 20 dB per octave and above this it decreases at about 30 dB per octave. Exceptions to this general pattern are obtained from birds of prey. Hawks, falcons, and owls

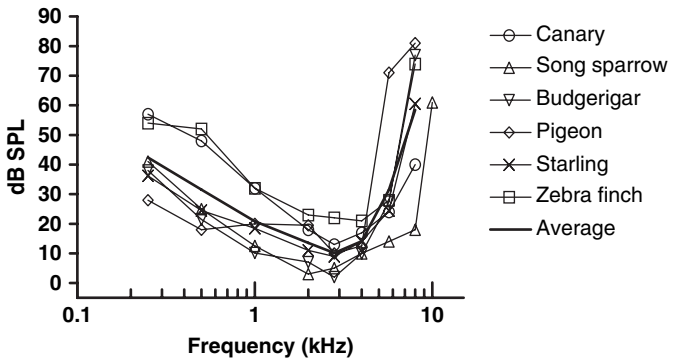


FIGURE 1 Behavioral audiograms of selected bird species. Most audiograms were obtained by operant conditioning with food reward. Data compiled in Fay (1988).

can hear a SPL as small as −20 dB at around 1 kHz. Hearing is one of the major tools for them for hunting and being able to hear at a low SPL was adaptive for them. However, these birds usually have narrower range of hearing especially at the higher frequency region. Small birds that are predated upon utilize the inaudible range for the prey when using alarm calls (Klump et al. 1986).

Within the narrow range of best hearing between 0.5 and 6 kHz, birds have an intensity difference limen of about 1–2 dB and frequency difference limen of about 2–8%. These are some of best performances within vertebrates almost rival to humans. Also as in humans, critical masking ratios increase as the center frequency increase, at the ratio of about 3 dB per octave in most birds. Minimum resolvable time difference is about 3 ms and maximum integration time is about 200 ms and both of these performance rival to humans (Fay 1988).

3 • Characteristics of Songs and Calls

In birds, vocalizations are roughly divided into two categories, calls and songs (Catchpole & Slater 1995). Calls are mostly used as non-sexual, social signals by both sexes. On the contrary, songs are mostly used as sexual communication signal by males. When used against other males, songs function as an agonistic signal and when used for females songs function as an attraction signal.

3.1 Types of Calls

Most species have several types of calls (Marler 2004). Calls are often classified in terms of behavioral context they are used. These contexts include separation, alarm, flight, aggression, courtship, mobbing, distress, contact, roosting, begging, and feeding. In several species of passerines, acoustic morphology of anti-predator alarm calls converged. These calls are mostly narrow banded and above 8 kHz, last about 100–200 ms with gradual rising and falling. These characteristics converged probably because of the context under which the alarm calls are used. The predators usually cannot hear signals above 8 kHz and the preys explored such higher region to convey alarm signals (Klump et al. 1986).

3.2 Song Syntax

Birdsong is delivered as an ordered sequence of discrete sound elements. In literature of birdsong studies, special terms are used to describe the hierarchical organization of song. Figure 2 is an example of a zebra finch song sequence. In this figure, the smallest unit of song is referred to as a note. A note can be defined in two ways. First, note is usually a discrete trace on the

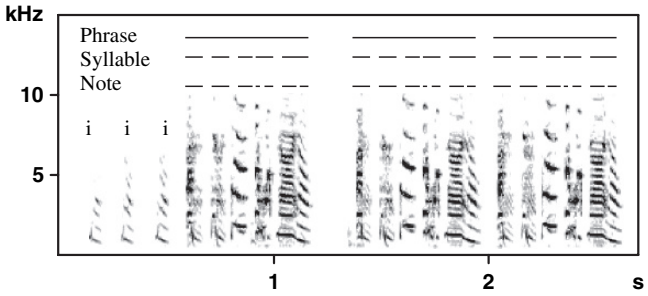


FIGURE 2 Example sonogram of a zebra finch song and terminology used in song learning literature.

sonogram, or a sequence of sound proceeded and followed by a silent interval. Second, a discrete trance on the sonogram can sometimes be composed of several notes when there is a sudden change such as frequency shift or intensity change in the trance continuity. In Figure 2, some of the seemingly continuous traces are separated into multiple notes by the second definition.

A syllable refers to a unit of appearance in the sonogram. Multiple notes can compose a single syllable when these notes always appear as a chunk. Single note can also be classified as a syllable when it appears as an independent entity. When several syllables are ordered in a stereotyped sequence, this sequence is called as a phrase. A motif is also used as a synonym for a phrase. The song notes denoted by "i" in Figure 2 are introductory notes and usually repeated several times before a phrase is initiated. A song bout consists of introductory notes and several repetitions of phrases. Most birds sing stereotyped song sequence but there are species with more elaborated song sequences. Starlings (Eens 1997), nightingales (Todt & Hultsch 1996), and Bengalese finches (Okanoya 2004) are some of these species and song of these species could be expressed as a finite-state syntax.

3.3 Song Phonology

Song phonology is determined by several factors. Birds use bilateral syrinx for phonation and interaction of these two structures results in complex sound interferences (Suthers 2004). The resulting source sound is transmitted through trachea and aural cavity, and finally broadcast through the beak. The widths of beak opening can also influence harmonic structures of song notes. The beak opening correlates with the fundamental frequency of the song notes. In helium air since velocity of the sound changes in such medium harmonic structure of song notes changes (Nowicki 1987) and this further influences syntax of song production (Figure 3) (Yamada & Okanoya 2003).

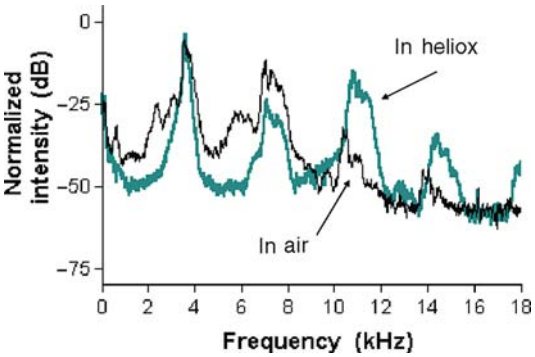


FIGURE 3 Effect of helium air in song phonology. The power spectrum of a song note of a Bengalese finch recorded in the normal air was compared with that recorded in the helium air. Note that the intensity of higher harmonics increased in the helium air. (re-plotted from Yamada & Okanoya 2004).

3.4 Song Learning

Behavior and neurobiology of song learning have been well documented (Marler 2004; Zeigler & Marler 2004) and this topic is only outlined here. Song learning occurs in two phases: sensory phase and sensorimotor phase. During sensory phase, chicks remember species typical song by selectively listening adult songs based on chicks' innate auditory preference. An auditory model of the species-specific song is established during this process. During sensorimotor phase, birds try to shape their own vocal output so that they can sing a similar song as the remembered auditory template. This process is enabled by auditory feedback and error correction learning. Naturally, sensory phase proceeds to sensorimotor phase, but species are different in the timing of these two phases. In migratory species these two phases usually do not overlap while in non-migratory species the two phases overlaps considerably.

Neural substrates for song learning have been studied extensively and here I will outline the current hypothesis (Figure 4). Based on the innate preference (neural substrates for this have yet to be identified), chicks attend to species-specific songs and specific excitation pattern is represented in the NCM during sensory learning phase. These patterns are further organized in the CMM and the neural representation of auditory template is probably established on these structures (Bolhuis & Gahr 2006). During sensorimotor phase, random excitation occurred in the posterior pathway by hormonal stimulation causes vocal output and that is fed back to the auditory system including NCM and CMM and then to HVC. The auditory template and auditory

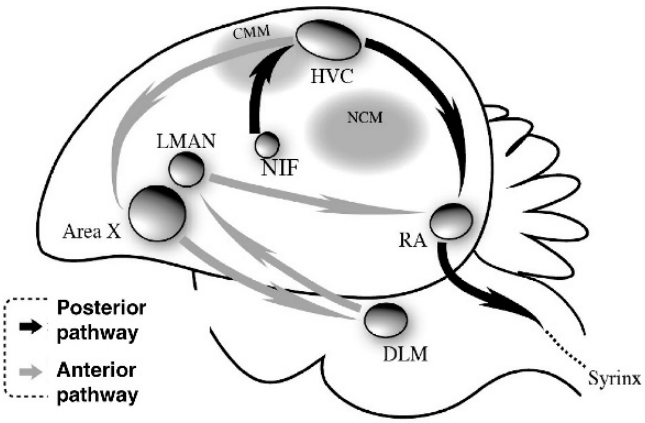


FIGURE 4 Brain regions important for song learing. CMM (Caudomedial mesopallium); HVC (Nucleus HVC of the nidopallium); NCM (Caudomedial nidopallium); LMAN (Lateral magnocellular nucleus of the anterior nidopallium); NIF (Interfacial nucleus of the nidopallium); RA (Robust nucleus of the arcopallium); DLM (the medical nucleus of the dorsolateral thalamus).

feedback are compared at the anterior pathway including area X and LMAN and if there are discrepancies neural connections within the posterior pathway are rewired. This will enable error minimization and song crystallization.

4 • Archives and Databases of Signals

There is a wealth of archives and databases for avian vocal signals. The most well known and complete is maintained by the Macaulay Library at Cornell University. Also, Avisoft Bioacoustics maintains a web site from where readers can easily try playback of some bird sounds. Many books and CDs are available and these are easily located through search engines. *Bioacoustics* is a journal dedicated to recording and analysis of biological sound and avian sounds are often treated.

References

Bolhuis JJ & Gahr M (2006) Neural mechanisms of birdsong memory. Nat Rev Neurosci 7, 347–357.

Catchpole PK & Slater PJB (1995) Bird Song: Biological themes and variations. Cambridge: Cambridge University Press.

Charif RA, Clark CW & Firstrup KM (2004) Raven 1.2 user's manual. New York: Cornell Laboratory of Ornithology.

Clements M (1998) Digital signal acquisition and representation. In: SL Hopp, MJ Owren & CS Evans (Eds.) *Animal Acoustic Communication: Sound analysis and research methods*. Berlin: Springer-Verlag, pp 27–58.
Dooling RJ (2004) Audition: can birds hear everything they sing? In: P Marler & H Slabbekoorn (Eds.) *Nature's Music: The Science of Birdsong*. New York: Elsevier Academic Press, pp 206–225.
Eens M (1997) Understanding the complex song of the European starling: an itegrated ethological approach. Adv Stud Behav 26, 355–434.
Fay RR (1988) Hearing in Vertebrates: a Psychophysics Databook. Winnetka: Hill-Fay Associates.
Gleich O & Manley GA (2000) The hearing organ of birds and Crocodilia. In: RJ Dooling, AN Popper & RR Fay (Eds.) *Comparative Hearing: Birds and Reptiles*. Berlin: Springer-Verlag, pp 70–138.
Klump GM, Kretzschmar E & Crio E (1986) The hearing of an avian predator and its avian prey. Behav Ecol Sociobiol 18, 317–323.
Marler P & Slabbekoorn H. (2004) Nature's Music: The Science of Birdsong. New York: Elsevier Academic Press.
Marler P (2004) Bird calls: a cornucopia for communication. In: P Marler & H Slabbekoorn (Eds.) *Nature's Music: The Science of Birdsong*. New York: Elsevier Academic Press, pp 132–177.
Nowicki S (1987) Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. Nature 325, 53–55.
Okanoya K (2004) Song syntax in Bengalese finches: proximate and ultimate analyses. Adv Std Behav 34, 297–346.
Stoddard PK (1998) Application of filters in Bioacoustics. In: SL Hopp, MJ Owren & CS Evans (Eds.) *Animal Acoustic Communication: Sound analysis and research methods*. Berlin: Springer-Verlag, pp 105–128.
Suthers RA (2004) How birds sing and why it matters. In: P Marler & H Slabbekoorn (Eds.) *Nature's Music: The Science of Birdsong*. New York: Elsevier Academic Press, pp 272–295.
Thorpe W (1961) Bird-song: the biology of vocal communication and expression in birds. Cambridge: Cambridge University Press.
Todt D & Hultsch H (1996) Acquisition and performance of song repertoires: ways of coping with diversity and versatiliy. In: DE Kroodsma & EH Miller (Eds.). *Ecology and Evolution of Acoustic Communication in Birds*. Ithaca: Cornell University Press, pp 79–96.
Wickstorm DC (1982) Factors to consider in recording avian sounds. In: DE Kroodsma & EH Miller (Eds.) *Acoustic Communication in Birds, Vol. 1*., New York: Academic Press, pp 2–52.
Yamada H & Okanoya K (2003) Song syntax changes in Bengalese finches singing in a helium atomosphere. Neuroreport 14, 1725–1729.
Zeigler HP & Marler P (2004) Behavioral neurobiology of birdsong. Annals of the New York Academy of Sciences 1016.

Web sites

http://www.birds.cornell.edu/macaulaylibrary/
http://www.avisoft-saslab.com/
http://www.zi.ku.dk/zi/bioacoustics/homepage.html

104

Cetacean Bioacoustics with Emphasis on Recording and Monitoring

Tomonari Akamatsu

National Research Institute of Fisheries Engineering, Fisheries Research Agency of Japan, Kamisu, Ibaraki, Japan

More than 80 cetacean species live in oceans, lakes, and rivers. For underwater navigation and recognition, whales and dolphins have developed unique sensory systems using acoustic signals. Toothed whales, such as dolphins and porpoises, have sonar using ultrasonic pulse trains called echolocations (Au, 1993). As top predators in the water, dolphins and porpoises rely on accurate and long-range sensory systems for catching prey. Dolphins have another type of vocalization called a whistle that is narrowband with a long duration. Most of the fundamental frequencies of whistles are in the human audible range, and this relatively low-frequency signal travels a longer range than sonar signals. The role of whistles are considered to maintain the mother–calf bond or to recognize each other in a breeding or migrating group of animals.

Baleen whales produce extremely low-frequency sounds, which can be as low as 17 Hz for blue whales. The low-frequency calls

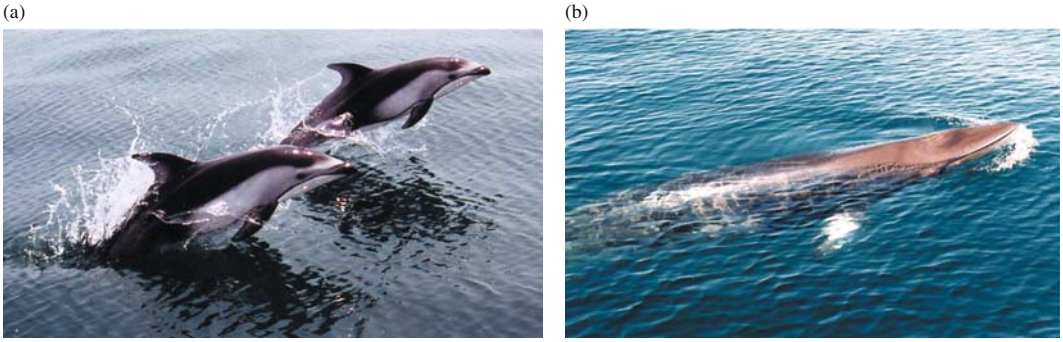


FIGURE 1 Pacific white-sided dolphins *Lagenorhynchus obliquidens* (toothed whale), *left*, and a minke whale *Balaenoptera acutorostrata* (baleen whale), *right*. These two animals have totally different life histories and use different vocalization frequency bands for sensing and communication. Photo by K. Sasamori.

from baleen whales are thought to be used for ultra-long-range communication. These calls span an area that is a distance of tens of kilometers or greater from the sound source. Baleen whales migrate thousands of kilometers seasonally. They breed in low-latitude waters during the winter and forage in high-latitude waters during the summer. A long-range communication tool is therefore advantageous in their vast breeding and foraging area. Their prey species are zooplanktons, small fish, and squids. They catch groups of prey in seawater and then filter the prey out of the water with their baleens. Thus, baleen whales do not need to sense and eat individual prey.

Toothed and baleen whales have totally different foraging and migration behaviors. The frequencies of their vocalizations are strongly affected by their life histories. Overview lists of the sound characteristics of dolphin and whale vocalizations are available from different sources, such as Wartzok and Ketten (1999), Richardson et al. (1995), and Au (1993).

1 • Recording and Monitoring Systems

Basic recording techniques, and tips and pit-falls for the recording and analysis of cetacean sounds are introduced in this section. Recording and monitoring systems for cetacean vocalizations are changing rapidly every year. Updated information on hardware systems, as well as software, journals, research group sites, and a directory of researchers, are available on the Internet. An excellent gateway for these resources is "Animal

Bioacoustics" at http://www.pmel.noaa.gov/vents/bioacoustics/Bioacoustics.htm.

1.1 Hydrophones

Underwater microphones, called hydrophones, are used to detect marine mammal vocalizations. A hydrophone should be chosen according to the frequency range of the target animal's vocalizations.

For dolphin whistles, commonly used hydrophones that are adjusted for the human audible range are appropriate. A high-pass filter is applied to eliminate background noise contamination. The cut-off frequency of the filter should be set between the target frequency of the animal and the noise frequency. Underwater noise characteristics are well summarized in Richardson et al. (1995). Many types of natural and anthropological noise sources are found at lower frequencies (Wenz, 1962). An appropriate cut-off filter frequency can be estimated for each species by examining the sound characteristic lists mentioned above.

Hydrophones adapted for high frequencies are used to record dolphin sonar signals. These hydrophones are relatively small and less sensitive than those used for lower frequencies. The dominant frequency of the sonar signals from dolphins and porpoises are out of the range of human audible frequencies, above 20 kHz. The source levels reach 177 dB for porpoises and 228 dB for bottlenose dolphins, *Tursiops truncatus*. This gives a good signal-to-noise recording ratio. However, the sonar signals of dolphins and porpoises are highly directional (Au, 1993). Therefore, the acoustic characteristics of the received sonar signals depend on the relative direction of the animal with respect to the hydrophone. The waveforms of off-axis signals are different from those of on-axis signals owing to greater attenuation of the higher frequency components.

Hydrophones adapted to low frequencies are required to detect baleen whale sounds. Various types of background noises inundate the frequency range of baleen whale vocalizations. This frequency range has been monitored for several decades for naval affairs purposes. Attempts have been made to record baleen whale sounds during this monitoring, these are described in Section 2, Locating and Tracking.

1.2 Recording Systems

Digital recording systems are now widely used to record cetacean vocalizations. Reliable portable equipment is desirable for field recording. Digital audio tape recorders (DATs) were commonly used, but hard disk recorders and laptop PCs are becoming more popular. Data recorders based on tapes, disks, and solid-state memory are suitable for multi-channel recording of hydrophone array systems.

It is essential to maintain the recording level within the dynamic range of the recording system to obtain undistorted data. All recording equipment have a specific dynamic range. First, the dynamic range of the recording system of the recording system and the target signals should be determined. Then, the amplification ratio must be adjusted so that the dynamic range of the recording system and target signal match. Clipped signals create an artificial frequency component that does not exist in the original signals, making it difficult to analyze the original sound characteristics. It is straightforward to determine whether clipped data have been included in a digital recording because the waveform becomes flat at the maximum points of the dynamic range of the recording system. It is useful to examine the recorded waveform for the exercise in advance to avoid clipping.

In all cases, an anti-aliasing filter is required to eliminate the phantom frequency component of the signal. This component is introduced during the analog to digital conversion of signals that have a frequency component higher than the Nyquist frequency. Most stand-alone recording systems, such as DATs, have anti-aliasing filters as a default. However, this should be verified before using the equipment.

A pit-fall is encountered when a band-limited recording system is used to record sonar signals. Most commercial systems only record a limited frequency band that is within the human audible range. The frequency band of biosonar signals from dolphins is wide. As the lower end of this spectrum is within the human audible range, the low-frequency components of biosonar signals can be recorded by a band-limited recording system. Even when an anti-aliasing filter is used, the recorded waveform is totally destroyed and is useless for acoustical analyses. It is only possible to measure the inter-pulse intervals from these recordings.

Mini disk (MD) and integrated circuit (IC) recorders must be used very carefully. The data compression algorithms used by these systems are adapted for human auditory systems. Most of the algorithms are one-way procedures, so it is impossible to recover the original waveform. Therefore, a waveform or spectrum analysis of an animal's vocalization using compressed data can potentially lead to wrong results.

Animal vocalizations have complicated characteristics, and it is difficult to obtain good-quality recordings in the wild. A quick examination and categorization of vocalizations is usually performed using a sonogram examined by human eyes. A sonogram is a convenient way to visualize the spectral and temporal characteristics of animal sounds, but the size of the fast Fourier transform (FFT) and the color scale of the display

strongly affect the visual impression, especially when analyzing pulse trains (Watkins, 1967).

2 • Locating and Tracking

In the water, long-range sensing and signal transmission depend on sounds instead of electromagnetic waves. This imposes a strong physical constraint. Many marine organisms use acoustic signals for their sensing and communication. We can detect mammals in the water by eavesdropping on their vocalizations. Passive acoustical monitoring methods are becoming popular for locating and tracking marine mammals.

Cetaceans' vocalizations are suitable for passive acoustical monitoring because of their frequent production and their unique acoustical characteristics. Finless porpoises, *Neophocaena phocaenoides*, produce several sonar pulse trains per minute, on average. They frequently use sonar to scan the distance beyond their visible range (Akamatsu et al., 2005). Baleen whales often produce sounds during their breeding season. Some calls have species-specific characteristics. In this case, it is possible to identify the species remotely.

Long-term continuous observations of cetaceans are quite difficult for human visual observers. Acoustical surveys of cetaceans have several advantages compared with visual surveys. An acoustical survey can be performed automatically and will maintain an identical detection threshold over long periods of time. An acoustical survey can perform continuous observations, even nocturnally. In contrast, the advantages of visual surveys are that they directly report the number of animals. In a large survey area, the number of animals can be estimated using a line transect survey, which is a well-established methodology (Buckland et al., 1993). Acoustical cetacean surveys still do not provide accurate estimates of the number of animals.

2.1 Monitoring By Using a Monaural Hydrophone

The detection distance of an underwater sound from a sound source depends on source level, background noise level, and detection threshold (Urick 1983). Sound propagation, absorption, and reflection also affect the maximum detection distance to the target sound. Roughly speaking, ultrasonic signals from dolphins and porpoises can be detected within several hundreds meters from the vocalizing animal. Detection ranges for such signals are limited due to absorption. For example, Wang et al. (2005) applied a stationary recording system for research on the occurrence of finless porpoises and showed that animals could be detected within distances of 150 m from the recording system.

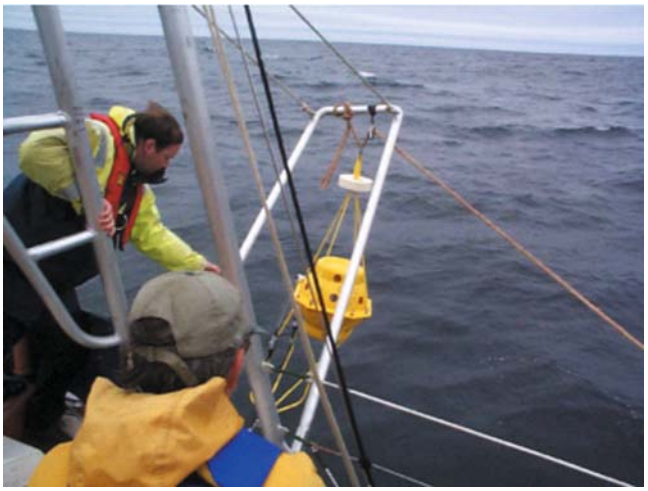


FIGURE 2 The pop-up recorder lowered to the ocean for the monitoring of baleen whale sounds, courtesy Christopher W. Clark, Bioacoustics Research Program/ Cornell Lab of Ornithology.

Many baleen whale vocalizations are of low frequency (<1000 Hz) and travel much greater distances than those of toothed whales. A device with a low sampling rate is good enough to record the baleen whales' sounds, thus enabling the application of long-term recording systems. For example, autonomous recording units (ARUs), deployed as a ocean-bottom recorder are used to monitor for baleen whales (Figure 2).

2.2 Tracking by Using a Hydrophone Array

The short baseline (SBL) method is an effective technique for locating and tracking dolphin sonar signals. Hydrophone arrays have been used to measure the relative location of an animal (Au and Benoit-Bird, 2003). The pulse duration of dolphin and porpoise sonar signals is less than 100 microseconds, during which time the sound travels 15 cm in water. Even using the SBL method, there is an observable time difference between the pulses received in different channels. Multiple stand-alone SBL systems work when locating the sound source. One stereo recording system provides the relative direction of the sound source, and two or more sound source directions obtained from different SBL systems provide the location of the sound source.

Long baseline (LBL) methods based on time-synchronized detectors on the seabed or sea surface are used to locate and track cetaceans (see, e.g., Janik et al., 2000). Hydrophone array systems (such as the U.S. Navy's SOSUS) have been used to locate

the source of underwater low-frequency baleen whale vocalizations (Nishimura and Conlon, 1994; Mellinger and Clark, 2003). The song of male humpback whales, *Megaptera novaeangliae*, produced in a train of vocal is the most famous baleen whale vocalization. This song is composed of winter units consisting of a predictable series of structures and is sung over and over again (Payne and McVay, 1971). The high redundancy and the predictable sequence of stereotypical units are typical characteristics of the song. Using a hydrophone array system, Charif et al. (2001) reported that humpback whales not only sing in their breeding areas but also on their migration routes.

3 • Echolocation Models

3.1 Transmission System

The sonar signal sound transmission system for dolphins is similar to a flashlight, which has a point light source, a reflector, and a lens. Dolphins and porpoises do not have vocal chords. In dolphins, the sound generator is a pair of fatty bursae embedded in soft connective tissue lips (Cranford et al., 1996). This is located in the upper nasal passage. The structure of the sound source is called MLDB, which stands for monkey lips/dorsal bursae (see Figure 3). The vibration of the fatty organs is caused by high-pressure air passing between them, which produces high-intensity short-duration pulse trains. The sound that propagates backward is reflected forward by the cranium. The sound that travels forward is focused by a melon organ, which is an acoustic lens. The sound speed at the outer shell of the melon is greater than that at the inner core. A numerical simulation of the sound

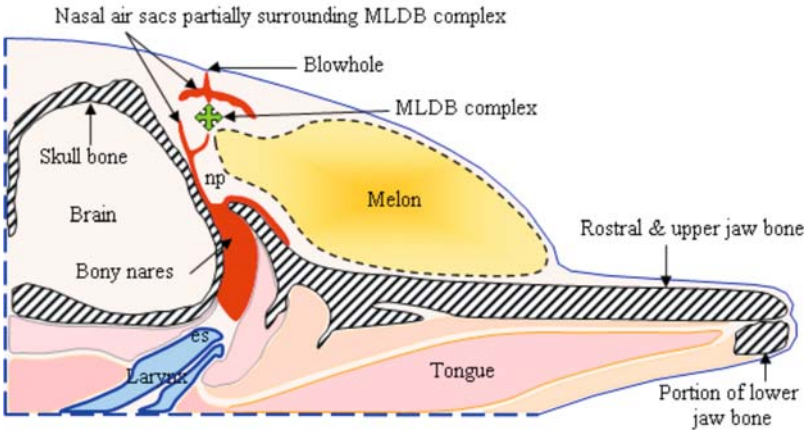


FIGURE 3 Diagram of selected dolphin head tissues, including the sound source (MLDB complex). Adapted from Aroyan, 1990.

propagation in the head region of a dolphin (Aroyan et al., 2000) produced results consistent with the experimentally measured directional beam pattern.

3.2 Receiving System; Comparing with Commercial Fish Echo Sounders Comparison Between the Cetacean Sonar and a Commercial Fish Echo Sounder

Underwater sound enters the lateral side of a dolphin's lower jaw; the ear hole does not function as a sound-receiving pathway. The high sensitivity of the jaw area has been confirmed by electrophysiological and behavioral studies (see, e.g., Bullock et al., 1968). Sound travels through the acoustic fat inside the lower jaw to the tympanoperiotic bone, which surrounds the inner ear system.

The inner ear system is similar in all mammals, including cetaceans. Toothed whales, such as dolphins and porpoises, have thicker and narrower basilar membranes to detect high-frequency sounds. On the other hand, baleen whales have an inner ear system adapted for low frequencies. Au et al. (2000) provided a good overview of the sound-receiving system of cetaceans while Supin et al. (2001) focused on the physiological aspects of cetacean hearing systems.

A great number of classical conditioning and auditory physiological experiments have revealed that the hearing range of dolphins and porpoises is sensitive up to ultrasonic frequencies. It is natural for fish eaters to use high-frequency sounds in water because ultrasonic sound is reflected well by their prey such as fish, especially by the swim bladders. From a physical point of view, higher frequency sounds would be reflected from even smaller targets, but this sound would travel a shorter distance owing to absorption. A good compromise corresponds to the frequency range of toothed whales' sonar, which extends from several tens to one-hundred kilohertz. The trade-off between the sensing distance and the echo level of the prey species determines the appropriate frequency that can be used to detect fish. The estimated target range of free-ranging dolphins and porpoises is several tens to one hundred meters. A bottlenose dolphin is reported to detect the existence of a metal sphere target, 7.62 cm in diameter, at distances up to 113 m (Au, 1993). A commercial fish echo sounder employs a frequency band similar to that of the dolphins' sonar, such as 38, 70, and 120 kHz, to detect fish resources up to several hundreds of meters away (MacLennan and Simmonds, 1992). The frequency selection of dolphins' sonar and fish echo sounders appears to be ideal for the acoustical detection of fish.

The directional receiving system and source level of dolphins' sonar and fish echo sounders are also comparable. Au (1993) is a useful reference for the remaining topics in this section.

Dolphins have directional hearing ability. They are sensitive to sound that comes from in front of them over an approximate +/− 20-degree range, depending on the sound frequency. A sharper receiving beam pattern has been observed for higher frequency sounds. Fish echo sounders also have directional receivers to improve their signal-to-noise ratio. The source level of a bottlenose dolphin reaches 228 dB re. 1 uPa peak-to-peak, which is just below the cavitation threshold. This is also similar to the source level of a fish echo sounder.

However, the spectrum characteristics of the sonar signals of dolphins and porpoises are quite different from those of a fish echo sounder. A sonar signal pulse only has several oscillations and lasts a very short period of time, approximately 100 microseconds. This means, the biosonar signals of small odontocetes have a broadband spectrum, whereas the fish echo sounder employs narrowband monochromatic signals. Narrowband signals improve the signal-to-noise ratio by using a frequency filter to eliminate the noise component outside of the signal band. However, the echo of a narrowband signal only provides the echo level for a specific frequency. In contrast, broadband sonar signals provide various types of frequency response information for a target, but they have a low signal-to-noise ratio.

The echolocation performance of toothed whales has been studied extensively, both for basic science and for the application of underwater exploitation tasks. A bottlenose dolphin can discriminate a 0.3 mm difference in the thickness of a cylindrical target. Material composition, such as bronze, aluminum, or coral rock, can be immediately judged by dolphins. Interestingly, once the echo frequency is modulated to the human audible range, humans, like dolphins, can also discriminate material composition by listening to the echo (summarized in Au, 1993). A dolphin can match target shapes to visual images (Harley et al., 2003), but the signal processing mechanisms of this task have not yet been revealed.

For an updated overview, Thomas et al. (2003) is a useful reference. A series of books, including Kastelein et al. (1994), Thomas et al. (1992), Thomas and Kastelein (1990), Nachtigall and Moore (1988), and Busnel and Fish (1980), detail various aspects of biosonar studies in cetaceans.

References

Akamatsu, T., Wang, D., Wang, K., and Naito, Y. (2005), Biosonar behaviour of free-ranging porpoises, Proceedings of the Royal Society of London Series B, 272, 797–801.

Aroyan, J.L., McDonald, M.A., Webb, S.C., Hildebrand, J.A., Clark, D., Laitman, J.T., and Reidenberg J.S. (2000), Acoustic Models of Sound Production and Propagation. In: Hearing by Whales and Dolphins. Au W.W.L., Popper A.N., Fay R.R. (eds.), New York: Springer-Verlag, pp. 409–469.

Au, W.W.L. (1993), The sonar of dolphins. Springer-Verlag, New York, Berlin, London, Tokyo, pp. 277.

Au, W.W.L., Popper, A.N., and Fay, R.R. (eds.) (2000), Hearing by Whales and Dolphins. Springer Handbook of Auditory Research, Springer-Verlang, New York, pp. 485.

Au, W.W.L. and Benoit-Bird, K.J. (2003) Automatic gain control in the echolocation system of dolphins, Nature, 423, 861–863.

Buckland, S.T., Anderson, D.R., Burnham, K.P., and Laake, J.L. (1993), Distance Sampling Estimating Abundance of Biological Populations. Chapman & Hall, London, pp. 446.

Bullock, T.H., Grinnel, A.D., Ikezono, E., Kameda, K., Katsuki, Y., Nomoto, M., Sato, O., Suga, N., and Yanagisawa, K. (1968), Electrophysiological studies of central auditory mechanisms in cetaceans, Zeitschrift Fur Vergleichende Physiologie, 59, 117–156.

Busnel, R.-G and Fish, J.R. (eds.) (1980), Animal Sonar systems. New York, Plenum Press, pp. 449–467.

Charif, R.A., Clapham, P.J., and Clark, C.W. (2001), Acoustic detections of singing humpback whales in deep waters off the British Isles, Marine Mammal Science 17, 751–768.

Cranford, T.W., Amundin, M., and Norris, K.S. (1996), Functional morphology and homology in the odontocete nasal complex: implications for sound generation, Journal of Morphology, 228, 223–285.

Harley, H.E., Putman, E.A., and Roitblat, H.L. (2003), Bottlenose dolphins perceive object features through echolocation, Nature, 424, 667–669.

Janik, V.M., van Parijs, S.M., and Thompson, P.M. (2000), A two-dimensional acoustic localization system for marine mammals, Marine Mammal Science 16(2), 437–447.

Kastelein, R.A., Thomas, J.A., and Nachtigall, P.E. (eds.) (1994)., Sensory systems of aquatic mammals. De Spil Publishers, Woerden, The Netherlands, pp. 588.

MacLennan, D.N. and Simmonds, E.J. (1992), Fisheries acoustics, Chapman & Hall, London, pp. 325.

Mellinger, D.K. and Clark, C.W. (2003), Blue whale (Balaenoptera musculus) sounds from the North Atlantic, Journal of the Acoustical Society of America, 114, 1108–11119.

Nachtigall, P.E. and Moore, P.W.B. (eds.) (1988), Animal sonar: processes and performance. Plenum Press, New York, pp. 862.

Nishimura, C.E. and Conlon, D.M. (1994), IUSS dual use: monitoring whales and earthquakes using SOSUS, Marine Technology Society Journal, 27(4), 13–21.

Payne, R.S. and McVay, S. (1971), Songs of humpback whales, Science, 173, 585–597.

Richardson, W.J., Greene, Jr. C.R., Malme, C.I., and Thomson, D.H. (1995), Marine mammals and noise, Academic Press, San Diego, New York, Boston, pp. 576.

Supin, A.Y., Popov, V.V., and Mass, A.M. (2001), The sensory physiology of aquatic mammals, Kluwer Academic Publishers, Boston, pp. 332.

Thomas, J.A., Cynthia, F.M., and Marianne V. (eds.) (2003), Echolocation in Bats and Dolphins, The University of Chicago Press, Chicago, p. 631.

Thomas, J.A., Kastelein, R.A., and Supin, A.Y. (eds.), (1992) Marine mammal sensory systems. Plenum Press, New York, pp. 773.

Thomas, J.A. and Kastelein, R.A. (eds.) (1990), Sensory abilities of cetaceans: laboratory and field evidence, Plenum Press, New York, pp. 710.

Urick, R.J. (1983), Principles of underwater sound / 3rd edition, McGraw-Hill, New York, pp. 423.

Wartzok, D. and Ketten D. (1999), Marine Mammal Sensory Systems, In: Biology of Marine Mammals. Reynolds III J.E., Rommel S.A. (eds.), Smithonian Institution Press, Washington and London, pp. 117–175.

Watkins, W.A. (1967), The harmonic interval: fact or artifact in spectral analysis of pulse trains, In: Marine bio-acoustics, Vol. 2. Tavolga W.N. (ed.), Pergamon Press, New York, pp. 15–43.

Wang, K., Wang, D., Akamatsu, T., Li, S., and Xiao, J. (2005), A passive acoustical monitoring method applied to observation and abundance estimation of finless porpoises, Journal of the Acoustical Society of America, 118, 1180–1185.

Wenz, G.M. (1962) Acoustic ambient noise in the ocean: Spectra and sources. Journal of the Acoustical Society of America, 34,1936–1956.

List of Important Abbreviations

A- Interface	Interface between the BSS and the MSC, manages the allocation of suitable radio resources to the MSs and mobility management
A-bis	BSS internal interface linking the BSC and a BTS
ACELP	Algebraic CELP (with special codebook, see CELP)
ACR	Absolute Category Rating
ADPCM	Adaptive Differential PCM (with adaptive prediction and quantization)
AHr,dt	Attenuation Range Receiving, Double Talk
AHs,dt	Attenuation Range Sending, Double Talk
AMR	Adaptive Multi-Rate coding
APC	Adaptive Predictive Coding (also “block-ADPCM”)
APP	A-Posteriori Probability
ATC	Adaptive Transform Coding
A-ter:	BSS internal interface linking the BSC and a TRC
ATM	Asynchronous Transfer Mode
BB-RELP	Base-Band (i.e., low-frequency band) RELP (see RELP)
BER	Bit Error Rate
BFH	Bad Frame Handling
BFI	Bad Frame Indication
BSC	Base Station Controller
BTS	Base Transceiver Station
C/I	Co-channel interference
CCR	Comparison Category Rating
CELP	Code-Excited Linear Predictor/Prediction
COVQ	Channel-Optimized Vector Quantization
CRC	Cyclic Redundancy Check
CSI	Channel-State Information
DCR	Degradation Category Rating

DCT	Discrete Cosine Transformation
DFT	Discrete Fourier Transformation
DSP	Digital Signal Processor
DTX	Discontinuous Transmission
D-value	Relationship between direct sound field sensitivity and diffuse sound field sensitivity
EDGE	Enhanced Data-rate for GSM Evolution (an add-on to GSM)
EFR	Enhanced Full-Rate system (with the original GSM data rate and better quality)
EMBE	Enhanced MBE (see MBE)
ETSI	European Telecommunication Standardization Institution
eXCELP	Extended CELP (with classified codebooks, see CELP)
FER	Frame Error Rate
FIR	Finite (-length) Impulse Response (also "non-recursive") filter
FR	Full-Rate system (= the first standard GSM speech coding at 13 kb/s)
GSM	Global System for Mobile communication
HC	Harmonic Coding (= SM with a frequency grid, see SM)
HR	Half-Rate system (with half the original GSM data rate and less quality)
HVXC	Harmonic Vector-eXcited Coding (a HC / VQ combinaztion, see HC, CELP, VQ)
IMBE	Improved MBE (see MBE)
IP	Internet Protocol
ISCD	Iterative Source-Channel Decoding
ISO	International Standardization Organization
ITU	International Telecommunication Union (formerly: CCITT, Conseil Consultative de la Téléphonie et Télégraphie)
JLR	Junction Loudness Rating
LMS	Least Mean Squares (algorithm)
LP	Linear Prediction (= linear estimation of a new sample from n past samples)
LPC	Linear Predictive Coding (used generally for the filter inside all predictive coders or specifically for the LP-based VOCODER)
LSF	Line Spectral Frequencies (= transformed LPC coefficients)
LTP	Long-Term Prediction (with a large lag in the order of a "pitch" period)
MBE	Multi-Band Excitation (of an LP filter)
Mops	Mega (i.e., million) Operations per Second
MOS	Mean Opinion Score
MOS-LQO	Mean Opinion Score, Listening Quality Objective
MPE	Multi-Pulse Excitation (with a reduced number of entries on a general time grid)
MPEG	Moving-Picture Expert Group (also standardizing speech and audio systems)
MSC	Mobile Switching Center

MSE	Mean-Square Error
NLMS	Normalized Least Mean Squares (algorithm)
NLP	Non Linear Processor
OLR	Overall Loudness Rating
PCM	Pulse-Code Modulation (= analog-to-digital conversion, with linear or possibly non-linear quantization)
PLMN	Public Land Mobile Radio
PSD	Power spectral density
PSTN	Public Switched Telephone Network
R	Rating Factor, used in the E-model for transmission planning
RELP	Residual (signal)-Excited Linear Predictor
RLR	Receiving Loudness Rating
RLS	Recursive Least Squares (algorithm)
RPE	Regular-Pulse Excitation (with a reduced number of entries on a regular time grid)
SBC	Sub-Band Coding
SIRP	Spherically Invariant Random Process
SISO	Soft-In/Soft-Out
SLR	Sending Loudness Rating
SM	Sinusoidal Modeling (of signals by sums of general sinusoids)
SNR	Signal-to-Noise (power) Ratio (usually on a log. scale, in dB)
SOVA	Soft-Output Viterbi Algorithm
TC	Transform Coding
TDM	Time Division Multiplex
TELR	Talker Echo Loudness Rating
TELRdt	Talker Echo Loudness Rating, Double Talk
TRC	Transcoder Controller
UEP	Unequal Error Protection
UMTS	Universal Mobile Telecommunication System
VAD	Voice Activity Detection
VOCODER	Voice CODER with parametric speech description and synthesis
VoIP	Voice over Internet Protocol
VQ	Vector Quantization (also “block quantization”)
WB	Wide-Band (speech, with ∼7 kHz bandwidth rather than 3.5 kHz)

Index

Abnormal child births, 1226
ABR wave V, 166
Absolute category rating (ACR), 500, 833
Absolute magnitude estimation, 219, 1198, 1205
Absorption coefficient, 680
Accelerometer
  compression mode, 1318
  shear mode, 1317–1318
  highimpedance charge output, 1319
  see also Piezoceramic accelerometers
Accreditation, 1307–1308
Accuracyregularization frontier, 523
Acoustic glints, 1840–1841
Acoustic impedance
  flowmeter, 996–997
  tube measurements, 103
Acoustic intensity map, 1129
Acoustic measurement, 662–663, 1302
Acoustic micrometer, 984–986
Acoustic middle ear reflex (AMER), 161–162
Acoustic optical modulator (AOM), 1333
Acoustic particle motion, 1851–1854
Acoustic rate gyro, 984, 995–996
Acoustic reflex threshold (ART), 161
Acoustic surface area meter, 989–994
Acoustic telemetry, 1857
Acoustic volume meter, 984–989
Acoustical loading, 629
Acoustical properties, measurement, 949
Acoustic-gravity waves, 1511
Acoustic-tube modeling, 402
Acousto-mechanical design, 1269
Active and reactive sound fields, 1111
Active noise cancellation (ANC), 169
Active sonar, 1737–1754
Actuator, 1141–1151
  electrodynamic, 1144–1145
  electromagnetic, 1144–1145
  electrostatic, 1304
  hydraulic, 1142–1143
  intertial, 1143
  magnetostrictive, 1145
  magnetostrictive, 1145–1150
  MEMS, 1151
  piezoceramic, 1079
  piezo-electric (electrostrictive), 1148
  pneumatic, 1143
  shape memory alloy, 1147
Adaptive filters, 887, 890, 1030
Adaptive line enhancer (ALE), 128
Adaptive multirate (AMR), 602, 876
Adaptive procedure, 217
ADAT format, 706
Advanced authoring format (AAF), 717–718
Aerodynamic noise, 963, 967
Affine projection (AP) algorithm, 888
Air absorption, 764
Air conditioner noise, 225
Aircraft landing noise, 1014
Aircraft noise
  detection of, 972–977
  induced annoyance, 1009–1010
Air-gun, 1546–1547, 1551–1552
Akaike information criterion, 1605–1606
Akaike weights, 1606
Alcoholic Gasserian rhizolysis, 1220
Algebraic code-excited linear prediction (ACELP), 590, 601, 603, 605, 608, 610, 863, 864
Algorithmic composition, 339
Aliasing, 20, 22–24, 39–40, 43, 54–55, 68, 74, 88, 137, 142, 407, 695, 727, 802, 888, 902, 1030–1031, 1040, 1102, 1136, 1403, 1426, 1428, 1488–1489, 1510, 1550, 1555–1556, 1559, 1764, 1863, 1864
  leakage, 40
Alpha-trimmed mean, 1614
Ambi-handed subjects, 1202
Ambisonics, 753, 758, 912
Amplitude
  clipping, 774
  modulated noise signals, 488
  modulated wideband noise, 248
  modulation following responses (AMFR), 168
  modulations (AM), 179
  panning, 750–753, 756
  range relationship, 1512
  tracking, 1078

AMR-WB, 599–601, 604, 607, 610
Analog-to-digital converter, 19, 21, 29, 40, 94, 1598, 1760
Analogue controllers, 1086
  features, 1103
Analogue recording gear, 76
Analysis-by-synthesis, 368, 590, 597, 862, 867, 875
Anechoic listening, 750
Anti-aliasing filter, 20, 22–24, 55, 103, 1031, 1102, 1900
Anti-aliasing, 22, 55
Aperture function, 1488–1489
Aperture length, 1136, 1776
Approximate transcription, 318
AR model, 776–777, 780
ARIB EFR-PDC, 603
ARMA modeling, 780
Array
  broadside, 901, 1711, 1718, 1763, 1767
  conformal, 1718–1719, 1811, 1822
  delay-and-sum, 1404–1405
  differential microphone, 691, 696, 912
  line, 670, 676, 963, 965, 970, 1134, 1282, 1474, 1663, 1711, 1713–1715, 1717–1719, 1759, 1762–1765, 1767, 1773–1775, 1777, 1821, 1829
  planar, 140, 696, 1498, 1719, 1808, 1822–1823
  processing, 1465
  rectangular, 1134
  response, 1426, 1488
  stack, 1552, 1554
  three-dimensional, 1718–1719
  See also Infrasound
    Loudspeaker arrays
    Microphone arrays
    Pipe arrays
    Seismic arrays
Articulation Index (AI), 201, 486
Articulation, 201, 369, 375, 440, 486, 488, 513, 733–735
Articulation-to-non-articulation ratio (ANR), 513
Articulatory modeling/synthesis, 581
Articulatory synthesis, 368, 561
Asper, 264
Asynchronous signal, 69, 80
Atmospheric models, 1439, 1501, 1513
Atmospheric variability, 1439, 1502–1503
Atmospherical tomography, 94, 95
ATRAC, 122, 720, 811
Attack behavior, 422
Attenuation, 765, 778, 845, 954, 959, 1070, 1530–1531
Audibility of the error, 406
Audio
  bandwidth, 595–598
  bitrate, 800
  browsing, 310
  coders, 599
  coding, 277, 597–598, 602, 605, 614, 801, 866, 1419
    algorithms, 614
  compression systems, 231
  Engineering Society Anthology series, 625
  restoration, 773–775, 780
  thumbnail, 320, 327
Audiological diagnostical methods, 149
Audiovisual
  interaction, 731–749
  perception, 731, 733, 738
  space, 731
  speech, 731–738, 741
Auditory
  brainstem responses (ABR), 166
  cortex, 170, 187, 215, 739–741, 1844
  distance (AD), 508
  filters, 185, 232, 234, 236
    bandwidth, 234–235
  function, physiological measures, 159–171
  information, 154, 1868
  models, 175, 187, 236, 281, 287
  applications of, 187
  nerve, 154, 156, 161, 166–177, 215, 236, 281–282, 1868, 1873
  non-intrusive quality estimation (ANIQUE), 512–513
  pathway, 154–158, 166, 168, 1868, 1870
  perception
    models of, 180
    principles of, 206
  periphery model, 281
  processing models, 175–188, 206
  scene analysis, 286, 1852
  signal processing, 175, 187
  steady-state responses (ASSR), 167
  system, functional model, 203
Auralization, 81, 677, 762–768
Autocorrelation, 10, 45–46, 48, 88, 91–93, 102, 183, 279–280, 283–284, 295, 318, 367, 466, 470–471, 569, 598, 858, 872, 886, 888–889, 891, 1046, 1057, 1777
Automatic event location, 1468
Automatic gain control (AGC), 206, 920, 1760, 1851, 1864
Automatic microphone scanning system, 966
Automatic music composition, 339
Automatic network processing, 1467–1469
Automatic phase association, 1468
Automatic repeat request (ARQ), 608
Automatic segmentation, 327
Automatic speech recognition (ASR), 186, 188, 203, 483, 521–522, 524
Automatic tone modeling, 287
Automatic unit selection, 563
Autoregressive (AR) models, 776
Autoregressive (AR) predictors, 609
Autospectra, 30, 45, 47, 69–71
Autospectrum, 42, 46–47, 70–71
A-weighting, 107, 109, 111, 117, 1078
Azimuthal deviation, 1438, 1449, 1451, 1500, 1503–1504

Babbling, 676, 735–736
Background noise, 825, 846
Back-plate, 1247, 1252, 1257
Bad frame handling (BFH), 870
Bad frame indicator (BFI), 611
Baleen whales, 1891, 1895, 1897–1898, 1901–1902, 1904
Band cancellation filtering, 283
Band frequency scales, 486
Band widening, 233
Band-limited random noise, 248
Band-pass filters
  function, 1366
  values, 116

Bandwidth Limitations, 1085
Bark
  loudness spectra, 503, 506
  power spectrum, 501
  scale, 268, 480, 486, 501, 504, 507, 509, 856
  spectral distortion, 486, 501
  value, 270
Barometer, 24–25
Barometric equalization, 1269
Barrel stave, 1792, 1796–1797
Bartlett estimator, 1488
Basilar membrane, 151, 156, 159, 166, 176, 215, 230, 232, 236, 265, 268, 451, 798, 1193, 1896, 1904
Basilar papilla, 1868–1869
Basis spectrum, 293, 378–379, 381, 383–384
Bat detectors, 1836–1837, 1847
Baum-Welch algorithm, 312
Bayesian
  Cramér-Rao bound, 1608
  framework, 291
  inversion, 1637
  network, 523, 526–527
  probability network, 287
BCJR algorithm, *see* Viterbi algorithm
Beam pattern, 1490, 1716, 1764–1765, 1802, 1804–1805, 1813–1814, 1904–1905
Beamformer, 141, 692–694, 696, 698, 700, 901–903, 1714, 1717, 1762–1767, 1770–1771, 1776
  delay-and-sum, 140–141, 694, 697, 1762
  high-resolution, 1717
Beamforming, 4, 139–140, 143–144, 691–693, 695–700, 900–902, 911–913, 1487–1489, 1656–1658, 1661, 1714, 1719, 1762, 1764, 1766–1768, 1770, 1774, 1776
  blind, 699, 904
  data-independent, 693–696, 900–902
  filter-and-sum, 696
  least-squares, 1489–1490
Beampattern, 695
Beamwidth-aperture product, 1713
Bearing estimation, 1490–1492, 1709, 1717, 1773, 1775
Bearing tracker, 1778
Bearing-time intensity, 1771
Beats, 232, 236
Bel canto, 361
Bell clapper, 424, 442
Bender bar, 1792
Best linear unbiased estimator, 1703
Best modulation frequencies (BMFs), 180
Bidirectional delay line, 404, 409
Binary gold sequences, 87
Binaural
  cells, 1873
  cue coding, 803
  dereverberation, 209
  hearing, 200, 207–208
  masking level difference, 237
  perception, 181
  processing, 178
  reproduction, 755
Biophysics, 1206
Biosonar, 1900, 1905
Biot theory, 1541–1542
Birth-death detection algorithm, 1369
Bit error rate (BER), 602, 827–828, 868–871, 876, 919
Bitrate reduction, 800
Blackman window, 58, 268
Blade passage frequencies (BPFs), 1076–1077, 1079
Blind source separation, 292–293, 691, 699, 904, 911
Blocked pressure, 1825
Blocking matrix, 698, 903–904
Blow hole, 429
Blown instruments, 420–421, 424, 431–432
Boatwhistle, 1860
Bodes sensitivity, 1092
Bootstrapping, 1616
Bottom loss, 1674–1675, 1725
Boundary element method (BEM), 762
Boundary layer, 640, 992, 1000, 1285, 1439, 1455, 1473–1474, 1483, 1525, 1822–1824
Bounded input bounded output (BIBO), 13
Bowed-string, 403, 406
Bowing pressure, 421
Bowstring
  dynamics, 401
  junction, 405
Braille, 1233, 1216
Brainstem processing models, 178
Brainstem-evoked response audiometry, (BERA), 154, 166
Bright sound, 421
Brightness matching, 392
Broadband
  detector, 1768, 1771–1772
  models, 1503
  noise, 265
  signal, 27, 29, 31, 57, 113, 697, 1668, 1682, 1687, 1768, 1771, 1777, 1859
Brunt-Väïsälä frequency, 1444
Bubaran, 348
Bucket brigade, 1358
Buried Markov models, 528, 550
Butterworth, 22–23, 624, 652–653

Calculated roughness, 268, 272–273
Calibration, 24–25, 1120, 1288
  couplers, 1298
Call types, 1836, 1846, 1870, 1875–1878
Capacitive sensors
  application, 1344
  electric circuit, 1341
  environmental consideration, 1343
  operation principle, 1339–1341
  selection of, 1343
  target and guard plate, 1340
Cardiac gating, 170
Cardioids, 756, 1279
Carpal tunnel syndrome, 1219, 1221
Category scales, 219
Cauchy distribution, 1389–1391, 1397, 1596, 1600–1601
Causalgia, 1220
CDMA2000, 607
CELP, 605–606, 610, 862, 865, 876, 922
Central auditory pathway, 1868
Cepstral, 188, 307, 365–366, 449–450, 475–476, 478, 480–481, 498–499, 522, 1399, 1579–1581
  analysis, 475
  distance (CD), 498–499

representation, 365–366
transform, 475–476, 478, 1579
Cepstrum, 1579, 1399
Cerebellar dysfunction, 1225
Channel codewords, 857
Channel-optimized vector quantization (COVQ), 609, 875
Chaotic dynamics, 7
Charge sensitivity, 1315
Chebyshev, 22–24,
Checkerboard, 314–316, 328
Chirp, 10, 29, 90, 166, 1375, 1377
Choip, 63
Chopper microphone, 984, 997–999
Chorus, 789, 1878
Chorusing, 1878
Choruslikeness, 324
Chowning law, 751
Chroma spectrum, 307
Chroma vector, 307–309, 323, 328
Chromagram, 307
Chronic neuralgia, 1220
Cinema sound, 748, 800, 1280
Cipher notation, 335
Clarinet, 409–410, 412, 420, 431–432
Classic horn theory, 639
Classical guitar, 436
Click detection, 776–777
Click evoked emissions, 163
Click suppression, 777
Closed coupler, 1874
Closed-loop experiments, 1873
Cluster synthesis, 384
Clustering algorithms, 328
Clustering segments, 305, 315
Clustering, 310
CNS, 1870, 1873
Co-array function, 1489
Coarticulation, 560
Cochlea, 151
Cochlear
amplification, 177
excitation pattern, 182
frequency map, 181
implant, 176, 1239, 1241
transformation, 176
Cocktail party phenomenon, 200
Codebook, 368, 588–590, 605–606, 608–610, 860–862, 865, 875–876
adaptive, 588–590, 605–606, 610, 862
fixed, 588
innovative, 588–590, 605, 608–609
Codec
AMR narrowband, 603
free lossless audio, 801
variable bit rate, 606
Code-excited linear prediction (CELP), 368, 480, 603
Coder
harmonic vector-excited, 877
tandeming of speech, 921–923
Coding
adaptive transform acoustic (ATRAC), 122, 720, 811
binaural cue, 803
channel, 602–603, 608–610, 709, 857, 869–868, 874–877, 918–919
frequency-domain, 865
gap, 856
general audio, 600
generic model, 608
harmonic, 866
linear predictive, 367, 449–450, 464, 466, 469–470, 480–481, 588
loss-less, 122
multipulse linear prediction, 572
nonnegative sparse algorithm, 297
RELP, 858
sparse, 292–293, 296–297
transparent, 801, 804
See also Audio coding
Cognitive information processing, 157
Cohen class distributions, 49
Coherence function, 70
Coherent field, 1392
Coherent power, 47, 71
Coincidence detection, 179, 247, 1837
Collage, 350–351
Columella, 1866–1867
Comb function, 137, 473
Comb-filter effects, 658, 672, 685, 753, 757
Comfort noise, 603, 606, 828, 910, 916, 918–918, 924
Command signal, 1082–1085
Commercial seismology, 1531–1533
data acquisition costs, 1546
Common chunk, 712
Common focus point (CFP) technology, 1568
Common-midpoint (CMP) stacking, 1552–1553, 1559
Communication, 1692, 1852, 1862, 1895
man-machine, 176
Commuted waveguide synthesis, 409
Comodulation masking release (CMR), 237
Companding, 455
Comparison calibration, 1293, 1303, 1307
Comparison category rating (CCR), 833, 838
Compensation schemes, 205–206
Complementary sensitivity function, 1085–1086, 1090, 1101
Complex smoothing, 659
Complex stimulus, 243
Compositional interface, 334
Compressional wave, 1526–1528, 1537–1538, 1542, 1547, 1586, 1673, 1675, *see also* P waves
Computational auditory scene analysis (CASA), 209
Computer assisted composition, 339
Computer music composition, 333–334, 340–341, 355
Computer notation programs, 335
Concatenation cost, 558, 564–565, 573–574, 577
Concatenative, 480, 558, 561, 569, 573, 575, 577–578, 580–581
Concentration matrix, 526–529
Conditional probability density function, 1729
Conditions for stability, 1046
Conservation of sound energy, 1110
Consistency parameter, 1428
Consonant-vowel syllables, 736
Constant beamwidth transducer, 1801
Constant delay, estimation, 494
Constant percentage bandwidth, 33–34, 43–44, 122
Constant-Q transform (CQT), 308
Continuous judgment, 220–221, 224–227

Continuous wave
models, 1503
parabolic equation, 1506–1508
Controlled illumination, 1568
Convective peak, 1825, 1827
Convergence zone, 1671, 1687–1688
Conversation tests, 833
Conversational situation, 822, 824–825, 831, 835, 834, 923
Convolution theorem, 35–36, 40
Copolymer, 1788
Corpus-based synthesis, 563
Correlation function, 44, 1427, 1488–1489, 1491, 1824
Cost function, 125, 296–297, 700, 887–888, 1038, 1058–1060, 1062, 1066, 1068, 1070–1072, 1078, 1090–1091, 1152, 1155, 1577, 1636, 1731
Coupling effect, 1384
Covariance matrix, 310, 312, 526–528, 531, 535, 545, 547, 549, 1404, 1491, 1607, 1611–1613, 1707, 1715
COVQ, 609, 875–876
Cramér-Rao bound, 1490
Bayesian, 1608
Cranial nerves, 1876
Cricothyroid muscle, 1846
Critical angle, 1676–1675, 1684, 1687
Critical band, 229, 233, 235, 255
filterbank, 254–255, 501
Critical bandwidth, 122, 233–234, 236–237, 840
Critical distance, 1393
Critical ratio, 1859
Cross-channel effects, 1217
Cross-correlation
factors, 272
function, 44–46, 94, 178, 495, 957, 959–961, 972–973, 1488, 1491
generalized, 904
Crossfading, 725
Crossover network, 410
Cross-spectrum, 46, 69
Curie point, 1784
Curve-fitting method, 951
Cut-and-spliced, 727
C-weightings, 109, 110
Cyclic redundancy check (CRC), 608, 870
Cyclostationary processes, 10
Cymbal, 1804, 1806–1807
Damping, 41, 169, 200, 292, 401, 406, 411, 429–430, 624, 636, 641, 1141, 1147, 1167, 1175–1176, 1178, 1252, 1254, 1257, 1267, 1270–1272, 1276–1277, 1391, 1635, 1637–1638, 1640, 1645, 1648, 1786
filter, 406
material loss, 636
Dark sound, 421
DAT format, 704–706
Data acquisition, 17, 25
Data compression, 720, 854, 1900
Data matrix, 1560–1561
Data validation, 31
Data window, 26, 40, 53, 56–60, 62, 898, 1494
See also window
Day-night average sound level (DNL), 1006–1014, 1017–1018
Decimation, 21
Decision process, 218
De-clicking, 776
Decoders, APP, 874
Decoding, iterative source-channel, 874
Deconvolution
Bayesian, 1578
blind, 1575
deterministic, 1572–1575, 1580
gapped, 1577
homomorphic, 1579, 1581
minimum entropy, 1578
MISO, 1610
predictive, 1577–1578
spiking, 1577
statistical, 1575–1576, 1580–1581
surface-consistent, 1579–1581
Wiener-Levinson, 1574–1575
Deep sound channel, 1671, 1673
Degradation category rating (DCR), 500, 833, 836–838
Degree of modulation, 264–265, 271–272
De-hissing, 778–780
Delay estimation technique, 494–495
Delay variation, 652, 659–660
Delays, 672, 786–787, 1049
Delay-sensitive responses, 1839
Delta-Sigma (DS), 19, 21–22, 24
DEMON, 1773
Denoising, 49, 779, 792
Depth migration, 1589
Design frequency, 1764–1765, 1767
Detection
outlier, 1615
parameter, 1740
probability, 1749, 1754
threshold, 234, 237, 804, 1218, 1190–1194, 1235, 1429, 1744–1745, 1751, 1753–1754, 1761, 1769, 1771, 1901
Detector
energy, 236, 1699, 1768, 1772
optimal, 185–186
Deterministic signal, 4–5, 34, 69, 72, 73–74, 80
De-thumping, 775–776
DFT, 37–39, 53–54, 136–138, 280, 457–461, 779, 884, 888–890, 892, 899, 904–905, 1132, 1368, 1774
Diabetes, 1219
Diaphragm
acoustic impedance, 1258–1259
damping, 1254, 1260
displacement, 1249, 1261
protection grids, 1302
radiation, 637
tension, 641
Dichotic pitch, 248
Dielectric constant, 1313–1315, 1339, 1341, 1343, 1785, 1788
Differential micro-phone, 209
Diffraction, 637, 679, 765, 1269
Diffraction kernels, 1647
Diffuse field equalization, 764
Diffuse sound field method, 936–938, 945–946
Diffuse zone, 1631
Diffuse-field, 1255, 1294, 1300–1301, 1303
Digital analyzers, 119
Digital audio
broadcast, 719, 800
editing, 723
restoration, 774
stationary head, 705–706
Digital compact cassette (DCC), 800
Digital filter, 368
Digital formats, 706, 718
Digital processing schemes, 210
Digital recorders, 704, 1863–1865

Digital recording, 22, 187, 323, 704, 1858, 1863–1864, 1899–1900
Digital signal processing (DSP), 3, 111, 205, 209, 400, 457, 558, 651, 655, 658, 794, 885, 984, 1759,
Digital-to-analog converter, 20, 560, 1772
Digital versatile disk (DVD), 720–723
Digital waveguide, 368, 400–404, 766–768
Dijkgraaf, 1835
Diphones, 560–563, 572
Direct stream transfer (DST), 722
Direction of arrival, 286, 1403, 1498, 1773
  estimation, 1402–1405, 1407, 1774
Directional filtering, 763, 765, 1655
Directional hearing, 1872, 1905
Directionality, 1278, 1281, 1284, 1761, 1867, 1873–1874, 1894
  figure-of-eight, 1278–1279
Directivity index, 694, 697, 763, 1744, 1761, 1823
Directivity patterns, 1279–1280, 1411
Disc of uncertainty, 1089
Discontinuous transmission (DTX), 603, 606, 828, 919
Discrete cosine transform (DCT), 802
Discrete fourier transform (DFT), 37–39, 53–54, 136–138, 280, 457–461, 778–779, 884, 888–890, 892, 899, 904–905, 1132, 1368, 1774
Discrete measurements, 136
Discrimination limens, 1871
Dispersion error, 767–768
Displacement overload, 660
Dissonance, 263
Distance measure, 308, 500, 503–506, 508, 510, 580
Distortion, 201
  components, 72
  emissions, 164
  peaks, 74
  product, 72, 74–75, 78, 163–165, 236
Distributed systems, 14
Disturbance, 812, 1015, 1082, 1091
  density, 510
  index (DIX) model, 813
  rejection, 1082
DIVA auralization system, 762
DOA *see* Direction of arrival
Dolby digital surround, 720
Dolph-Chebyshev window, 58–59, 696
Dolphins, 1897, 1899–1901, 1903–1905
Domain transfer functions, 1098
Doppler effect, 789, 791, 967, 1332, 1778, 1839, 1841
Doppler shift, 1330, 1332–1333, 1735, 1741–1742, 1777, 1841–1842
Doppler vibrometer, 1179, 1402–1404
Dorsomedullary nucleus, 1868
Double-modulator/single-carrier (DMSC) FM, 389
Double-talk detection, 890
DPOAE fine structure, 165
DPOAE I/O functions, 164–165
DSD-IFF File Format, 716
3-D triangular surveying, 1408
Dual channel FFT analysis, 69
Dual resonance nonlinear (DRNL), 177
Duplex theory of pitch perception, 183
Dynamic linearity, 1308–1309
Dynamic processing, 793–794
Dynamic range, 136, 1249, 1260, 1264
  compression, 205–206
Dynamic thermal model, 634
Dynamic time warping (DTW), 320
Dynamic tortuosity, 1168
Dynamite, 1546, 1548, 1554

EAAC+, 600–601
Ear
  bone, 1866
  defender, 169, 1094, 1097
  dermis, 1186
  epidermis, 1186
  pinna, 147
  stapes, 147, 150–151, 1866
  tympanum, 149, 1866–1867
  VIIIth cranial nerve, 1868
Eardrum vibration, 1873
Earth
  anisotropy, 1529
  composition and structure, 1523
  seismic wave attenuation, 1530
  velocity, 1528
EC theory, 182
Echo
  cancellation, 825–828, 844, 883–885, 887, 891, 897, 904, 909–912, 916–919, 923–924
  canceller, 825, 828, 832, 836–837, 844, 913, 916–920, 922–924
  colors, 1840
  free connection, 826
  loss, 824, 844–845, 916–917, 919, 923
  path models, 890–891
  suppression, 883–884, 891, 912, 916
Echolocation, 1903, 1905
  performance, 1838–1844
Eckart filter, 1492
Edit decision list (EDL), 717, 725, 727
Edit point, 724–728
Editing digital tape recordings, 706
Editing modes, 727–728
  effects, 236
EFR-GSM, 591, 603
Eigenfrequencies, 421, 433, 436, 438–442, 1382–1385, 1388, 1391
  distribution, 1386
  spacing, 1383
Eigenvalue spread, 888, 1047, 1051, 1060–1061, 1064, 1074
  adaptive decomposition, 904
Eikonal equation, 1441, 1443, 1445–1456, 1504
Electrical analog circuits, 1874
Electrical stimulators, 1236–1237, 1240
Electroacoustic analogies, 176
Electrodynamic shaker, 1143, 1145, 1148
Electroencephalography (EEG), 167–168, 738–739, 780
Electro-mechanical coupling, 1253
Elicitation, 1598, 1603–1605
EM algorithm, EM, 291, 521, 525, 530

Embedded Codec, 604–606
Embouchure control, 411
Emission sound pressure level, 1126
Empirical models, 1501, 1511–1512, 1594
Endfire, 697, 1763, 1767
Energy spectral density (ESD), 34, 42
Energy splitting loss, 1747
Entropy-satisfying upwind-differencing scheme, 1456
Envelope extraction, 255
Envelope fluctuations, 265
Equal temperament, 289, 333, 337, 346
Equalization cancellation (EC), 179
Equal-loudness contours, 251–252, 1194
Equivalent circuit model, 1798
Equivalent rectangular bandwidth (ERB), 235–236, 255
Error sensor, 1024–1026, 1030, 1032, 1061
ESPRIT, 1490, 1774
Estimation
  algorithms, 294, 296, 485, 506
  parametric, 1699
Estimator, 1610–1611, 1616–1617, 1700, 1703
  perception-based, 499–500
  robust, 1610, 1614
Euler's equation, 1110, 1116, 1282
Euler-Lagrange equation, 1442
European Broadcasting Union (EBU), 715
Eustachian tube, 149–150, 1867
Evanescent
  filter, 1134
  waves, 133–135, 1129, 1131, 1133, 1136
Evoked response audiometry (ERA), 156, 166
EVRC, 607
Excitation
  components, 476
  pattern, 154, 182, 238, 254, 268–270, 1893
  -plus-resonance (EpR), 367
  -resonance model, 361
  signal, 66–67, 69, 71–75, 78–81, 87, 102, 362, 364, 366, 451, 459, 464, 469, 471, 475–476, 478, 480, 572, 575–576, 588–589, 855, 869
  spectra, 270
Expectation-maximization (EM), 290, 521–522
Extratympanic sensitivity, 1873
Eyring formula, 680

F0 estimation, 278
  multiple, 284
Fast field program, 1680
Fast Fourier Transform (FFT), 26, 33, 38, 53, 67, 88, 99, 377, 805, 1130, 1765, 1900
Fast MLS Transform (FMT), 97
Fast Transversal Filters (FTF), 888
Feature vector, 305, 307–313, 316, 318, 320, 323–325, 328, 521, 526, 528
Federal Aviation Administration, 1017
Federal Interagency Committee on Noise (FICON), 1008
Feedback
  amplifier, 408
  comb and allpass filters, 794
  control systems, 1081
  controller, 1026, 1048, 1051–1052, 1054, 1082–1083, 1098–1103
  FM, 390, 393–394
  linear resonant controller, 1054
  loops, 185
  nonlinear, 408
  reduction, 684, 686–687
Feedforward, 1024–1025, 1029–1030, 1032–1033, 1044, 1051–1054, 1081–1082, 1092, 1099–1103
Feedforward comb filter, 788
Female response call, 1875
Fermat's principle of stationarity, 1442
Ferroelectric, 1313–1315, 1317, 1783–1784, 1788–1789
Field-effect transistor (FET), 1251
Filters
  allpass, 401, 406, 786, 790, 794, 892
  crossover, 651
  digital design, 401
  dispersion, 405–406
  elliptical, 23
  exponential, 135
  extended multidelay, 890
  FIR, 116, 592, 654–655, 766, 888, 1042, 1051
  high-slope analogue crossover, 654
  in-phase/quadrature, 121
  inverse, 67
  Kalman, 978, 1701, 1705, 1708–1709, 1778
  lattice digital, 403
  linear phase, 120, 406, 654, 660
  loop, 406
  matched, 49, 67, 157, 185, 775, 777, 1699, 1738–1744, 1869
  mel, 308
  multidelay, 889
  Octave- and fractional-octave-band, 112–113
  perceptual, masking, 122
  pluck, 409
  polynomial smoothing, 776
  prediction-error, 1641
  pure-state, 1492, 1494
  synthesis, 572, 588–589, 592, 597, 599, 609, 865, 888
  Tikhonov, 135
  transversal, 125
  waveguide digital, 368
  See also Finite Impulse Response (FIR) filters
Filterbank, 156–157, 179, 186–187, 206, 238, 254–255, 270, 501–502, 504, 805, 809, 865, 888–889, 1366, 1370, 1372, 1840, 1842, 1893
  nonlinear, 238
Filtered-reference LMS, 1044–1045, 1047–1049, 1051
Filtering, statistical techniques, 209
Filters and delays, 786
Filters discriminate signals, 119
Finite and boundary element methods, 662
Finite difference (FD), 401, 424, 1284, 1456
  approximation, 1117–1118
  error, 1117
  formulation, 427
  time domain, 401, 762

Finite element, 151, 363, 369, 419, 424–425, 429, 628, 638, 762, 1154, 1388, 1676–1677, 1798
Finite impulse response (FIR), 120, 125, 176, 654–655, 885
Fish vocalizations, 1859, 1857
Fisher F-statistic, 1466, 1492
Fisher information, 1607, 1609
Flanger, 789
Flanking bands, 234, 237
Flanking maskers, 237
Flat cardioid characteristics, 1282
Flow resistivity, 1168–1169
Flow-induced noise, 1822
Fluctuation strength, 263
Fluid-filled (audio) horns, 638
Flute-like instruments, 428
Flutter tonguing, 333, 342
FM slide, 1741–1742
FM synthesis, 377, 385, 389, 391, 394
Forced-choice discrimination test, 811
Formant, 361–362, 364–365, 369–370, 387, 390–394, 449–451, 458–461, 463–464, 466, 468, 472, 474–475, 478, 480–481, 499, 508, 560, 566, 574–576, 578–579, 597–599, 855, 862, 1237
  FM, 387, 392–394
  Singer's, 361
  vocoder, 364, 369
  wave functions, 369
Forward error correction (FEC), 604, 608, 613
Forward problem, 1642, 1724, 1726, 1728
Forward propagators, 1130
Four-channel model, 1202–1204
Fourier series, 5, 36–38, 54, 57
Fourier synthesis, 1510, 1682, 1688
Fourier-t transformation (FTT), 254
Four-point microphone method, 1408
  arrangement, 1408
  correlation, 1410
  directivity patterns, 1411
  short-time intensity, 1410
  virtual image, 1409–1410
Fractional-octave-band analysis, 116–117
Frame erasure concealment, 606, 608, 610–611, 613
Frame loss recovery, 611
Frame of reference, 220–221, 227, 1444
Free-field, 936, 1255, 1294
  comparison, 1307
  condition, 200, 662, 993, 952, 1115, 1268
  corrections, 1300–1301, 1307
Free-nerve endings, 1189–1190
Frequency bands, 44, 81, 108, 142, 154, 156, 181, 201, 209, 237–238, 486, 649, 838, 1404, 1410, 1428–1429, 1493, 1868–1870
  broadening of, 201
Frequency chirp, 29
Frequency-dependent amplitude, 84, 1683
Frequency-dependent group delay, 71
Frequency-domain
  adaptive filtering, 889
  calibration, 26, 29
  coders, 866
  coding, 865
  LMS algorithm, 127
  models, 289, 291
Frequency-independent delay, 120, 766
Frequency mapping, 843
Frequency modulation matching, 391
Frequency response, 655, 1252
Frequency scaling, 323, 450
Frequency subband filtering, 888
Frequency tracker, 1778
Frequency warping, 768
Frequency weighting, 107, 109
Fresnel zone, 1647
Frogs, 1861–1863, 1866, 1870–1872, 1875–1878, 1880
  earless, 1870
Full wave modelling, 1440
Functional magnetic resonance imaging (fMRI), 168–170, 738–740

G.722.1, 601–602
G.729, 593, 605–606, 611
G.729.1, 605–606, 611
Gain-before-feedback boost, 684, 686
Gain error, 73
Gain estimation, 494
Gain margin, 207, 1088, 1090, 1096–1097
Gamelan instruments, 347
Gamma distribution, 1599–1600
Gamma-tone, 1877
Gap detection, 185, 1198, 1234–1235
Gauss elimination algorithm, 427
Gauss's theorem, 1110
Gaussian
  beam model, 1505
  distribution, 289–290, 310, 527–528, 1595, 1599–1601, 1614, 1725, 1747
  error, 1725
  mixture, 524, 531–532, 550
  models, 521–522, 531, 1613
  noise, 26–27, 29, 1611, 1697, 1702, 1705
  random fields theory, 170
  sparse, 522, 524, 527, 529
  windowed sinusoids, 50
Gaussianity, 1578
Gauss-Markov model, 1705, 1709
Gene expression, 1870
Generalized sidelobe canceller, 698, 903
Generalized sweep synthesis, 81
Genetic algorithm, 380, 382, 385, 391–393, 1152, 1158, 1640, 1642, 1731–1732
Geodynamo, 1524
Geometric diffraction theory, 638
Geometric modeling of musical instruments, 424
Geometric models, 1503
Geometrical dispersion, 1684
Geophones, 1547, 1550
Georges' model equation, 1444
Geotomography, 1640
Ghost image, 142
Ghost reflections, 1576
Ghost source, 142–143
Gibbs sampling, 1729
Glabrous skin, 1188
Glasberg and Moore's formula, 235
Global propagation, 1683, 1702
Glottal
  closure, 560, 567–568, 571
  excitation, 366, 368, 570, 575–576
  formant, 576, 578–579

pulses, 365, 450, 459, 462–464, 469, 471–472
source, 362
Golay codes, 87–88, 103
Gold sequence, 87, 100–102
See also Kasami sequence
Gonadotrophin releasing hormone, 1877
Gong morphing, 333, 341
Gradient noise amplification, 127
Graphical models, 521, 524, 527, 551
Grating lobes, 142, 1489, 1714, 1766
Green's function, 14, 1133–1134, 1541, 1678
Griffin, 1224–1226, 1835
Ground loops, 18
Group additive synthesis, 382
Group delay, 66
Guitar, 313, 315–316, 401, 407–409, 420–424, 436, 438–440, 669, 788–789
body, 409, 421, 423, 436, 438–440
electric, 407–408

H_2/H_∞ control, 1091, 1101
Habituation, 224
Hadamard transform-Based FMT, 98
Hair cells, 281
Hair-follicle network, 1190
Half-wave rectification (HWR), 281
Hamiltonian equations, 1439, 1441–1443, 1445–1446, 1456
Hamiltonian treatment, 1441–1443
Hammersley-Clifford theorem, 527
Hamming, 56, 377, 457–458, 502, 513, 609
Hamming windowed, 458, 502
Hand-arm vibration, 1222–1223
Hankel function, 1136, 1506, 1508, 1678
Hann window, 56, 496, 1877
Hanning, 41
Hannon album, 221
Harmonic coding (HC), 866
Harmonic complex tones, 183, 234, 247
Harmonic distortion, 74, 76, 663, 811, 1261
Harmonic impulse responses, 75
Harmonic noise model (HNM), 569
Harmonic vector-excited coder (HVXC), 877
HARPA ray tracing computer program, 1504, 1438
Harvey Fletcher, 486
Head-related transfer function (HRTF), 87, 90, 160, 747, 755, 764, 766
Hearing aids, digital, 176, 187, 205
Hearing impairment, 201, 206
listeners, models for, 255
Hearing model, 500
Hearing sensation, 251, 254, 260, 263
Hearing, threshold, 166, 168, 245, 272, 501, 504, 1260, 1293, 1309
Heisenberg uncertainty principle, 35, 48
Helmholtz equation, 132, 767, 1506, 1508
Hemi-free-field method, 934–936, 943–944
Heschl's gyrus, 171
Hidden Markov model, 311, 297, 311–312, 324, 328, 371, 522, 525, 550, 580
High-definition television (HDTV), 800
Histogram method, 1603–1604
Hit and false alarm rates, 218
Hologram surface, 1130, 1135
Holography
cylindrical nearfield acoustical, 1133
Nearfield acoustical (NAH), 131, 1129–1137
patch, 137
planar nearfield acoustical, 131, 1130
Homomorphic vocoder, 365
Hooke's law, 1536
Horizontal wind model (HWM), 1501
Hot-wire anemometers, 1284
Human auditory system, 499, 1900
Human segmentation, 327
Husebye's method, 1426
Hybrid adaptive procedure, 218
Hybrid feedback controllers, 1102
Hybrid methods, 780, 1732
Hydraulic actuator, 1142, 1150
Hydrophone, 1547, 1550–1552, 1581, 1685, 1711, 1758, 1760, 1767, 1851, 1857, 1862–1863, 1899, 1901–1903
Hypercomplex, 1869, 1870
Hypoglossal, 1876

IEC-standard timecode, 705
Image source method, 762, 765
Impedance
adjustment, 150, 431
audiometry, 150
tube, 765, 1179
Impulse response, 12–14, 22, 35–36, 47, 49, 55, 65–66, 74–75, 79, 88, 90, 93, 96–99, 102–103, 116, 120, 128, 160, 176–177, 238, 406, 409, 440–441, 444, 468, 475, 654–655, 657, 662–663, 679–680, 683, 692–693, 758, 765–765, 789, 794, 885–887, 890, 904–905, 957, 959–960, 1032, 1034–1035, 1043–1044, 1048, 1051, 1366, 1391, 1395, 1398–1400, 1402, 1408–1410, 1412, 1417, 1483, 1488, 1561, 1564–1567
measurements, 103
Incompressibility condition, 429
Independent component analysis (ICA), 292–293, 295, 298
Independent subspace analysis (ISA), 295
Index of refraction, 1506, 1670
Indonesian gamelan, 333, 335, 347, 420
Inertial actuator, 1143
Infinite impulse response (IIR), 654
Informational masking, 237
Infrasound
array data processing, 1487
from earthquakes, 1432–1435
noise-reducing pipe arrays, 1473
phase identification, 1464
signals, 1432, 1462, 1470, 1498, 1502, 1513
travel time tables, 1452, 1469
waves, propagation of, 1497
from volcanoes, 1430–1431

Infrasound data processing at IDC, 1464–1467
  network processing, 1463–1464, 1466
  reviewed event bulletin, 1463
  standard event list, 1463
  station processing, 1463–1465, 1467–1468
Inharmonicity, 7, 278, 290, 380, 420–421, 438, 560–561
Inhibitory network, 183
Initial transient, 420–422, 424, 436, 438–440
Inner and outer hair cells, 151
Inner ear, 147, 150–151, 153–154, 159, 161–162, 164, 215, 281, 451, 798, 856, 1840, 1867, 1869–1870, 1904
  oval window, 150–151, 159, 164, 176, 1867
Input-output function, 184
Instantaneous frequency, 1370
Instantaneous impression, 221–224, 226
Instantaneous judgments, 222–225
Instrumental clustering in families, 420
Integer-coded strings, 1160
Integrated octave band analyzer, 118
Intelligibility function, 483, 197, 199–201
Interactive music, 338
Interactive processing, 1469–1470
Interaural
  cross-correlation function, 178
  intensity difference, 182
  level difference, 748, 750, 1873
  phase difference, 237
  time difference, 178–179, 182, 748, 750, 766, 1873
Interleaved pulse positions (IPP), 590
Internal inflammation, 1226
Internal model control, 1099–1100
Internal representation, 157, 185–186, 813, 843
Internal stability, 1086
International Electrotechnical Commission (IEC), 1293
International microphone standards, 1262
Interpolation schemes, 13
Intersexual selection, 1878
Invariance property, 1610
Inverse problem, 1300, 1580, 1629, 1640, 1642, 1696, 1723–1726, 1728–1729
Inverse repeat sequences, 93, 100
Inverse-vocal-tract, 862
Inversion, iterative, 1639, 1642
Irradiated wave, 1332
IRS filtering, 592
Isolated large amplitude, 1465

Jew's harp, 246
Jitter, 560, 580, 612, 826, 917
Judgement Model, 500
Just intonation, 333, 344–346, 349

Kaiser and Gaussian windows, 58
Kaiser-Bessel, 41, 811
Kalman filter, 978, 1701, 1705, 1708–1709, 1778
Kandinsky tests, 834, 837
Karaoke, 367, 369
Kasami sequence, 87, 97, 100–102
Kelly-Lochbaum model, 368, 403
Kernel methods, 523, 1604
Kirchhoff migration, 1589
KTH rule, 370
Kullback-Leibler (KL), 310, 574–575, 1605

Lamb wave, 1438, 1500
Lamé constant, 1526
Laminar flow, 429
Laplace-domain transfer function, 35, 1032, 1030–1031, 1083, 1098
Laryngeal, 1846, 1876–1877
Larynx, 360–361, 855, 1845–1846, 1876
Laser Doppler method, 1330
Laser Doppler Velocimeter, 1329–1338
Laser Doppler Vibrometers (LDV)
  applications, 1337–1338
  construction, 1333–1336
  features, 1331–1332
  one-dimensional, 1333–1335
  operation principle, 1332
  three-dimensional, 1335–1336
Late auditory evoked potentials (LAEP), 166–167
Late packet processing, 613
Lateral inhibition, 1873
Lateral suppression, 238
Lateralization, 178, 242, 244
Leakage, 22, 55, 141
  wavenumber 137
Leaky precursors, 1689
Least absolute deviation (LAD), 1612
Least squares
  damped, 1636
  recursive (RLS), 126–127, 887–888
Legendre sequence (LGS), 87, 100, 102–103
L-estimators, 1614, 1615
Level recorder, 225
Likelihood function, 287, 530, 550–551, 1601, 1605, 1607–1608, 1610, 1699, 1701, 1703, 1726, 1729
Liljencrants-Fant (LF) model, 363
Line spectral pair (LSP), 480
Line spectrum frequencies (LSFs), 574
Linear
  acoustic wave propagation, 1499
  chirps, 63
  deconvolution, 74
  distortion, 246, 408,
  phase filters, 120, 406, 654, 660
  prediction PSOLA (LP-PSOLA), 571
  predictive coding (LPC), 367, 449–450, 464, 466, 469–470, 480–481, 588
  resonant feedback controllers, 1054
  source-filter model, 361
  wave
    equation, 1500
    theory, 1440
Linearization, 623
Linearly constrained minimum variance (LCMV), 698, 903
Linkwitz-Riley functions, 652–651
Lip reading, 206, 1239, 1241
Lippmann-Schwinger equation, 1590
Listening
  -only tests, 837
  situation, 822–823, 839, 842
  tests, 654, 664, 751, 804, 833, 836–837, 868
Lithosphere, 1525
Lloyd-Max algorithm, 454

LMS algorithm, 126–127, 887, 1025, 1030, 1033, 1036, 1038, 1043–1051
Localization
  collapsing method, 1631
  double difference method, 1631
  Geiger's method, 1629, 1631–1632
  grid search, 1630
  inside-head, 755
  joint hypocenter determination (JHD), 1630–1632
Lofargram, 1769–1770
Log spectral error (LSE), 498
Logarithmic frequency scale, 33, 289–290, 323
Logistic function, 508
Log-lag correlogram model, 283
Lombard effect, 910
Long baseline, 1902
Long-term spectrum, 236
Loop gain, 153, 1084–1085, 1092–1094, 1096
Loss mechanisms, 1673–1675
Loudness, 162, 251, 504, 823
  adaptation, 224
  dynamic model, 254
  Equal contours, 251–252, 1194
  fluctuation, 263
  models of, 182
  overall of instrument, 422
  overall rating (OLR), 842, 849
  pattern, 182, 254
  rating, 838, 841–842, 845, 916–917
  rating, sending (SLR), 842
  scaling, 165, 252, 258, 843
  specific, 182, 254, 257–258, 813
  summation, 259
  transformation, 256–258
  Zwicker model of, 254, 257
  See also Bank loudness spectra
Loudspeaker
  active, 650–651, 655, 658
  arrays, 758
  design, 649–664
  direct-radiating, 635, 637
  distributed-mode, 638
  dodecahedron, 940
  dynamic, 624, 631–632, 640–641, 1349–1350, 1362
  electrostatic, 1354–1356
    composition, 1354–1355
    transducer efficiency, 1355–1356
  enclosure, 626, 652
  horn, 638
  layouts, 748–750
  passive, 649–650
  performance evaluation, 649–664
  superconductive, 1351–1354
    composition, 1351–1353
    equivalent circuit, 1353
    properties, 1351
  theory developments, 625
  thermal model, 635
  time alignment, 684
  transmission line, 634
Love wave, 1527, 1538–1539
Low-noise-treated cabling, 1319, 1321
LSQR, 1640
Lyapunov equation, 533

Mach number, 1821, 1827
Magnetic bias, 1147
Magnetic induction, 17
Magnetic resonance images (MRI), 168–170, 369
Magnetoencephalography (MEG), 167–168, 738–739
Magnitude estimation, 219, 254, 1195, 1197
Magnitude response, 66, 78, 501, 649, 651–652, 656, 695, 1091, 1390–1391, 1399
Mahalanobis distance, 310, 574–575
Mainlobe, 141–143, 695–695, 902
Male aggressive behavior, 1878
Markov Chain Monte Carlo (MCMC), 291
Markov models, 297, 311, 328, 521–522, 526, 528, 550
Maskers, flanking, 237
Masking, 122, 181, 201, 229–232, 234, 236–238, 256, 1850
  backward, 185, 233, 803, 808
  binaural level difference, 237
  central effects, 237
  comodulation release, 237
  forward and backward, 185, 803
  forward, 232
  frequency, 122
  growth functions, 230
  informational, 237
  partial, 229, 866
  psychoacoustic models, 807
  simultaneous, 230, 232, 501, 504, 803, 806
  spectral, 229, 256, 258
  temporal, 122, 187, 236, 255–256
  tone-on-tone patterns, 231–232
  two-tone, 234, 237
  upward spread of, 230–233
Mass Spectrometer-Incoherent Scatter Radar Model (MSIS), 1501
Mass-spring-damper system, 633
Matched field inversion, 1729
Matching pursuit algorithm, 286, 1403–1405, 1417
Mate attraction, 1878
Material absorption, 765
Mating-call detector, 1869
Maximum a posteriori (MAP), 536–537, 1613, 1729
Maximum connect set (MCS), 96–97, 102
Maximum length sequences (MLS), 67, 87, 90–91, 663, 683
Maximum likelihood
  estimator, 1610
  linear regression, 525, 547
  sequential procedure, 217
  training, 525, 529–531
Maxwellian optics system, 1193
MBROLA, 565
McGurk effect, 734–738
Mean listening quality-opinion score (MOS), 833, 843
Mean opinion score, degradation, 833
Mean-squared-error (MSE), 898, 902, 1575
Measurement equation, 1704–1705, 1708–1709
Measuring normalizing block (MNB), 505–506
Mechanical stimulators, 1236
Mechanical-acoustical transformation, 631
Mechano-acoustical conversion, 629
Mechano-electric transduction, 1207
Mechanoreceptor, 1187, 1189, 1207
Media Exchange Format (MXF), 717–718
Meissner corpuscle, 1189–1190, 1192, 1205
Mel filters, 308

Mel-frequency cepstral coefficient (MFCC), 307–309, 328, 365–366, 480, 522, 525, 574, 578
Melodic, 318
Mel-scale cepstral feature, 188
Meridian lossless packing (MLP), 722
Merkel-cell neurite complex, 1189, 1204–1205
Mesh Topologies, 767–768
M-estimators, 1615
Metabolism, 1878
Method of moments, 1599, 1603–1604
Metropolis algorithm, 1731
Micro electrical mechanical systems (MEMS), 1331
Microflown, 1116, 1120, 1284, 1286, 1288–1290
Microflown intensity probe, 1121–1122, 1288–1290
Micro-mechanical devices, 1264
Microphone
  accurate theory, 640
  arrays, 691, 696, 758, 901–912, 1125, 1263, 1281
  calibration, 1293–1311
  condenser, 632, 640–641, 987–988, 1116, 1120, 1247, 1254, 1262, 1284, 1286, 1288, 1295, 1299, 1308–1309, 1862
  dipole, 756–757
  directional, 207, 594, 686, 757–758, 912, 1275
  dynamic, 641–642, 1267–1268, 1271–1272
  electret, 1249
  interference, 1280
  laboratory standard, 1295, 1297, 1300, 1302–1303, 1307
  measurement, 160, 687–688, 1247, 1249, 1252, 1257, 1263–1264, 1295, 1303
  phase-matched pressure, 1284
  pressure, 1275
  pressure-sensing condenser, 1247
  pressure-sensitive omnidirectional, 1279
  ribbon, 631, 671, 1278–1279
  techniques, 671, 756–758
  working standard, 1302
Middle ear, 147
Middle ear input impedance, 160
Middle latency auditory evoked potentials (MAEP), 166
MIDI, 277, 297, 333, 335–339, 370, 712
MIDI dance floor, 338
Midrange, 649–650
Millington-Sette formula, 680
Minidisc, 719
Minidisk recorders, 1864, 1865
Minimum mean squared error (MMSE), 698, 873, 1609, 1613
Minimum phase zero, 1388–1390, 1394, 1396–1398
Minimum variance distorsionless response, 1717
Misadjustment, 126
Mismatch problem, 1709–1710
Mismatched filter, 67
Mixed excitation linear prediction (MELP), 480, 607, 866
Modal equalization, 659
Mode projector, 1804–1805, 1808
Model parameter, 289–290, 369, 371, 465, 492, 536, 661, 1603, 1616, 1636–1637, 1641–1642, 1710, 1723–1733, 1844
Model selection, 523, 636, 1594, 1605, 1617
Model-based processing, 1696
Modeling pitch, 241–242
Modelling error, 1043, 1047–1048, 1073–1075, 1709
Modified discrete cosine transform (MDCT), 600, 802, 809, 811
Modulation filterbank, 179, 186–187
  frequency, 154, 264
  index, 385–387, 389–392, 489–490
  transfer functions, 180, 488, 492
Modulation, 179, 185, 264, 322, 385, 391, 959
Moessbauer method, 215
Mohorovicic discontinuity, 1525, 1527
Moment generating function, 1602
Monitor wedges, 670, 676, 686
Monophonic, 279
Monte Carlo method, 1642
Motion sickness, 1224, 1226
Motoneurons, 1875–1876
Motor nuclei, 1876
Motorneuron map, 1875
Mountain associated waves (MAW), 1430
Moving coil, 1267
MP3, 122, 187, 614, 716, 808–810, 1864, 1895
MPEG audio file formats, 716
  MPEG-1, 804, 808–811
  MPEG-2, 720, 810–811
  MPEG-4, 803
MS stereo, 757
Multiband resynthesis OLA (MBROLA), 572
Multichannel techniques, 900–904
Multidimensional scaling, 220
Multigrid approaches, 1642
Multi-input multi-output, 13
Multimodel inference, 1606
Multi-path sound propagation, 961
Multiphonic playing technique, 435
Multiphonic sounds, 431–433, 436
Multiple scattering, 1564, 1568
Multiple sensory inputs, 1216
Multiplexed channel phase correction, 29–30
Multisensory neurons, 739
Multitone techniques, 663
Multi-variable binary string (MVBS), 1155, 1160
Murich segmentation, 309
Music
  segmentation, 310
  structure detection, 327
  summarization, 309
  summary, 324
  synthesis, 375, 385, 391
  thumbnails, 325–326
  transcription, 277–278, 283, 285, 291, 293, 295, 297–298, 328
Music and singing synthesis equipment (MUSSE), 364–365, 370
Music transcription, manual, 297
Musical instrument digital interface *see* MIDI
Musical instrument, structure of, 419, 422
Musical tones, 284, 421, 899
Musique Concrete, 350

Muting, 1587
Mutual information, 308, 905

Nagra-D recorder, 705
Narrowband, 461, 1600, 1768–1769, 1772, 1774, 1777, 1837, 1897, 1905
  detector, 1768–1772
  masking of, 232
  noise, 233, 248, 265, 866
  telephony, 492, 823
National metrology institutes (NMIs), 1296
NATO STANAG, 607
Naval Research Laboratory-Ground to Space Model (NRL-G2S), 1502
Navier-Stokes model, 429–430
Nearfield acoustical holography (NAH) *see* Holography
Negative binomial, 1597
Neighboring critical bands, 268
Neonatal screening tool, 165
Nested modulator FM, 392–394
Neural adaptation, 185
Neural hearing disorders, 156
Neural networks, 283, 371, 523, 524, 1606, 1072
Neural structures, 1188–1190
Neuralgia, 1220
Neurectomy, 1220
Neuronal filter bank, 1846
Neyman-Pearson detector, 1696
Night-time "penalty", 1006
Noise
  aircraft landing, 1014
  barriers, 103, 953, 958, 963
  cancellation, 162, 166, 169, 825, 827, 844, 918, 1024, 1032
    electrical, 1025, 1030, 1033, 1043, 1052
    self-adaptive (SANC), 128
  of car, 225
  control, 1167
  disturbance, 505
  electromagnetic, 19
  electronic, 17–18
  elimination, 984
  environmental, 216, 219–220, 222, 1006–1008, 1550
    annoyance, 1008–1016
    community reaction, 1018
  floor, 676
  induced annoyance, 1008–1012, 1014
  internal, 157, 257
  level, 27, 480, 486, 492, 592, 600, 807, 846, 910, 912–913, 963, 965, 1005–1008, 1015, 1018, 1097, 1260, 1264, 1429–1430, 1484, 1491, 1577, 1744, 1749, 1757–1769, 1821, 1850–1859, 1901
  mapping, 1124
  narrowband 233, 248, 265, 866
  notched, 234–236
  Nyquist (or Johnson), 1260
  periodic, 72
  pseudo, 1826–1828
  pseudo-random, 74–75, 80
  radiation problems, 131
  railroad, 1358–1362
  ranking of sources, 1124
  reduction, 206–209, 607, 779, 825–826, 832, 844, 846, 897–905, 910–925, 1124, 1465, 1474, 1482, 1485
    algorithms, 594–595, 847
    schemes, benefit of, 209
  road, 225
  self, 1759, 1821–1822
  shaping, 21, 599, 806, 810
  -to-Masked-Ratio (NMR) model, 812
  traffic, 222
  uncorrelated, 70–71, 268, 902, 1715
  uniform masking, 234
  white, 233
  wideband random, 248
Noisy channels, 852–877
Nonideal harmonicity, 292, 491, 774, 842
Nonlinear
  feedback, 408
  filterbank, 238
  effects, 236
  models, 636, 639
  progressive wave equation (NPE), 1513–1514
  reflection coefficient, 411
  transmission line model, 164
Non-minimum phase, 1033, 1092, 1388–1390, 1394–1398
Nonnegative matrix factorization (NMF), 292–298
Nonnegative sparse coding algorithm, 297
Non-parametric processing, 1746
Non-periodic Sweeps, 74
Non-stationary random signal, 10
Normal-Incidence Sound-Absorption, 950–951
Normalized Least Mean Square (NLMS) algorithm, 126–127, 887, 892
Numerical magnitude balance, 1195
Numerical solutions, 1508
Numerical weather models, 1502
Nyquist
  (or Johnson) noise, 1260
  criterion, 1087–1098, 1714, 1764
  frequency, 20–23, 54–55, 68, 377, 451, 1739, 1863–1864, 1900
  sampling, 1738

Objective audio signal evaluation (OASE) model, 812
Objective evaluation, 812, 832, 838
Objective function, 530, 1158, 1608, 1611, 1613, 1637, 1642, 1728–1732
Off-axis signals, 1899
Offset function, 92, 93
Omnidirectional, 1275
Open media framework interchange (OMFI), 717–718
Open-loop experiments, 1873
Operculum, 1867
Optimization problem, 1629–1630, 1725, 1728
Optohapt, 1233
Orthotelephonic reference position, 821, 823
Oscillation, 35, 243, 246, 394, 405, 429, 812, 1047, 1357, 1531
Ossicles, 150
Otic capsule, 1867
Otoacoustic emissions (OAE), 153, 162–165
Otoscopy, 147, 149
Overall impression, 220
Overlapping partials, 287
Overshoot, 236
Overtone series models, 288
Overtone singing, 246
Oxygen consumption, 1878

P waves, 1526–1528, 1533, 1538
Pacinian corpuscle (PC), 1188, 1192, 1217
Pacinian system, 1192, 1201, 1217
Packet based Transmission, 826, 916–918
Packet loss concealment, 611, 826, 916, 918
Packetization, 825–826, 844, 916–917
Packet-switched transmission, 877
Pair-wise panning, 751–753
Parabolic equation (PE), 1503, 1506, 1510, 1678–1677, 1678, 1681, 1683, 1730
Parameter sensitivity, 1733
Parametric processing, 1696
Paraxial approximation, 1444, 1647, 1681
Paraxial ray tracing, 1444–1446
Parkinsonism, spastic disorders, 1225
Parseval's theorem, 34, 108
Partial coherence, 139
Particle velocity sensor, 1282
Partitioned block frequency-domain adaptive filters (PBFDAF), 890
Passive radiators, 634, 659
Pattern theory, 1202–1203
Peak-picking algorithms, 474
Pendopos, 340, 348
Pentode, 794
Perceptual
  audio quality measure (PAQM), 813
  brightness, 391
  coding, 120, 122
  evaluation of speech quality (PESQ), 509–510, 514, 610, 824, 843, 848
  grouping rules, 287
  linear predictive (PLP), 480
  masking filters, 122
  models, 175, 206, 580, 602
  objective measurement (POM), 813
  speech quality measure (PSQM), 504
  transformation, 500, 503, 505–506, 510
Percussion transcription, 278, 295
Periodic
  autocorrelation functions (PACF), 88
  coded signals, 100
  cross-correlation, 93–94, 99
  noise, 72
  pseudo-random noise, 74
  signal, 5, 7, 10, 36–37, 243, 279, 567
  to aperiodic ratio (PAPR), 579
Periodotopic mapping, 156
Peripheral auditory, models of the, 281
Peripheral hearing model, 283–284
Permeability, 1168
Phase error, 30, 1042–1043, 1049–1051, 1116–1117, 1121, 1339, 1507, 1513
Phase linear filters, 119–120
Phase margin, 1088, 1090, 1093–1094, 1097
Phase response, 22–23, 108–109, 120, 178, 406, 651, 653, 655, 663, 765, 790, 1118, 1120, 1270, 1288, 1390–1391, 1397, 1399–11400, 1477, 1798, 1807, 1810
Phase vocoder, 353, 364, 377–378, 1365–1366
Phase-lag compensator, 1093, 1096–1098
Phaser, 800
Phon, 251
Phonation, 360–361, 575–576, 1892
Phoneme transitions, 371
Phoneme, 199, 298, 361–362, 370–371, 451, 456–457, 459–460, 462, 559–562, 574
Phonology, 1892–1893
Phonotaxis, 1872
Physical acoustics series, 625
Physiological mechanisms, 180, 216
Piano
  performance, 221, 283
  roll, 292, 319
  string, 7, 400–401, 406
Picket fence effect, 40–42, 54
PICOLA algorithm of Morita, 568
Piezoceramic accelerometers
  construction, 1316
  compression mode, 1318
  flexural mode, 1318
  high impedence charge, 1319–1321
  low impedence designs (IEPE), 1331–1322
  model analysis, 1322
  mounting
    stud, 1323
    screw, 1324
    adhesive, 1324
    magnetic, 1324
    shear mode, 1315–1316
  NVH, 1322
Piezo-ceramic, 1079, 1146, 1233, 1313–1314, 1316–1319, 1325, 1784, 1789–1794, 1804
  plates, 1148
  stacks, 1150
Piezo-composite, 1788, 1799, 1801, 1804–1805, 1807–1811, 1813–1814
Piezo-electric, 625, 632, 635, 671, 675, 1239, 1314–1317, 1319, 1322–1323, 1329–1330, 1547, 1783–1785, 1789–1790, 1807, 1809
  acceleration pickup, 1330
  crystal, 1316
  film, 1148
  materials, 1239, 1789
  polymer, 1786, 1788
  transducers, 625, 632, 642, 1547
Pipe array
  acoustic frequency response, 1477–1482
  acoustics of noise-reducing, 1474
  electrical equivalents for, 1477
  electro-acoustic model (for), 1477
  fundamental resonance frequency, 1480
  high frequency cut-off, 1482
  resonances, 1480
  response, 1474
  simulations, 1484
  transfer function for, 1478
Pistonphones, 24, 1302, 1310
Pitch
  bend, 349
  harmonics, 462, 473, 479
  height, 244
  manipulation, 369
  marks, 567
  mismatch, 571

modification, 566, 569, 571, 785, 789, 1365, 1379
pattern completion, 247
percept, 168, 183, 242–243, 246, 278, 281, 283
perception, 168, 183, 241–243, 278, 281–283
of random stimuli, 248
shifting, 790–791, 793
strength, 245
theory, 183, 241–242, 247
units, 244
Pitch-synchronous overlap-add (PSOLA), 369, 566
Plane-wave horn theory, 638
Playback, 197, 231, 336, 409, 484–485, 560, 587, 660, 677, 719, 773–775, 790, 814, 838, 1865, 1880, 1895, 1894
Plucked instrument, 420, 439–440
Plucked or struck stringed, 408
PODX (Poisson Distribution) system, 340
Point impedance models, 177
Point of subjective equality (PSE), 216
Poisson model, 1383
Polarization, 1249
Pole-zero diagram, 466–468
Polymer composites, 1809
Polyphonic, 277–279, 284, 287, 291, 297, 299, 306, 318, 341
Polyphonic transcription, 297
Pore radius, 1173
Pore shape, 1168
Pore walls, 1173
Poroelastic foams, 1178
Poroelastic media, 1175
Porosity, 1168
Porous materials, 1167
Porpoises, 1897, 1899, 1901, 1903–1905
Positron emission tomography (PET), 156, 168, 738–739
Post-location processing, 1470–1471
Postmasking, 232
Poststack time migration, 1588
Power spectral density (PSD), 27, 34, 42, 115, 495, 898
Power spectrum, 26–27, 45–46, 92, 102, 160, 248, 280, 289, 295, 298, 316, 475–476, 487–489, 501, 764, 778, 859, 1825, 1850, 1893
p-p method, 1116, 1118
Preamplifier, 81, 674, 1120, 1250–1251, 1259–1260, 1308, 1863
Predation risk, 1861, 1878, 1880
Prediction error, 465–466, 1577
Predictive quantization, 609
Pre-echo, 654, 804, 809
Preferred pairs, 95–96, 101
Premasking, 233
Preoptic nucleus, 1868
Preprocessing, 162, 175, 185, 187–188, 205, 295, 514, 559, 804, 1567, 1578
Pressure field, 1255, 1294
Pressure gradient, 631, 999–1000, 1276–1277, 1283–1284, 1851
Pressure-intensity index, 1118–1120
Pressure-residual intensity, 1119–1120
Prestack time migration, 1589
Prestress, 1147
Pretrigeminal nucleus, 1876–1877
Prewhitening factor, 1575, 1577
Primary calibration, 1294
Primary frequencies, 1804
Primary wave, 955, 1537
Primitive polynomial (PP), 91
Principal components analysis (PCA), 294–296, 380, 383
Principal co-ordinates, 1060, 1074
Probabilistic relational models (PRMs), 552
Probability density function (PDF), 10, 312, 452, 454, 873, 1602, 1604, 1610, 1729
Probability distribution,
Bernoulli, 1597
beta, 1598, 1600, 1603
binomial, 1597
chi-square, 1600
exponential, 1596, 1597
gamma, 1599
generalized beta, 1598
generalized lambda, 1598
geometric, 1597
Laplace, 1596
levy-stable, 1595–1596
lognormal, 1595, 1600
maximum entropy, 1599
Maxwell, 1600
Pareto, 1596
Poisson, 1596, 1597
power-law, 1596
Rayleigh, 1385, 1600, 1747
triangle, 1598
uniform, 1598, 1600
Weibull, 1597, 1601
Probe frequency, 232
Processing gain, 1741, 1746
Profile analysis, 237
Progressive multi channel correlation (PMCC), 1425–1435, 1465–1466, 1468, 1493–1494
Prominent portion, 224
Propagation
loss, 634, 949, 1754
modeling, 1434, 1498, 1503, 1679
models, 1503–1505, 1512–1513, 1515, 1729, 1731, 1754
phase, 1391
Propeller noise in aircraft, 1076–1079
Prosody, 558–559, 563–566, 577
modification, 558, 562–563, 566
transplantation, 566
Proximity effect, 1277
Proximo-distally Gaussian, 1207
Pseudo-periodic (voiced) signal, 591
Pseudo-random modulation, 491
Pseudo-random noise, 74–75, 80
Pseudo-reference speech signal, 514
Psychoacoustic analysis, 803–804, 810
Psychoacoustical model, 804
Psychoacoustical roughness, 263–273
Psychoacoustics, 215–227, 241, 801, 824
Psychophysical methods, 216–220, 223, 226
limitation of, 218–219
Psychophysical nomenclature, 1203
p-u sound measurement systems, 1116, 1120–1123
Public Switched Telephone Networks (PSTN), 825–827, 915–916, 918–919, 924
Pulse code modulation (PCM), 605, 704, 805, 854
Pulse-transfer function, 1031
Pythagorean tuning, 333, 345

Q values, 1844
$Q_{10\ dB}$, 1871
QCELP, 607
Quadruplex model, 1203
Quality chunk, 715
Quality factor, 1530, 1790, 1810
Quantile function, 1598–1599, 1602, 1604
Quantisation, 450–455, 480–481, 609, 704, 712, 805–806, 854
Quartz, 1314
Quasi-constant frequency, 1837
Quasi-free-field measurements, 78
Quasi-stationary Signals, 62
Questionnaires, 219

Radiation efficiency, 1123
Radiation impedance, 626, 637, 938
Random signal, 4
Range dependent, 1671, 1674–1677
Range estimation, 1737, 1741, 1743
Range resolution, 1737, 1741–1744
RASTI, 491, 677, 683
Rate of convergence, 126
Rating Method, 219
Ray
  based methods, 762, 1398
  model, 1443, 1503–1505, 1687, 1691
  parameter, 1443, 1447–1448, 1453
  theory, 1437, 1456–1457, 1503–1504
  tracing, 762, 794, 1386, 1437–1457, 1469, 1513, 1553, 1589, 1591, 1629, 1636, 1691
Rayleigh wave, 1179–1178, 1527–1526, 1538–1539
Razor blade cut-and-splice, 705
RCELP, 607–608
Reactive intensity, 1111, 1122
Real-time analyzers, 117
Receiver matrix, 1561–1562
Receiver Operating Characteristics, 1697, 1740
Reciprocal calibration, 641
Reciprocal interpretation, 1207–1208
Reciprocal pair, 95
Reciprocity technique, 1295, 1297–1299, 1303
Reconstruction error, 288, 297, 868
Reconstruction from Laser Projections, 1413, 1415–1417
Recording checklist, 1866
Recording gain, 1859, 1863
Recruitment, 153–154, 165, 206, 252, 257, 259, 1221
Recursive least squares (RLS), 126–127, 887–888
Reed, 279, 409–412, 420, 423, 431–432
Reel-rocking, 727–728
Reference measurement, 73
Reference signal, 125, 127, 129, 139, 495, 497–498, 500, 506, 508–509, 512, 688, 698, 847, 897, 904, 998, 1026, 1029, 1032, 1034–1037, 1040, 1043–1045, 1046–1049, 1051–1052, 1057, 1077, 1081, 1092, 1102
Reflected wave, 1332
Reflection
  emissions, 164
  free surface, 1542
  operator, 1562, 1567–1568
Reflection coefficient, 404, 411, 498–499, 951, 1538, 1563, 1576, 1595, 1598, 1674, 1725, 1733
Reflectivity estimation, 1568
Reflectivity, 1673
Reflex sympathetic dystrophy, 1220
Refractive index, 1413
Regularization, 134
Regular-pulse excitation (RPE), 860, 862
Relative amplitude spectral error, 381, 393
Release call, 1875
RELP coding, 858
Reoptimization of DOAs, 1406
Replica correlator, 1698, 1739
Residual echo suppression, 884, 913
Resonance chambers, 1268, 1271
Resonator, 176, 400, 408, 444, 1280, 1356–1358, 1790, 1792, 1868, 1876
Respiratory, 1185, 1876
Reverberation
  artificial, 1398
  control, 1399–1400
  phase, 1394–1395, 1397–1398
  spectrum, 1738, 1750
  strength, 1753
  surface, 1750
  time, 14, 66, 81, 116, 677–678, 766, 937, 1125, 1382
  water column, 1576
Revision of STI (STIr), 489
Reynolds averaging, 1824
Reynolds-Averaged Navier-Stokes (RANS), 430
Ridge regression, 536–538, 540, 549
RIFF (Resource Interchange File Format), 712–715
Robotization, 793
Robust stability, 1089–1091, 1094, 1099, 1101
Robust transmission, 856–857
Robust vector quantization, 609
Robustness, 126, 246, 284
Room equalization, automatic maintenance of, 687
Room equalization., 684
Room impulse response, 66, 90, 102–103, 677, 756, 794
Room response equalizers, 657–669
Room simulation, 786, 794
Rotary, 785, 788–789
Rotating diffuser, 103
Rotating reflector (diffuser), 938
Roughness, 254, 263, 1205
  of amplitude-modulated tones, 268, 272
  calculation, 265
  dependency, 271
  specific, 265
  threshold, 264, 268, 273
Ruffini/Paciniform end organ, 1189
Rules-based systems, 370, 560
Runge-Kutta four-step algorithm, 1444
Running autocorrelation function, 183
RWC Music Database, 327

S wave, 1526–1529, 1537–1538, 1540–1542, 1646
Sabine formula, 680
Sacculus, 1867
Sampled-time system, 1082, 1098

Sampling
period, 799
rate, 799
synthesis, 378, 409
theory, 799
Satellite, 826, 885, 912, 915, 1878
Saxophone, 422–423, 431–432
Scaling method, 219
Scanning laser doppler vibrometer (SVM), 1404, 1414, 1419
Scattering loss, 1673, 1675
Search algorithm, 590, 1630, 1692, 1725, 1728, 1730
Secondary calibration, 1293
Secondary wave *see* S wave
Segregating simultaneous sound sources, 1852
Seismic
arrays, 1655
signal processing, 1560–1561, 1566, 1568–1569, 1575, 1612, 1617
waves, 1526–1528, 1530–1531
Seismogram, 1525, 1527, 1535, 1537–1538, 1587
Seismology, 1506, 1523, 1526–1527, 1571–1581
exploration, 1538–1539, 1595, 1605, 1657
earthquake, 1539
Sensitivity function, 1053, 1083, 1085–1086, 1090–1091, 1098, 1100–1101
Sensor, kinematic viscosity, 984, 999
Sensory substitution, 1218–1219, 1231–1241
Separation of speech voices, 288
Sequence modulation method, 957–961
Sequential testing, 217
Servo-control systems, 1082
Servo-valves, 1142–1143
Shaft vibration measurement, 1344
Shallow water, 1670–1672, 1677, 1685, 1733, 1748, 1804, 1863
Shear layer, 1823–1824
Shear modulus, 424, 426, 1176–1177, 1179–1180
Shear waves, 1176, 1526–1531, 1537–1539, 1542 *see also* S waves
Shielding and scattering, 949
Shift-and-add, 92, 93
Shimmer, 560
Short-term memory, 206, 246, 812
Short-time Fourier transform (STFT), 47–49, 292–294, 307, 366, 1365–1368, 1370–1372
Short-time intensity, 1410
Short-time spectral attenuation (STSA), 778–779
Sidelobes, 40–41, 142, 144, 970, 1715, 1719
Sigma-Delta converters, 19
SIGNAL, 1895–1896
Signal
detectability, psychophysical theory of, 1011
detection theory, 218
digitization, 853–856
excess, 1754, 1760–1761
follower, 1778
model, 285–286, 289, 291, 293, 296, 362–363, 378, 1559–1569, 1669–1693
reconstruction, 777, 1376
subspace method, 1490
Signal classification, 4–12
almost periodic, 7
chaotic, 7
cyclo-stationary, 10
non-stationary random, 10
periodic, 5–6
stationary random, 10
transient, 7
Signal correlation, 888, 890, 1490, 1427
Signal processing
algorithms, 399, 484, 515, 691, 1560, 1568
package Matlab®, 38
scheme, 205, 207
theory, 1426
Signal-to-noise ratio (SNR), 17, 26, 71, 79, 79, 97, 103, 134, 169, 185, 201, 206, 487, 610, 663, 733, 778, 823, 849, 898, 955, 967, 1284, 1334, 1426–1427, 1429, 1432, 1462, 1466, 1470, 1474–1475, 1484, 1491, 1574, 1588, 1655, 1657, 1661, 1738, 1760, 1840, 1859, 1905
Signal-to-quantisation noise ratio (SQNR), 453
Silence descriptor (SID), 606
Similar material, 306
Similarity matrix, 313–318, 320–321, 324, 328, 1895
Simmering, 348
Simulated annealing, 1642, 1731–1732
Simultaneous iterative reconstruction technique (SIRT), 1639–1640
Simultaneous sample and hold, 29
Simultaneous-evoked OAE, 163
Sine law, 750
Singer's formant, 361
Singing physical articulatory synthesis model (SPASM), 368
Singing synthesis, 360, 364, 367, 369
Single crystal, 1352, 1783, 1789
Single degree-of-freedom (SDOF) system, 49
Single flux quantum (SFQ), 1351
Single input-single output, 12
Single scattering, 1562–1564, 1567, 1647, 1662
Single-ended or output-based approach, 511
Single-phoneme models, 370
Single-reed woodwind, 409
Singular value decomposition (SVD), 1063–1066, 1075, 1726–1727
Sinuosity or "shape factor", 1170
Sinusoidal model, 284–286, 288, 291, 366, 370, 572, 611, 778, 793, 866, 1365, 1367, 1370, 1379
Sinusoidal voice coding, 366–367
Sinusoids plus noise model, 285
Slant stacks *see* Vespagrams
Sleep disturbance, 1006, 1010, 1016–1017
Slotted cylinder, 1792, 1794–1795
Slowness vector, 1445, 1655–1659
Slow-phase modulation, 686
SmartMusicKIOSK, 326–327
Smooth time, 340–341, 344
Smoothing, 312, 563, 570, 572
SMV, 607
Soft clipper, 407
Soft formatting, 709
Somatosensation, 1194, 1215
SONAH (statistically optimal NAH), 137, 144
Sonar equation, 1744–1746, 1754, 1760–1761

Sonar, passive, 1717, 1743, 1757–1761, 1767–1769, 1772, 1773
Sone, 252, 257–258
Song learning, 1892–1893
Sonogram, 47, 1887, 1892, 1900
Sound
  absorbing and sound-insulating, 1168
  attenuation
  coefficient, 951
  measurement, 954, 959
  damping materials, 1176
  detection, 1851–1852
  energy level, 943–946
  exposure level, 43, 945, 1016–1017
  insulation, 1167
  intensity, 1109, 1283
  measurement, 1115, 1117, 1121
  localization, 182, 215, 731, 1845, 1862, 1872
  power, 933–946, 1109, 1121
  measurement of, 1289
  pressure method, 934–937, 944–946
  source distance (SSD), 1391–1392, 1397–1398
  speed profile, 1671, 1685, 1689–1690
  synthesis, 336, 399, 409
Sound designer I format, 711
Sound field
  calibrators, 1310–1311
  corrections, 1300, 1303
  visualizations, 1411
Sound file formats, 711
Sound recordings, 668, 800, 809, 1862–1863, 1280, 1880
Sound separation, 278, 284, 287
Soundscapes, 353
Source level, 1744–1745, 1749, 1751, 1760, 1804, 1807, 1899, 1901, 1904–1905
Source matrix, 1561–1562
Source separation, 165, 285–286, 295, 298, 793, 911
Source spectrum, 1682
Source-filter model, 361, 371, 464–466, 802
Spaced microphone techniques, 756–758
Spallanzani, 1835
Sparse coding, 292–293, 296–297
Sparse Gaussians, 522, 524, 527, 529
Sparse temporal sampling, 170
Sparseness, 297
Spatial aliasing, 142, 695, 902, 1403, 1426, 1428, 1488–1489, 1510, 1556, 1764
Spatial angular frequencies, 132
Spatial event plots, 1207
Spatial Fourier transform, 133, 136–137, 141, 1179–1180, 1488
Spatial hearing, 747
Spatial information, 1403
Spatial low-pass filter, 134, 136
Spatial processing, 1711, 1741, 1761–1762, 1765
Spatial resolution, 131, 136, 168–169, 739, 901–902, 1129, 1263, 1331, 1662, 1716, 1739
Spatial window effects, 137
Speaker adaptation, 524, 546
Speakerphones or piston-phones, 19
Specificity theory, 1202
Spectral analysis, 33
Spectral band replacement (SBR), 600–601, 803, 814
Spectral centroid, 308, 384, 391–392, 420–421
Spectral components, phase, 236
Spectral density, 233
Spectral domain, 67, 69, 80, 298, 457, 472
Spectral fluctuations, 420–421
Spectral magnitude subtraction, 898
Spectral pattern matching, 280
Spectral pitch, 244–249
Spectral resolution, 71, 259, 460, 480
Spectral smoothing, 165, 289, 578
Spectral subtraction, 594–593, 898
Spectral techniques, 579
Spectral tilt, 363, 384, 499, 508, 576, 578–579, 598–599
Spectral-temporal modulations, 187
Spectrogram, 293, 295–296, 307, 461–463, 481, 797, 798, 1470, 1769–1770, 1772, 1895
Spectro-temporal, 183, 185, 187, 254–255
Spectrum scaling, 42
Speech
  acquisition, 1219, 1232
  coding, 125, 188, 285, 449–450, 464, 588, 590, 592, 601, 603, 605, 607–609, 614, 802, 824–827, 844, 853, 855–857, 867, 876–877, 916–916, 921
  enhancement, 205, 898–899
  envelopes, 202, 494–497, 512
  intelligibility, 154, 187, 197, 200–202, 207, 209, 483–488, 491–493, 514–516, 822, 1273
  index, 201, 491
  test, 197, 484, 488, 491
  models, 571
  -plus-noise mixture, 594
  production, 450–455
  quality, 492, 504, 509, 918–919
    estimators of , 492
    tests, 484, 500, 504, 508
  reception threshold (SRT), 197–198
  recognition sensitivity model, 203
  recordings, tools, 580
  synthesizers, 561
  transmission index, 201–202, 488–489, 677
  waveforms, 455, 463
Speech processing applications, 450, 453, 457, 480
Speechreading, 732, 740
Speech-to-noise ratio (SNR), 197, 503
Split spread, 1547, 1551, 1554
Split-step Fourier (SSF) parabolic equation, 1506
Spontaneous OAE, 162, 164
Stable/transients separation, 793
Stagnation region, 1822–1823
Standard event bulletin (SEB), 1464, 1470
Standard screened event bulletin (SSEB), 1464
Stapedius muscle, 150–151, 161, 176
State equation, 1413, 1499, 1705, 1708–1709
State transition matrix, 1705, 1708
State vector, 1705, 1708
Stationary random process, 10
Statistical decision theory, 181, 203
Statistical parametric mapping, 170

Steady state genetic algorithm (SSGA), 1158
Steepest descent
  algorithm, 1038–1041, 1062–1063, 1066–1072, 1075, 1077
  method, 126, 1038, 1041
Stereophonic setup, 748, 755–756
Stereo tomopgrahy, 1643
Stevens power law, 1195
Stiff string, 401, 406
Stiffness matrix, 427
Stimulus pairs, 217
Stimulus parameters, 220, 242, 254, 1207
Stopwatch mechanism, 1839
Stratospheric returns, 1500, 1508, 1511
Stratum germinatium, 1186
Stress-strain formulation, 425
Striated time, 340–341
Strict stationarity, 10
Strictly positive real (SPR), 1047
String instruments, 278–279, 392, 400, 442, 788–789
Struck strings, 401, 406
Structural waves, 1822
Structure-based distance, 308
Subconvective, 1827
Subduction zones, 1526
Subhar-monic applied in the virtual-pitch theory, 247
Subharmonics, 76
Substitution schemes, 206
Successive approximation (SA), 19
Super Audio CD (SACD), 722, 800
Superconducting quantum interference devices (SQUIDs), 1351
Super-heterodyne principle, 77
Superior olivary nucleus, 1868
Superior temporal sulcus (STS), 739
Supervised learning, 297, 308
Suppression experiments, 237
Suprathreshold distortions, 154
Supremum, 1090
Surface duct, 1669, 1683, 1687–1688
Surface intensity method, 1121
Surface-related multiple removal, 1568
Sweep, 74–75, 90
Swim bladder, 1852–1853, 1856, 1904
Switching, 846
Syllable, 565, 734, 737, 1894–1895, 1892
Sylvester equation, 534, 536, 540, 544
Sylvester matrix, 534
Synchronized overlap-add (SOLA), 567
Synthesis of the Singing Voice, 370–371
Synthesis Tool Kit (STK), 406
Synthesizers, 336, 364, 375, 385, 390, 560–561, 563, 565–566, 578, 1880
Synthetic aperture, 1775
System identification, 125
System under test, 65, 67, 88, 97, 489

Tactile aid, 1231, 1233–1235, 1237, 1239–1241
Tactile communication, 1219, 1231, 1235, 1239
Tactile sensitivity, 1218
Tactile vocoder, 1232, 1239
Tactual roughness, 1205
Talk tests, 835–836
Talker echo loudness, 845
Talking situation, 822, 824
Tandem-free operation, 923
Tape recorders, 704, 709, 1863, 1866, 1894, 1899
Tape recording, 702–704, 706
Tape saturation characteristics, 794
Target cost, 564–565, 574
Tau-p model, 1438, 1446–1448, 1505
Teleseismic wave analysis, 1662
Template matching, 775
Temporal clipping, 924
Temporal discontinuities, 496, 513
Temporal fluctuations, 236
Temporal integration, 185, 238, 259
Temporal modulation transfer functions (tMTFs), 180
Temporal patterns, 219
Temporal resolution, 184
Temporal stream, 219–221, 223, 226–227
Temporal summation, 1192–1193, 1198
Temporal waveform, 236
Terminal differentiation, 1186
Testosterone, 1877
Text-to-speech (TTS), 557–559
Texture
  morphing, 333, 355
  patterns, 313–316
  score representation, 324
  sequences, 324
Thalamus, 1868, 1870
Theory vs. model, 241
Thermal characteristic length, 1168, 1174
Thermal overload, 660
Thermal runaway, 1809
Thermoacoustics, 1356–1362
  engines, 1356
  heat pump mode, 1356–1358
  resonator, 1356
  theory, 1350
Thermodynamic model, 405
Thermospheric returns, 1500, 1512
Thiele-small parameters, 633, 661
Third Generation Partnership Project (3GPP), 600
Tikhonov filter, 135
Tikhonov regularization, 1611
Timbre, 341, 346, 420, 422, 440
  attack and vibrato, 420
  models, 286
  morphing, 333, 341, 355
  variation, 423
Time aliasing, 74, 88
Time alignment, 672, 684, 843
Time and intensity pathway, 1873
Time bandwidth product, 1740–1741, 1771
Time delay
  closure, 1427, 1465, 1493
  estimation algorithm, 496–497
  spectrometry, 77
Time dispersion, 1684–1685
Time invariant, 12, 88, 97, 405, 408, 693, 901, 935
Time panning, 754, 756
Time resolution, 116
Time stretched pulse, 80, 955, 958
Time stretching, 351, 785, 790–791, 793
Time structure, 91, 246
Time warping, 1372–1379
Time-domain
  and frequency-domain algorithms, 566
  harmonic scaling (TDHS), 567
  model, 291–292, 579
  parabolic equation (TDPE) model, 1503, 1510

Time-frequency
  analysis, 47
  components, 286
Time-harmonic representation, 139
Time-lag matrix, 317, 320–323
Time-lapse method, 1546
Time-reversed excitation signal, 67
Time-series analysis, 3
Time-varying
  amplitudes, 292, 389
  delay, 496–497, 843
  systems, 14
Tinfoil cylinder phonograph, 773
Tomography
  autoregressive technique, 1641
  cross-borehole, 1643, 1647
  double-difference, 1646
  global, 1646, 1648
  irregular grids, 1640
  local earthquake, 1646
  nullspace shuttle, 1639
  refraction, 1643
  stereo, 1643
  travel-time, 1636, 1642–1643, 1646
  X-ray, 168
TomoVes experiment, 1643–1645
Tonal percept, 247
Tonal stimuli, 244–247
Tonotopic organization, 156
Tortuosimeter, 1171
Tortuosity, 1168
Torus semicircularis, 1868
Total energy density, 1289
Tracking, 126, 1707, 1710
Tractable, 1307, 1566–1568
Trajectory models, 524, 528
Transducer, 631–632, 642, 1349, 1783, 1798, 1801
  dynamic, 631–632, 660–661
  electric-to-acoustic, 1362
  electromagnetic, 632
  electromagnetic-acoustic, 631
  electronic data sheet, 1325
  electrostatic, 632
  flextensional, 1792, 1805
  membrane-based, 1268
  theory, 625, 641
Transduction, 1248
Transfer function (TF), 65, 1381
  complex, 65, 70
Transfer matrix, 1561
Transform coded eXcitation (TCX), 601
Transient detector, 1772
Transient signal, 7
Transient-evoked emissions, 162
Transitory hyperventilation, 1225
Transmission
  loss, 956, 1125, 1506–1507, 1562, 1725, 1744–1745, 1751, 1760–1761
  path, acoustic, 71
  performance, background noise, 825
  errors, 496, 602, 608–609, 827, 918
  error-free, 917
  packet based, 826, 916–918
  packet-switched, 877
Transmission line, 164, 176, 628, 634, 826, 1202, 1298, 1798
  loudspeakers, 634
  model, 177
    nonlinear, 164
Transmitting voltage response, 1799
Transparent coding, 801, 804
Transport equation, 1441, 1504
Transversal filter, 125
Transversal or tangential wave, 1538
Travel time, 1441–1443, 1447–1451, 1456, 1469, 1498, 1504–1505, 1510–1511, 1523, 1537, 1629, 1633, 1662, 1690–1691, 1724, 1728
Traveling wave, 176, 402
Tremolo, 343, 785, 788
Triangulation, 973, 1330, 1625, 1776
Triboelectric effect, 1320
Trigger zone, 1220–1221
Trills, 342–342
Triode, 794
Trombone, 423
Tube distortion, 407
Tukey window, 137
Tuning curves, 230, 232–233, 1871
Tuning fork tests, 151
Turbulent boundary layer, 1822–1824, 1828
Tweeter, 410, 649–650, 658, 660–661, 674
Two component approach, 257
Two-microphone technique, 1284
Two-tone masking, 234, 237
Two-tone suppression, 178, 1869, 1875
Two-way mismatch algorithm, 378
Tympanometry, 160
Tympanum, 149

Ultrasonic
  measurement techniques, 103
  non-destructive testing, 94
  tortuosimeter, 1172
  transducers, 624, 626, 637, 642, 1172
Uncomfortable level (UCL), 258
Underwater acoustics, 104, 691, 1676, 1678–1679, 1682, 1715, 1723, 1725, 1729, 1731, 1762
Unequal error protection (UEP), 610, 871
Uniform shading, 142–143
Unique Material Identifier (UMID), 716
Unitary pitch model, 283
Unmodulated noise, 265
Unpleasantness, 264
Unvoiced sounds, 361, 363, 856

Vagal, 1876
"Variable *d*" principle, 1280
Variance distortionless response, 903
Vector base amplitude panning (VBAP), 751–752
Vector quantization, 609, 860–861
Vellus hair, 1186
Velocity model building, 1590
Velocity vector field, 1129, 1131
Ventriloquism illusion, 732, 739–741
Vertebrate evolution, 1870
Vespagrams, 1658
Vibrating tools, 1222
Vibration damping, 1167
Vibrato, 385, 788–789
Vibrators, 1542, 1548
Vibrotactile sensitivity, 1201, 1216
Vibrotactile system, 1237
Vibrotaction, 1196, 1203, 1216, 1219, 1241
Virtual acoustics, 149, 200, 761–762, 1398
Virtual earth, 1084
Virtual musical instruments, 399
Virtual tape, 727
Virtual-pitch, 244–248

Viscous characteristic length, 1168, 1174
Viscous friction, 1167, 1170
Viterbi algorithm, 312, 868
VMR-WB, 595, 604, 606–607, 611
Vocal animals, 1861
Vocal cords, 450–451, 464, 590, 1845–1846, 1876
Vocal folds, 360–361, 363–364, 369, 560, 577, 580, 1845–1845
Vocal timbre, 362
Vocal tract, 360–361, 363–369, 450–451, 458, 464–466, 468–471, 473, 475–476, 478, 480–481, 514, 560, 576, 855, 858
  nodels, 364
  physical modeling, 368
Vocoder, 363–365, 369, 1365
  homomorphic, 365
  phase, 353, 364, 377–378, 1365–1366
  tactile, 1232, 1239
Voice activity detector (VAD), 594, 603, 605, 918
Voice coding, sinusoidal, 366–367
Voice coil impedance, 633
Voice-controlled attenuation, 825, 844
Voice enhancement devices, 925
Voice modeling, principles of singing, 360
Voice morphing, 367
Voice over Internet protocol (VoIP), 493, 496, 605, 612–613, 825, 844, 918
Voice quality, 575, 577–579
Voiced sounds, 360–361, 845, 856, 859
Volterra
  expansion, 14
  kernels, 93
  series, 92, 636
Vowel consonant-vowel stimuli, 733
Vowel recognition, 737

Walmsley's method, 291–292
Waterbed effect, 1092, 1096
Water-house's correction, 937
Wave equation, 400–402, 628, 762, 1108, 1130, 1380, 1440–1441, 1500, 1504, 1506, 1512, 1526, 1561, 1566–1567, 1589–1590, 1676–1678, 1681–1682, 1724, 1776
  migration, 1589
WAVE Format, 713, 715
Wave propagation
  anisotropic elastic media, 1536–1537
  heterogeneous elastic media, 1540
  homogeneous, isotropic media, 1537
  isotropic elastic media, 1537
  porous media, 1541
  transversely isotropic media, 1539
Wavefield
  operators, 1559, 1568
  synthesis, 912
Waveform similarity overlap-add (WSOLA), 369
Wavefront construction (WC) technique, 1456
Waveguide
  digital filter, 368
  mesh, 400, 762
Wavelet
  analysis, 49–50, 392
  mother, 49, 50
  shaping, 1574
  shrinkage, 779
  transformation, 439
Wavenumber
  integration, 1676–1678
  leakage, 137
  space, 1132–1134
  spectrum, 133–135, 137, 1714
  vector, 133–134, 141–142, 1445, 1659, 1764
Wavetable
  index synthesis, 380, 382–383
  interpolation synthesis, 380
  multiple synthesis, 378–380, 385, 389
  synthesis, 375–376, 378, 380–381, 385, 387, 389, 394
Weber's law, 231
Weberian ossicle, 1853
Weighted count test, 1468
Weighted speech signal, 597
Weighting filter, 597–592, 787
Weighting for sound measurements, 108
Weighting functions, 270
Wet end, 1754
Whales, 1757, 1836, 1895, 1897–1898, 1902–1905
Whisperization, 792
Whistles, 1897, 1899
White finger, 1222
Whole body vibration, 1224
Wideband
  random noise, 248
  spectrogram, 461–463
  speech, 595–597, 601, 606, 608
  telephony, 492, 607, 614
Wiener filter, 779, 886, 898–899, 902–903
Wiener-Levinson estimator, 1575
Wigner-Ville distribution (WVD), 48–49
  smoothed pseudo, 48
Wind-corrected amplitude, 1511–1512
Wind-generated turbulence, 1473
Window function, 40–41, 134, 137–138, 377, 791, 970, 1366, 1371, 1374
Window
  exponential, 41, 134–135, 1399–1400
  optimum data, 53, 58–60, 62
  rectangular, 26, 40–41, 56, 79, 134, 137, 457
  spatial effects, 137
  spectral, 57
  tukey, 137
Wireless telephony, 493, 607
WKBJ approximation, 1441
Woofer, 649–650, 675
Working standard (WS), 1302
Wow and flutter, 774, 1864
Wraparound, 1134
Write-once-read-many (WORM), 709

Youla parameterisation, 1100
Young Modulus, 424, 1177–1176

Zebra finch, 1894, 1896–1892
Zech's logarithm, 92
Zero padding, 99, 138, 474
Z-network, 109
Zone of silence, 1454
Z-transform, 54–56, 60, 367
Zwicker model of loudness, 254, 257
Zwicker's and Terhardt's formula, 235
Zwislocki's theory, 1192–1193

Cumulative Contents

Editor Biographies v
Editorial Board vii
Preface ix
Acknowledgments xi
Contributors xvii

PART I ACOUSTIC SIGNALS AND SYSTEMS 1

1 Signals and Systems 3

1 Temporal Signal Classification 4
2 System Definition and Classification 12

2 Acoustic Data Acquisition 17

1 Electronic Noise 17
2 Analog-to-Digital Conversion 19
3 Anti-aliasing Filters 22
4 Basic Digital Signal Calibration 24
5 Multichannel System Calibration 28
6 Data Validation and Diagnostic Techniques Summary 31

3 Spectral Analysis and Correlation 33

1 Introduction 33
2 The Fourier Transform and Variants 35
3 Practical FFT Analysis 40
4 Spectral Analysis Using Filters 43
5 Correlation Functions 44
6 Time–Frequency Analysis 47

4 The FFT and Tone Identification 53

1 Introduction 53
2 The FFT (and DFT) 54
3 Leakage 55
4 Windows 57
5 Estimation of Amplitude, Frequency, and Phase 59
6 Applications 62

5 Measuring Transfer-Functions and Impulse Responses 65

1 Introduction 65
2 Measurement Methods 66
3 Excitation Signals 79

6 Digital Sequences 87

1 Introduction 87
2 Golay Codes 87
3 Binary Maximum-Length Sequences 90
4 Fast MLS Transform (FMT) 97
5 Gold and Kasami Sequences 100
6 Legendre Sequences 102
7 Application Remarks 103

7 Filters 107

1 Introduction 107
2 Frequency Weightings 108
3 Octave- and Fractional-Octave-Bandpass Filters 112
4 Phase Linear Filters 119
5 In-Phase/Quadrature Filters 121
6 Perceptual Masking Filters 122

8 Adaptive Processing 125

1 Introduction 125
2 The LMS Algorithm 126
3 Adaptive Applications 128

9 Beamforming and Wavenumber Processing 131

1 Planar Nearfield Acoustical Holography 131
2 Beamforming 139

PART II AUDITORY SYSTEM AND HEARING 145

10 Anatomy, Physiology and Function of the Auditory System 147

1 External Ear 147
2 Middle Ear 149
3 Inner Ear 151
4 Auditory Pathway 154

11 Physiological Measures of Auditory Function 159

1 Head-Related Transfer Functions 160
2 Tympanometry 160
3 Acoustic Middle Ear Reflex 161
4 Otoacoustic Emissions 162
5 Evoked Potentials 166
6 Functional Imaging of the Auditory System 168

12 Auditory Processing Models 175

1 External and Middle Ear Transformation 176
2 Cochlear Transformation 176
3 Examples of Brainstem Processing Models 178
4 Models of Auditory Perception 180
5 Applications of Auditory Models 187

13 Speech Intelligibility 197

1 Speech Intelligibility 197
2 Measurement Methods 199
3 Factors Influencing Speech Intelligibility 200
4 Prediction Methods 201

14 Signal Processing in Hearing Aids 205

1 Basic Approach to the Rehabilitation of Hearing Deficiencies 205
2 Factors of Hearing Impairment and Rehabilitation Strategies 206
3 Limitations of Current Rehabilitation Strategies 209
4 Conclusions 209

PART III PSYCHOACOUSTICS 213

15 Methods for Psychoacoustics in Relation to Long-Term Sounds 215

1 Introduction 215
2 Methods for Psychoacoustics 216

3 Information Processing along the Temporal Stream – The Method of Continuous Judgment by Category 220
4 Information Processing in Hearing – Summary 226

16 Masking and Critical Bands 229

1 Introduction 229
2 Description of Masking Effects 230
3 Critical Bands in Psychoacoustical Experiments 233
4 Approximations of Auditory Filter Bandwidth 234
5 Limits of the Critical Band Concept 236
6 Modeling Auditory Masking 238
7 Conclusions 238

17 Aspects of Modeling Pitch Perception 241

1 Introduction 241
2 Basic Issues 241
3 Pitch of Tonal Stimuli 244
4 Pitch of Random Stimuli 248

18 Calculation of Loudness for Normal and Hearing-Impaired Listeners 251

1 Introduction 251
2 Model 254
3 Conclusions 260

19 Psychoacoustical Roughness 263

1 Introduction 263
2 Roughness in Dependence of Signal Parameters 264
3 Roughness Calculation 265

PART IV MUSICAL ACOUSTICS 275

20 Automatic Music Transcription 277

1 Introduction 277
2 Musical Sounds and F0 Estimation 278
3 Auditory Model-Based Multiple-F0 Analysis 281
4 Sound Separation Using Sinusoidal Modeling 284
5 Statistical Inference Within Parametric Signal Models 289
6 Unsupervised Learning Techniques 292
7 Summary and Conclusions 298

21 Music Structure Analysis from Acoustic Signals 305

1 Introduction 305
2 Features and Similarity Measures 307
3 Segmentation 309
4 The Similarity Matrix 313
5 Finding Repeating Sequences 317
6 Music Summary 324
7 Evaluation 327
8 Summary and Conclusions 327

22 Computer Music Synthesis and Composition 333

1 Introduction 333
2 Music with Live Performers 334
3 Form, Structure and Time 340
4 Timbre and Texture 341
5 Tuning Systems 344
6 Musique Concrète and Collage 350
7 Environments 353
8 Conclusion 355

23 Singing Voice Analysis, Synthesis, and Modeling 359

1 Introduction 359
2 Principles of Singing Voice Modeling 360
3 Models of the Glottal Source 362
4 Models of the Vocal Tract 364
5 Synthesis of the Singing Voice 370
6 Conclusion 371

24 Instrument Modeling and Synthesis 375

1 Introduction 375
2 Spectral Analysis 377
3 Wavetable Synthesis 378
4 Wavetable Matching 381
5 Frequency Modulation (FM) Synthesis Models 385
6 Frequency Modulation Matching 391
7 Conclusions 394

25 Digital Waveguide Architectures for Virtual Musical Instruments 399

1 Introduction 399
2 Vibrating Strings 400
3 Wind Instruments 409
4 Recommended Reading 412

26 Modeling of Musical Instruments 419

1 Introduction 419
2 Finite-Element Formulation 425
3 Finite-Difference Formulation 427
4 Turbulent $k-\epsilon$ Model of Flute-Like Instruments 428
5 Multiphonics in Blown Instruments in Terms of a Split-Phase Pressure Description 431
6 Time-Dependent Model of a Classical Guitar Body 436
7 Guitar Back-Plate Timbre Brightening by Stress Bending 440
8 Church Bell Clapper Time-Dependent Striking Force 442

PART V SPEECH 447

27 Display and Analysis of Speech 449

1 Introduction 449
2 Speech Production, Sampling and Quantisation 450
3 Display of Speech Waveforms 455
4 Analysis of Speech Waveforms 463
5 Conclusion 481

28 Estimation of Speech Intelligibility and Quality 483

1 Introduction 483
2 Estimators of Speech Intelligibility 485
3 Estimators of Speech Quality in Telecommunications 492
4 Conclusion 514

29 Gaussian Models in Automatic Speech Recognition 521

1 Introduction 522
2 Background 525
3 Training Regularized Gaussian Decompositions 532
4 Training Shared Regularized Gaussian Decompositions 541
5 A Generalized EM Procedure 549
6 Conclusions 550

30 Speech Synthesis 557

1 Concatenative Text-to-Speech Synthesis 558
2 Prosody Modification 565
3 Smoothing 570
4 Estimation of Unit Selection Costs 573
5 Voice Quality/Speaker Modification and Other Open Issues 575
6 Further Reading and Relevant Resources 581

31 Speech Coders **587**

1 Introduction 587
2 Speech Coding Concepts 588
3 Issues Related to Input Conditions 591
4 Issues Related to Channel Conditions 602
5 Current Developments and Future Challenges 614

PART VI AUDIO ENGINEERING **621**

32 Transducer Models **623**

1 Modelling Fundamentals 623
2 Historical Introduction 624
3 Current Literature Overview 625
4 Overview of the Models 625
5 Classical Principles 628
6 Fundamentals of Transducers 631
7 Dynamic Loudspeakers 632
8 Horn Loudspeakers 638
9 Electrostatic Loudspeakers 640
10 Microphones 640
11 Piezoelectric Transducers for Ultrasonic and Underwater Applications 642

33 Loudspeaker Design and Performance Evaluation **649**

1 Loudspeaker as a System Design Problem 649
2 Crossover Filters 651
3 Frequency Response Equalization 655
4 Room Response Equalizers 657
5 System Delay and Delay Variation 659
6 Protection 660
7 Understanding Loudspeaker Operation by Modeling 661
8 Acoustic Measurement Techniques 662
9 Listening Tests 664

34 PA Systems for Indoor and Outdoor **669**

1 Introduction 669
2 The PA System 671
3 Planning and Setting Up the Loudspeakers 676
4 Sound Indoors 679
5 Sound Outdoors 682
6 Measurement Methods 683
7 Special Issues on Signal Processing for PA Systems 684

35 Beamforming for Speech and Audio Signals 691

1 The Basic Concept 692
2 Data-Independent Beamforming 693
3 Differential Microphone Arrays 696
4 Statistically Optimum Beamformers 697
5 Blind Source Separation as Blind Beamforming 699
6 Concluding Remarks 700

36 Digital Audio Recording Formats and Editing Principles 703

1 Digital Tape Recording 703
2 Digital Recording on Computer Mass Storage Media 707
3 Consumer Digital Formats 718
4 Digital Audio Editing Using Mass Storage Media 723

37 Audiovisual Interaction 731

1 Audiovisual Space 731
2 Audiovisual Speech 732
3 Neural Basis of Audiovisual Perception 738

38 Multichannel Sound Reproduction 747

1 Introduction 747
2 Concepts of Spatial Hearing by Humans 747
3 Loudspeaker Layouts 748
4 Reproduction Methods 750
5 Capturing Sound for Multichannel Audio 755

39 Virtual Acoustics 761

1 Introduction 761
2 Auralization and Signal Processing 762
3 Waveguide Meshes in Room Acoustics 766

40 Audio Restoration 773

1 A Brief History of the Recording Technology 773
2 Aims and Problems 773
3 Overview of Digital Audio Restoration 774
4 Future Directions 780

41 Audio Effects Generation 785

1 Filters and Delays 786
2 Tremolo, Vibrato and Rotary Speaker 788
3 Chorus, Flanger and Phaser 789
4 Pitch Shifting and Time Stretching 790

5 Dynamics Processing 793
6 Room Simulation 794

42 Perceptually Based Audio Coding 797

1 Introduction 797
2 Transparent Coding: General Strategy 801
3 Transparent Coding: Specific Examples 804
4 Subjective Evaluation of Sound Quality 811
5 Objective Evaluation of Sound Quality 813
6 Concluding Remarks 814

PART VII TELECOMMUNICATIONS 819

43 Speech Communication and Telephone Networks 821

1 General Aspects of Speech Communication 821
2 Modern Telephone Networks 825

44 Methods of Determining the Communicational Quality of Speech Transmission Systems 831

1 Assessment Methods 831
2 Subjective Assessment 832
3 Objective Methodologies 838
4 Network Planning 848

45 Efficient Speech Coding and Transmission Over Noisy Channels 853

1 Introduction and Fundamentals 853
2 Speech Coding 857
3 Speech Transmission over Noisy Channels 867
4 Further Work, Results and Literature 877

46 Echo Cancellation 883

1 Echo Cancellation and Suppression 883

47 Noise Reduction and Interference Cancellation 897

1 Single-Channel Techniques 898
2 Multichannel Techniques 900
3 Blind Signal Separation of Convolutive Mixtures 904

48 Terminals and Their Influence on Communication Quality 909

1 Acoustic Interfaces 910

49 Networks and Their Influence on Communication Quality 915

1 PSTN (Public Switched Telephone Networks) 915
2 Packet-Based Transmission 916
3 Mobile Networks 918

50 Interaction of Terminals, Networks and Network Configurations 921

1 Tandeming of Speech Coders 921
2 Delay 923
3 Tandeming of Speech Echo Cancellers 923
4 Tandeming VAD and NLP 924
5 Other Signal Processing 925

List of Important Abberviations 927

PART VIII NOISE 931

51 Sound Levels of Noise Sources 933

1 Sound Power Level 933
2 Sound Energy Level 943

52 Outdoor Measurements 949

1 Measurement of Acoustical Properties of Ground Surfaces – *Hideki Tachibana* 949
2 Measurement of Shielding Effect of Noise Barriers – *Kohei Yamamoto and Yasushi Miki* 953
3 Measurement of Sound Attenuation for Long-Distance Propagation Outdoors – *Yasushi Miki* 959
4 Noise Sources in Road and Railway Transport – *Yasushi Takano* 963
5 Detection of Aircraft Sound Events – *Ichiro Yamada* 972

53 Measurements of Physical Characteristics Using Sound at Audible or Lower Frequencies 983

1 Acoustic Micrometer [59] 984
2 Acoustic Volume Meter [45,51] 986
3 Acoustic Surface Area Meter [56] 989
4 Acoustic Rate Gyro [35] 995
5 Acoustic Impedance Flowmeter [18] 996
6 Chopper Microphone [29] 997
7 Kinematic Viscosity Sensor [27] 999

54 Community Response to Noise 1005

1 Statistical Character of Environmental Noise in Populated Areas 1006
2 Predicting the Annoyance of Environmental Noise 1008
3 Predicting Sleep Disturbance 1016
4 Predicting the Likelihood of Noise-Induced Complaints 1017
5 Summary 1018

PART IX ACTIVE NOISE CONTROL 1021

Introduction 1023

55 Forward Control in Single-Channel Systems 1029

1 Introduction 1029
2 Control of Deterministic Disturbances 1033
3 Adaptive FIR Controllers 1043
4 Equivalent Feedback Controller 1051

56 Forward Control in Multichannel Systems 1057

1 Introduction 1057
2 Optimal Control of Tonal Disturbances 1058
3 Steepest Descent Algorithms 1062
4 Trade-Off Between Error Attenuation and Control Effort: The Leaky Algorithm 1070
5 Robustness to Plant Uncertainties and Plant Model Errors 1072
6 Applications: Controlling Propeller Noise in Aircraft 1076

57 Feedback Control Systems 1081

1 Introduction 1081
2 Analogue Controllers 1086
3 Application: Active Headsets 1094
4 Digital Controllers 1098
5 Hybrid Feedback Controllers 1102

PART X STRUCTURAL ACOUSTICS AND VIBRATION 1107

58 Intensity Techniques 1109

1 Introduction 1109
2 Sound Intensity and Sound Power 1109
3 Active and Reactive Sound Fields 1111

4 Measurement of Sound Intensity 1115
5 Applications of Sound Intensity 1122

59 Nearfield Acoustical Holography 1129

1 Introduction 1129
2 Planar Nearfield Acoustical Holography 1130
3 Cylindrical Nearfield Acoustical Holography 1135
4 Spherical and General Surface Nearfield Acoustical Holography 1137

60 Actuator Design and Deployment 1141

1 Introduction 1141
2 Actuator Design 1142
3 Actuator Deployment Using Genetic Algorithms 1151

61 Materials Testing 1167

1 Introduction 1167
2 Sound Propagation in Rigid Porous Materials 1168
3 Sound Propagation in Elastic Porous Materials 1175

PART XI BIORESPONSE TO VIBRATION 1183

62 Tactile Responses to Vibration 1185

1 Introduction 1185
2 Neural Structures 1188
3 Thresholds of Detection in the Temporal and Spatial Domains 1190
4 Suprathreshold Responses 1194
5 Subject Variables 1197
6 The Four-Channel Model of Cutaneous Mechanoreception 1202
7 Roughness and Texture 1205
8 Biophysics 1206

63 Uses and Applications of Somatosensation 1215

1 Introduction 1215
2 Intermodality Interactions 1216
3 Sensory Substitution 1218
4 Health 1219
5 Environmental Well-Being 1221
6 Whole Body Vibration 1224

64 Sensory Substitution 1231

1 Introduction 1231
2 Temporal Gap Detection 1234

3 User Variables 1235
4 Practical Design Considerations 1235
5 Mechanical Stimulators 1236
6 Electrical Stimulators 1236
7 Coding 1237
8 The Systems 1239
9 Training 1239

PART XII TRANSDUCERS FOR ACOUSTIC MEASUREMENT 1245

65 Condenser Microphones 1247

1 Introduction 1247
2 General Principle 1247
3 Transduction 1248
4 Preamplifiers 1250
5 Sensitivity 1252
6 Frequency Response 1252
7 Microphones and Sound Fields 1255
8 Critical Design Parameters and Materials 1257
9 Acoustic Impedance of Diaphragm 1258
10 Dynamic Range 1260
11 Microphones and International Microphone Standards 1262
12 Microphones of Other Designs and Materials 1263

66 Dynamic Pressure Microphones 1267

1 Introduction 1267
2 Driving the Membrane 1268
3 Generating the Electric Output 1270
4 Advantages and Limitations 1271

67 Dynamic Pressure Difference Microphones 1275

1 Introduction 1275
2 The Frequency Response of a Dynamic Pressure Difference Microphone 1275
3 The Directional Response of a Pressure Difference Microphone 1278
4 The Ribbon Microphone 1278
5 Cardioids and Other Directivity Patterns 1279
6 Interference Microphones 1280
7 Reflectors 1281
8 Fixed and Variable Arrays 1281

68 The Microflown Particle Velocity Sensor 1283

1 Introduction 1283
2 The Microflown Transducer 1285
3 Limitations 1286
4 Calibration 1288
5 Applications 1288

69 Microphone Calibration 1293

1 Introduction 1293
2 Calibrations Performed at Primary and Secondary Levels 1293
3 Principal Types of Sound Field and Microphone Responses 1294
4 Laboratory Standard Microphones Suited for Primary Calibration 1295
5 Principle of Microphone Reciprocity Calibration 1295
6 Primary Pressure Response Calibration by Reciprocity Technique 1297
7 Primary Free-Field Response Calibration by Reciprocity Technique 1299
8 Sound Field Corrections of Laboratory Standard and Measurement Microphones 1300
9 Determination of Diffuse-Field Frequency Response 1300
10 Working Standard Microphones 1302
11 Sensitivity and Frequency Response Calibration by Comparison 1303
12 Frequency Response Calibration by Electrostatic Actuator 1304
13 Free-Field Response Calibration by Free-Field Comparison 1307
14 Traceability and Accreditation 1307
15 Dynamic Linearity 1308
16 Field Calibration with Sound Field Calibrators 1310

70 Piezoceramic Accelerometer 1313

1 Introduction 1313
2 Technology Fundamentals 1313
3 Construction 1316
4 Selecting and Specifying 1319
5 Mounting 1323
6 Latest and Future Developments 1325
7 List of Applicable Standards 1326

71 Laser Doppler Velocimeter 1329

1 Introduction 1329
2 Operation Principle 1332

3 Basic Construction 1333
4 Applications 1337

72 Capacitive Sensors 1339

1 Principle of Operation 1339
2 Some Useful Advice for Capacitive Sensors 1343
3 Applications 1344

PART XIII ENGINEERING ACOUSTICS 1347

73 Energy Saving Transducers and Thermoacoustics 1349

1 Introduction 1349
2 Challenges to Improving Dynamic Speakers 1350
3 Superconductive Loudspeakers 1351
4 Electrostatic Loudspeakers 1354
5 One-Bit Driving 1356
6 Thermoacoustics 1356
7 Summary 1362

74 Sinusoidal Representation of Acoustic Signals 1365

1 Introduction 1365
2 Phase Vocoder 1365
3 Sinusoidal Model 1367
4 Sinusoidal Model for Rapidly Changing Signal 1370
5 Time-Scale and Pitch Modification 1379

75 Room Transfer Function 1381

1 Introduction 1381
2 Modal Expansion of the TF 1381
3 Distribution of Eigenfrequencies 1382
4 Chaotic Properties of Ray Tracing and Eigenfrequency Distribution 1386
5 Distribution of Zeros 1388
6 Phase Response of Room TF 1390
7 Control of Magnitude and Phase for Artificial Reverberation Rendering 1398

76 Spatial Information of Sound Fields 1403

1 Introduction 1403
2 DOA Estimation 1404
3 Spatial Information Measurement: Four Closely Located Point Microphones Method 1408
4 Sound Field Measurement Using a Laser 1411
5 Summary 1417

PART XIV INFRASONICS 1423

77 Infrasound Event Detection Using the Progressive Multi-Channel Correlation Algorithm 1425

1 Introduction 1425
2 Correlation: The Basic Component of PMCC 1427
3 PMCC Configuration 1429
4 Processing Examples 1430

78 Ray Tracing in an Inhomogeneous Atmosphere with Winds 1437

1 Historical Introduction 1437
2 Atmospheric Variability 1439
3 Geometrical Acoustics 1439
4 Concluding Remarks 1456

79 Infrasonic Data Processing at the Provisional Technical Secretariat of the Comprehensive Nuclear-Test-Ban Organization 1461

1 Overview of IDC Processing 1463
2 Infrasonic Station Processing at the IDC 1464
3 Automatic Network Processing 1467
4 Interactive Processing 1469
5 Post-location Processing 1470

80 The Design and Performance of Infrasound Noise-Reducing Pipe Arrays 1473

1 Introduction 1473
2 Historical Background 1474
3 Design Principles for a Robust Infrasonic Noise-Reducing System 1475
4 Noise-Reducing Pipe Array Design and Performance 1477
5 Conclusion 1485

81 Processing Infrasonic Array Data 1487

1 Array Processing 1487
2 Bearing Estimation 1487
3 Signal Detection and Enhancement 1492

82 Infrasound Propagation 1497

1 Introduction 1497
2 Relevant Physics and Issues 1499

3 Propagation Models 1503
4 Summary and Future Direction 1512

PART XV SEISMOLOGY 1521

83 Structure and Elastic Properties of the Earth 1523

1 Why Seismology is Studied 1523
2 Structure and Composition of the Earth 1523
3 Seismic Waves 1526
4 Elastic Properties of the Earth 1528
5 Commercial Seismology 1531

84 Seismic Wave Propagation 1535

1 Introduction 1535
2 Medium Models 1536
3 Reflections of Plane Wave 1542

85 Seismic Data Acquisition 1545

1 Introduction 1545
2 The Acquisition System 1546
3 Signal-to-Noise Quality 1550
4 Sampling for Migration Imaging 1554

86 Signal Models in Seismic Processing 1559

1 Introduction 1559
2 The Data Matrix 1560
3 Source, Transfer and Receiver Matrices 1561
4 Accuracy and Tractability in Seismic Signal Models 1566

87 Deconvolution and Optimal Filtering in Seismology 1571

1 Introduction 1571
2 Deconvolution with Known Source Signatures 1572
3 Deconvolution with Unknown Source Signatures 1575
4 Conclusions 1580

88 Seismic Migration Imaging 1585

1 Introduction 1585
2 The Migration Problem 1586
3 Time Migration 1588
4 Depth Migration 1589

89 Geophysical Parameter Estimation 1593

1 Introduction 1593
2 Model Selection 1594
3 Information Assessment 1606
4 Estimator Selection 1610
5 Estimator Validation 1616
6 Conclusion 1617

90 Earthquake Localization 1627

1 Introduction 1627
2 The Location Problem 1627
3 Geiger's Method 1629
4 Grid Search 1630
5 Methods to Reduce Uncertainty 1630

91 Seismic Tomography 1635

1 Introduction 1635
2 Seismic Travel-Time Tomography 1636
3 Examples of Travel-Time Tomography 1643
4 Extensions of Travel-Time Tomography 1646

92 Array Signal Processing 1655

1 Introduction 1655
2 Slowness, Backazimuth and the Complete Slowness Vector 1655
3 Array Methods 1656
4 Array Configurations 1663

PART XVI UNDERWATER AND OCEANOGRAPHIC ACOUSTICS 1667

93 Propagation and Signal Modeling 1669

1 Introduction 1669
2 The Ocean Acoustic Environment 1670
3 Numerical Modeling Techniques 1675
4 Time-Harmonic Solutions of Separable Problems 1678
5 Time-Harmonic Solutions of Non-Separable Problems 1679
6 Pulse Solutions by Fourier Synthesis 1682
7 Signal Simulation Examples 1682
8 Summary 1692

94 Statistical Signal Processing 1695

1 Background 1695
2 Detection 1697
3 Estimation 1699
4 Spatial Processing and Arrays 1710

95 Inverse Problems in Underwater Acoustics 1723

1 Introduction 1723
2 Linear Inverse Problems 1725
3 Linearized Inversion 1727
4 Nonlinear Inverse Problems 1728
5 Matched Field Inversion 1729

96 Active Sonar 1737

1 Introduction 1737
2 The Sonar Signal 1738
3 Sonar System Design 1743
4 Performance Analysis 1744
5 Discussion 1754

97 Passive Sonar 1757

1 Introduction 1757
2 Passive Sonar Systems 1758
3 The Passive Sonar Equation 1760
4 Spatial Processing and Array Gain 1761
5 Detectors 1768
6 Classification 1772
7 Localization 1773

98 Transducers for Sonar Systems 1783

1 Transduction Materials 1783
2 Conventional Designs and Modeling Approaches 1790
3 Broad Bandwidth Transducer Designs 1798
4 Considerations for Remotely Operated Undersea Vehicles 1807

99 Flow Noise 1821

1 Introduction 1821
2 Flow-Induced Noise in Hull-Mounted Sonar Systems 1822
3 Transition Zone Noise Scaling 1823
4 Turbulent Boundary Layer Wall Pressure Fluctuations 1823
5 Transducer Size and Array Effects 1827

PART XVII ANIMAL BIOACOUSTICS 1833

100 Bat Bioacoustics 1835

1 Introduction 1835
2 Recording and Monitoring Systems 1836
3 Call Characteristics 1837
4 Echolocation Performance and Models 1838
5 Hearing Performance and Models 1844
6 Stimulus generation 1845
7 Archives and Databases 1847

101 Fish Bioacoustics 1851

1 Recording and Monitoring Systems 1851
2 Hearing Performance 1851
3 Vocalization Sounds 1855
4 Locating and Tracking 1857
5 Archives and Databases 1858

102 Amphibian Bioacoustics 1861

1 Recording and Monitoring Systems 1862
2 Hearing Performance and Modelling 1866
3 Characteristics of Anuran Calls 1875
4 Stimulus Generation – Synthesized Calls 1880
5 Archives and Databases of Signals 1880

103 Avian Bioacoustics 1887

1 Recording and Monitoring Systems 1887
2 Hearing Performance 1890
3 Characteristics of Songs and Calls 1891
4 Archives and Databases of Signals 1894

104 Cetacean Bioacoustics with Emphasis on Recording and Monitoring 1897

1 Recording and Monitoring Systems 1898
2 Locating and Tracking 1901
3 Echolocation Models 1903

Index 1909