TRIBOLOGY SERIES, 41
EDITOR: D DOWSON

TRIBOLOGICAL RESEARCH AND DESIGN FOR ENGINEERING SYSTEMS

D.DOWSON
M. PRIEST
G. DALMAZ
A.A. LUBRECHT
(Editors)

ELSEVIER

TRIBOLOGICAL RESEARCH AND DESIGN FOR ENGINEERING SYSTEMS



TRIBOLOGY SERIES, 41
EDITOR: D DOWSON

TRIBOLOGICAL RESEARCH AND DESIGN FOR ENGINEERING SYSTEMS

edited by

D DOWSON, M PRIEST, G DALMAZ, AA LUBRECHT

Proceedings of the 29th Leeds-Lyon Symposium on Tribology
held at Bodington Hall, University of Leeds, UK
3rd - 6th September, 2002.



ELSEVIER
Amsterdam • Boston • London • New York • Oxford • Paris
San Diego • San Francisco • Singapore • Sydney • Tokyo 2003
For the Institute of Tribology, The University of Leeds
and
Institut National des Sciences Appliquées de Lyon

ELSEVIER B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

© 2003 Elsevier B.V. All rights reserved.

This work is protected under copyright by Elsevier, and the following terms and conditions apply to its use:

Photocopying
Single photocopies of single chapters may be made for personal use as allowed by national copyright laws. Permission of the Publisher and payment of a fee is required for all other photocopying, including multiple or systematic copying, copying for advertising or promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish to make photocopies for non-profit educational classroom use.

Permissions may be sought directly from Elsevier via their homepage (http://www.elsevier.com) by selecting 'Customer support' and then 'Permissions'. Alternatively you can send an e-mail to: permissions@elsevier.com, or fax to: (+44) 1865 853333.

In the USA, users may clear permissions and make payments through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA; phone: (+1) (978) 7508400, fax: (+1) (978) 7504744, and in the UK through the Copyright Licensing Agency Rapid Clearance Service (CLARCS), 90 Tottenham Court Road, London W1P 0LP, UK; phone: (+44) 207 631 5555; fax: (+44) 207 631 5500. Other countries may have a local reprographic rights agency for payments.

Derivative Works
Tables of contents may be reproduced for internal circulation, but permission of Elsevier is required for external resale or distribution of such material.
Permission of the Publisher is required for all other derivative works, including compilations and translations.

Electronic Storage or Usage
Permission of the Publisher is required to store or use electronically any material contained in this work, including any chapter or part of a chapter.

Except as outlined above, no part of this work may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the Publisher.
Address permissions requests to: Elsevier Global Rights Department, at the mail, fax and e-mail addresses noted above.

Notice
No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

First edition 2003

Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

British Library Cataloguing in Publication Data
A catalogue record from the British Library has been applied for.

ISBN: 0-444-51243-8 (Vol. 41)
ISSN: 0-444-41677-3 (Series)

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

INTRODUCTION

The 29th Leeds-Lyon Symposium on Tribology was held at Bodington Hall in Leeds under the title "Tribological Research and Design for Engineering Systems" from Tuesday 3rd September until Friday 6th September 2002. This was the first Leeds-Lyon meeting of the 21st century to be held in Leeds and once again terrific support was received from the tribological community. The attendance was excellent with 137 delegates from 18 countries reflecting the true international status of the Symposium. Papers were presented in 22 sessions, many parallel, and for the first time in Leeds computer-based projection facilities were made available in all presentation rooms.

There have been spectacular advances in understanding in the field of tribology in recent years. Progress has been made in relation to the engineering and interaction of surfaces, micro- and nano-tribology, elastohydrodynamics, surface films, surface texture, tribochemistry, wear and life prediction, with both experimental and theoretical contributions. This Symposium presented an opportunity to review these advances and to consider the impact of this new understanding of fundamentals upon total engineering activity embracing design, manufacture and the efficient and reliable operation of machinery.

Michael Neale OBE of Neale Consulting Engineers (UK) presented the Keynote Address with the simple but somewhat daunting title of "Tribology in Design". He stressed the vital interaction between high quality research and innovative design solutions as a two way process. Not only should designers take advantage of the latest improvements in basic understanding but also researchers must tackle the exciting and challenging practical problems of real engineering systems. This was a highly appropriate starting point for the meeting. During the subsequent three days of presentations the full range of such activities was explored with topics such as fundamental studies of dry and lubricated surfaces, tribochemistry, elastohydrodynamic lubrication, wear and life of machine elements, surface engineering, design and performance of tribological components and systems, plain and rolling element bearings, bio-tribology, automotive engine tribology and transmissions. Invited contributions were sought and generously provided by leading international figures in these fields, namely J-M. Martin, C.J. Hooke, R.A. Rowntree, Z.M. Jin, A. Kapoor, S.B. Glavatskih, N.D. Spencer, A. Jackson, R. Castle, R.W. Snidle, A.V. Olver, D. Zhu and E. Ciulli. Our thanks go to them and their chosen co-authors. In total, 85 papers are included in this Volume of Proceedings as a fully reviewed record of a thought-provoking meeting.

Successful scientific meetings require high quality presentations, enjoy lively discussion and must keep to schedule. In this regard we thank all the authors, the delegates and the Session Chairmen for their efforts throughout the Symposium. We draw attention to the fact that every effort is made to record questions from delegates and the response of the authors during Leeds-Lyon Symposia as a valuable additional contribution to the proceedings. This year, the discussion was vibrant and extensive with over 150 written questions and responses, which are documented towards the end of this Volume of Proceedings.

Naturally a social programme complemented the high quality technical sessions. On the afternoon and evening of Thursday 5th September, delegates experienced the delights of Bowness on the shores of Lake Windermere in the English Lake District. After exploring the hilly streets of this popular resort the delegates enjoyed an evening cruise and dinner on Lake Windermere. The Symposium Dinner was held at the historic Castle Howard on Tuesday 3rd September, with the after dinner speech delivered masterfully by Professor Dick Coy.

We are sincerely grateful to those who helped in the detailed planning and execution of the Symposium. Colleagues in Leeds gave unstintingly of their time and we would like to thank Mrs. Sheila Moore, Mrs. Cath Goulborn, Professor Tom Childs, Dr. Peter Dearnley, Dr. Karl Dahm, Dr. Daniel Mehenny, Mr. Ron Harding (who took a break from retirement to help), Mr. Paul Banks, Mr. Balbir Randhawa and all our postgraduate students. A special note of appreciation goes to Dr. David Barrell whose tireless contributions were critical to the smooth management of the computer-based presentations at the Symposium and the formal refereeing of the technical papers. As ever, our long-term collaborators and friends at the Laboratoire de Mécanique des Contacts (LMC) at the Institut National des Sciences Appliquées (INSA) de Lyon gave valuable support and encouragement.

The 29th Leeds-Lyon Symposium on Tribology marked the last meeting in Leeds before the formal retirement from the University of Leeds of Mrs. Sheila Moore. Sheila has been a key member of the Leeds team since the very beginning of the Symposia in 1974. We offer our sincerest thanks for all her help, encouragement and kindness over the last 29 years.

The Volume of Proceedings has been produced in excellent style by Elsevier Science Publishers BV to whom we give grateful thanks. We have maintained rigorous review procedures for the papers that appear in this volume. International referees review all papers and the editors require the corresponding author to effect appropriate amendments in the traditional style of a learned society journal if the work is accepted. We hope that such procedures will continue to enhance the high quality which has been synonymous with the Leeds-Lyon Symposia.

We are grateful to the following who gave financial support to the 29th Leeds-Lyon Symposium on Tribology

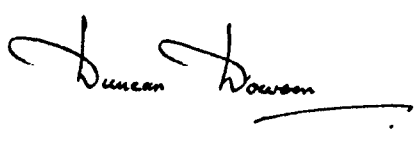
DePuy International Ltd
SKF Engineering Research Centre BV

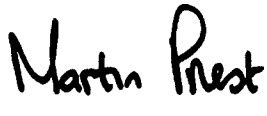
Leeds-Lyon Symposia have addressed an extensive range of topics over the last 29 years as outlined below

1.	Cavitation and Related Phenomena in Lubrication	Leeds	1974
2.	Superlaminar Flow in Bearings	Lyon	1975
3.	The Wear of Non-Metallic Materials	Leeds	1976
4.	Surface Roughness Effects in Lubrication	Lyon	1977

5.	Elastohydrodynamic Lubrication and Related Topics	Leeds	1978
6.	Thermal Effects in Tribology	Lyon	1979
7.	Friction and Traction	Leeds	1980
8.	The Running-In Process in Tribology	Lyon	1981
9.	The Tribology of Reciprocating Engines	Leeds	1982
10.	Numerical and Experimental Methods Applied to Tribology	Lyon	1983
11.	Mixed Lubrication and Lubricated Wear	Leeds	1984
12.	Global Studies of Mechanisms and Local Analyses of Surface Distress Phenomena	Lyon	1985
13.	Fluid Film Lubrication - Osborne Reynolds Centenary	Leeds	1986
14.	Interface Dynamics	Lyon	1987
15.	Tribological Design of Machine Elements	Leeds	1988
16.	Mechanics of Coatings	Lyon	1989
17.	Vehicle Tribology	Leeds	1990
18.	Wear Particles: From the Cradle to the Grave	Lyon	1991
19.	Thin Films in Tribology	Leeds	1992
20.	Dissipative Processes in Tribology	Lyon	1993
21.	Lubricants and Lubrication	Leeds	1994
22.	The Third Body Concept: Interpretation of Tribological Phenomena	Lyon	1995
23.	Elastohydrodynamics - 96	Leeds	1996
24.	Tribology for Energy Conservation	London	1997
25.	Lubrication at the Frontier	Lyon	1998
26.	Thinning Films and Tribological Interfaces	Leeds	1999
27.	Tribology Research: From Model Experiment to Industrial Problem	Lyon	2000
28.	Boundary and Mixed Lubrication: Science and Applications	Vienna	2001
29.	Tribological Research and Design for Engineering Systems	Leeds	2002

In 2003 the Symposium returns to INSA de Lyon from 2nd to 5th September 2003 and addresses "Transient Processes in Tribology". This throws down a challenge to tribologists to step out of the ideal world of steady state into the often-harsh reality of dynamic systems.





Duncan Dowson

Martin Priest

Leeds, January 2003

This Page Intentionally Left Blank

REFEREEING

The Leeds-Lyon Symposia on Tribology were launched in 1974. They were originally envisaged as informal gatherings but pressure arose almost immediately for the presentations to be published in an annual Volume of Proceedings. The large number of references to original work published in the Proceedings of the Leeds-Lyon Symposium over many years confirms the quality of the published papers. It also indicates that the volumes have served their purpose and become a recognised feature of the tribological literature.

Two developments prompted the extension of our refereeing process in recent years. The number of papers offered to the programme organisers each year has increased significantly. Secondly, international interest in the publication of peer reviewed papers to high standards has prompted a number of authors to favour such a process for papers included in the Volume of Proceedings.

The evolution of a full peer review process has been evident for a number of years. Initially individual members of the *Institute of Tribology* at Leeds and *the Institut National des Sciences Appliquées* de Lyon were able to oversee the bulk of the refereeing process, but more recently it has been necessary to involve many colleagues outside the Leeds-Lyon groups of tribologists. The procedure was gradually extended and formalised and standardised Report Forms were adopted.

An important feature of the Leeds-Lyon Symposia is the presentation of current research findings. This remains an essential feature of the meetings and great care has been taken to design a review procedure that does not detract from this vibrancy. Authors are invited to submit their papers for review a few weeks in advance of the Symposium and international referees are asked to complete specially prepared report forms and submit their recommendations to the Editors where possible at, or in advance of, the Symposium. If a paper is accepted, the Editors may then require authors to modify their manuscripts in the manner of a learned society journal.

The success and value of this formalised peer review process for the Leeds-Lyon Symposia relies heavily upon the co-operation of all authors and particularly on the speedy, expert and respected work of our referees. We are pleased to record our appreciation of the invaluable efforts of the referees listed below.

P. Andersson	D.J.W. Barrell	Y. Berthier	J.F. Booker	B. Bou-Said
P.M. Cann	T.H.C. Childs	R.J. Chittenden	E. Ciulli	R.C. Coy
K.L. Dahm	G. Dalmaz	P.A. Dearnley	D. Dowson	R.S. Dwyer-Joyce
H.P. Evans	J. Fisher	L. Flamand	J. Frene	R.J. Gamble
S.B. Glavatskih	J.A. Greenwood	M. Hadfield	S. Hogmark	K. Holmberg
C.J. Hooke	I.M. Hutchings	Z.M. Jin	M. Kaneta	A. Kapoor
R. Larsson	A.A. Lubrecht	C.N. March	A. Matthews	J.B. Medley
D.S. Mehenny	R.M. Mortier	M.J. Neale	D. Nelias	A. Neville
A.V. Olver	D.D. Parker	G. Poll	M. Priest	W.M. Rainforth
R.A. Rowntree	R.S. Sayles	I. Sherrington	R.W. Snidle	N.D. Spencer
H.A. Spikes	T.D. Stewart	R.I. Taylor	A. Unsworth	J.A. Williams
R.J.K. Wood	D. Zhu			

TRIBOLOGY SERIES
Editor
D. Dowson (Gt. Britain)

Advisory Board
W.J. Bartz (Germany)
R. Bassani (Italy)
B. Briscoe (Gt. Britain)
H. Czichos (Germany)
K. Friedrich (Germany)
N. Gane (Australia)
W.A. Glaeser (U.S.A.)
H.E. Hintermann (Switzerland)
K.C. Ludema (U.S.A.)
W.O. Winer (U.S.A.)

Vol. 6 Friction and Wear of Polymers (Bartenev and Lavrentev)

Vol. 10 Microstructure and Wear of Materials (Zum Gahr)

Vol. 11 Fluid Film Lubrication – Osborne Reynolds Centenary (Dowson et al., Editors)

Vol. 12 Interface Dynamics (Dowson et al., Editors)

Vol. 13 Tribology of Miniature Systems (Rymuza)

Vol. 14 Tribological Design of Machine Elements (Dowson et al., Editors)

Vol. 15 Encyclopedia of Tribology (Kajdas et al.)

Vol. 16 Tribology of Plastic Materials (Yamaguchi)

Vol. 17 Mechanics of Coatings (Dowson et al., Editors)

Vol. 18 Vehicle Tribology (Dowson et al., Editors)

Vol. 19 Rheology and Elastohydrodynamic Lubrication (Jacobson)

Vol. 20 Materials for Tribology (Glaeser)

Vol. 21 Wear Particles: From the Cradle to the Grave (Dowson et al., Editors)

Vol. 22 Hydrostatic Lubrication (Bassani and Piccigallo)

Vol. 23 Lubricants and Special Fluids (Stepina and Vesely)

Vol. 24 Engineering Tribology (Stachowiak and Batchelor)

Vol. 25 Thin Films in Tribology (Dowson et al., Editors)

Vol. 26 Engine Tribology (Taylor, Editor)

Vol. 27 Dissipative Processes in Tribology (Dowson et al., Editors)

Vol. 28 Coatings Tribology – Properties, Techniques and Applications in Surface Engineering (Holmberg and Matthews)

Vol. 29 Friction Surface Phenomena (Shpenkov)

Vol. 30 Lubricants and Lubrication (Dowson et al., Editors)

Vol. 31 The Third Body Concept: Interpretation of Tribological Phenomena (Dowson et al., Editors)

Vol. 32 Elastohydrodynamics – '96: Fundamentals and Applications in Lubrication and Traction (Dowson et al., Editors)

Vol. 33 Hydrodynamic Lubrication – Bearings and Thrust Bearings (Frêne et al.)

Vol. 34 Tribology for Energy Conservation (Dowson et al., Editors)

Vol. 35 Molybdenum Disulphide Lubrication (Lansdown)

Vol. 36 Lubrication at the Frontier – The Role of the Interface and Surface Layers in the Thin Film and Boundary Regime (Dowson et al., Editors)

Vol. 37 Multilevel Methods in Lubrication (Venner and Lubrecht)

Vol. 38 Thinning Films and Tribological Interfaces (Dowson et al., Editors)

Vol. 39 Tribological Research: From Model Experiment to Industrial Problem (Dalmaz et al., Editors)

Vol. 40 Boundary and Mixed Lubrication: Science and Applications (Dowson et al., Editors)

CONTENTS

INTRODUCTION v

REFEREEING ix

SESSION I **Keynote Address**

Tribology in Design 3
M J NEALE, OBE, FREng

SESSION II **Chemical Actions in Tribology**

Invited Paper : Tribochemistry: Friction-induced Lamellar Solids from Lubricant Additives 15
R LADAVIÈRE, J-M MARTIN, T Le MOGNE, B VACHER, B CONSTANS and S IOVINE

Assessing Friction, Wear and Film Formation Characteristics in Formulated Lubricants in Severe to Moderate Boundary Lubrication Conditions 23
A MORINA, A NEVILLE, J H GREEN and M PRIEST

Approaches to Sensitising Engine Valve Train Friction Models to Lubricant Formulation Characteristics 35
J H GREEN, M PRIEST, A MORINA and A NEVILLE

Electrochemical Control of Wear : A Third Body Approach 47
S MISCHLER

The Effect of Anti-Wear Additives, Detergents and Friction Modifiers in Boundary Lubrication of Traditional Fe-Base Materials 57
M REYES and A NEVILLE

SESSION Iil **Elastohydrodynamic Lubrication (1) - Dynamic Effects**

Invited Paper : Dynamic Effects in EHL Contacts 69
C J HOOKE

Comparison of Transient EHL Calculations with Start-up Experiments 79
M J A HOLMES, H P EVANS and R W SNIDLE

Comparison of Transient EHL Calculations with Shut-down Experiments 91
M J A HOLMES, H P EVANS and R W SNIDLE

Optical Interferometric Observations of the Effect of Moving Transverse Asperities on Point Contact EHL Films 101
M KANETA, N TANI and H NISHIKAWA

Mechanism of Dimple Formation under Elastohydrodynamic Conditions 111
K YAGI, K KYOGOKU and T NAKAHARA

SESSION IV Tribological Coatings - Advances in Surface Engineering (1)

Deformability of the Near Surface Layer of the First Body 123
A ELEÖD, L BAILLET, Y BERTHIER and T TÖRKÖLY

Thickness Dependency of the Lubricity of Ag Thin Films Deposited on Si (111)7 x 7 Surface 133
M GOTO, T YAMAMOTO and F HONDA

Sliding Wear Behaviour of Nanocrystalline Cr_3C_2-25(Ni20Cr) Coatings at Elevated Temperature 141
M ROY, A PAUSCHITZ and F FRANEK

SESSION V Tribological Studies of Dry and Lubricated Surfaces (1) - Dry

Modeling of Variable Friction in Contact with Fretting 153
M CHEIKH, S QUILICI and G CAILLETAUD

Micro-abrasive Wear of Semi-crystalline Polymers 165
Y J MERGLER and A J HUIS IN'T VELD

Wear Testing of Acetal for the Prediction of Polymer Rolling Bearing Durability 175
A F BEKE, R S DWYER-JOYCE, F R JONES and J GREGORY

Self-Lubricating Dry Rubbing Bearings - Fundamentals and Methods of Calculation 183
L DETERS, F MUELLER and M BERGER

SESSION VI Tribological Studies of Dry and Lubricated Surfaces (2) – Dry and Lubricated

Invited Paper : Tribological Design Criteria and Standards for Spacecraft Mechanisms 197
R A ROWNTREE

A Simple Numerical Procedure to Calculate the Input Data of Greenwood-Williamson Model of Asperity Contact for Actual Engineering Surfaces 205
E TOMANIK, H CHACON and G TEIXEIRA

The Study of Self-lubricating Sintered Nickel Matrix Material and its Applications 217
Z M LIU, Y ZHOU, Y ZHANG, Z GAO and E WANG

Behaviours of Ultra-thin Lubricant Films Between Solid Surfaces Having Atomic Scale Roughness 225
T OKUMURA, J SUGIMURA and Y YAMAMOTO

SESSION VII Bio-Tribology (1) – Hip and Knee Joints

Invited Paper : Fluid Film Lubrication in Artificial Hip Joints 237
Z M JIN, J B MEDLEY and D DOWSON

Wear of Alumina/Zirconia Nano-composite for Ceramic/Ceramic Joint Prostheses 257
K IKEUCHI, Y MORITA, K NAKATA, Y H KIM, T SEKINO and K NIIHARA

Surface Characteristics and Tribological Behaviour of New Zirconia Ceramics for Cementless Knee Joint Prostheses 263
Y SAWAE, T MURAKAMI, T SAWANO, I NODA and T SHIMOTOSO

The Relationship Between Steady-state Wear Rate and Theoretical Film Thickness in Metal-on-Metal Total Replacement Hip Joints 273
D DOWSON

On the Role of the Lambda Parameter in Simulator Wear of Metal-on-Metal Hip Implants 281
P PARÉ, J B MEDLEY, F W CHAN and S K YOUNG

Advances in Simulator Testing of Orthopaedic Joint Prostheses 291
T D STEWART, S WILLIAMS, J L TIPPER, E INGHAM, M H STONE and J FISHER

SESSION VIII Tribological Coatings - Advances in Surface Engineering (2)

Rolling Contact Fatigue of Nitrided 32CrMoV13 Steel 299
C JACQ, D NÉLIAS, G LORMAND, A VINCENT and G DUDRAGNE

Effect of Hf Addition on the Tribological Performance of ZrN Coatings 309
E ATAR, H MİNDIVAN, H ÇİMENOĞLU and E S KAYALI

Tribological Properties of TiAlSiN Film Coated by AIP Equipment onto SKD61 Substrate 315
R MURAKAMI, Y MATSUDA and D YONEKURA

Wear Mechanisms in Thermal Spray Steel Coatings 323
A EDRISY, T PERRY, Y T CHENG and A T ALPAS

SESSION IX **Friction, Wear and Life of Contacts in Machine Elements (1) – Life Predictions**

Invited Paper : The Role of Wear in Enhancing Rail Life 331
A KAPOOR, D I FLETCHER AND F J FRANKLIN

Laboratory-based Wear Life Testing of Sliding and Rolling Contacts – The First Step Towards Technology Transfer 341
T P SPERRING and B J ROYLANCE

Life Prediction for Surface Crack Initiated Contact Fatigue of Silicon Nitride Bearing Balls 349
Y WANG and M HADFIELD

Experimental Investigations on Rolling Contact Fatigue of Dented Surfaces Using Artificial Defects: Sub-surface Analyses 359
Y DIAB, S COULON, F VILLE and L FLAMAND

Incorporating a Railway Wheel Wear Model into Multi-body Simulations of Wheelset Dynamics 367
A WARD, R LEWIS and R S DWYER-JOYCE

SESSION X **Component Tribology (1) - Plain Bearings**

Invited Paper : Tilting Pad Thrust Bearings 379
S B GLAVATSKIH

Study on the Development of a Thin-film Sensor for Measuring Oil-film Pressure in Engine Bearing and Piston Pin-boss 391
Y MIHARA, K SEKINE and T SOMEYA

Optimisation of Static Behaviour and Stability Level of a Multilobe Journal Gas Bearing 403
G GRAU, I IORDANOFF, B BOU-SAÏD AND Y BERTHIER

SESSION XI **Bio-Tribology (2) – Biological Systems**

Invited Paper : Biotribological Approaches to the Lubrication of Engineering Systems 411
N D SPENCER, S S PERRY, S LEE, M MÜLLER, S PASCHE, S M De PAUL, M TEXTOR, X YAN and M S LIM

Eyelid/Lens Contact During a 100 metres Sprint. A Tribological Approach 417
B BOU-SAÏD

Biological Lubrication of Hydrated Surface Layer in Small Intestine 425
H YOSHIDA, Y MORITA and K. IKEUCHI

Tribology of a Snail (Terrestrial Gastropod) 429
A KOBAYASHI, I YAMAMOTO and T AOYAMA

SESSION XII **Tribological Studies of Dry and Lubricated Surfaces (3) – Lubricated**

Invited Paper : The Role of Tribology Research in the Development of Advanced Lubricants 439
A JACKSON and M N WEBSTER

Ultrasonic Characterisation of an Interference Fit 449
R LEWIS, M B MARSHALL and R S DWYER-JOYCE

SESSION XIII **Component Tribology (2) - Bearings and Mechanisms**

Tribologically-based Design of Precise Positioning Stage 461
K ADACHI, T YAMAGUCHI, Y ISHIMINE and K KATO

Journal Bearing Oil Film Measurement Using Ultrasonic Reflection 469
P HARPER, B HOLLINGSWORTH, R S DWYER-JOYCE and B W DRINKWATER

THD Lubrication of Mechanical Face Seals During Transient Period After Start-up : 2D Modelisation 477
J C DANOS, B TOURNERIE and J FRÊNE

SESSION XIV **Engine Tribology (1)**

Invited Paper : Modelling Lubricant Related Fuel Economy in Heavy Duty Diesel Engines 491
R CASTLE and S ARROWSMITH

A New Tribometer for Piston Ring Friction Measurements 501
P ANDERSSON

The Wear Response of Ceramic Matrix Nano-composite Coatings in Simulated Piston-Ring/Cylinder Wall Test Environments 511
K L DAHM, K PANAGOPOULOS and P A DEARNLEY

An Initial Investigation of the Lubricant Condition in the Automotive Ring Zone Under Cold Start Conditions 517
ANG DATOO and M F FOX

SESSION XV **Elastohydrodynamic Lubrication (2) - EHL - Mixed Lubrication**

Slip at the Wall - Evidence and Tribological Implications 525
H A SPIKES

Calculation of EHL Contacts in Mixed Lubrication Regime 537
A C REDLICH, D BARTEL and L DETERS

Some Insights into Lubricant Film Thickness in Mixed EHD Point Contacts 549
I KŘUPKA, M HARTL and M LIŠKA

On the EHL Film Breakdown 559
D ZHU

SESSION XVI Component Tribology (2) - Transmissions

Invited Paper : Gearing Tribology 575
R W SNIDLE, H P EVANS and M P ALANOU

Invited Paper : Tribology of Spline Couplings 589
A V OLVER, S MEDINA

Prediction of the Friction Coefficient Between Gear Teeth Using a Mixed Film Lubrication Model 603
J CASTRO, A SOTTOMAYOR, A CAMPOS and J SEABRA

Influence of Temperature Distribution on the Tribological Performance of Automotive Synchronisers 613
G POLL and M SPRECKELS

SESSION XVII Friction, Wear and Life of Contacts in Machine Elements (2) – Wear

Feasibility of Using Electrostatic Monitoring for Oil Lubricated Ceramic to Steel Sliding Contacts 625
L WANG, R J K WOOD, T J HARVEY, S MORRIS, H E G POWRIE and I CARE

High Temperature Abrasion Behaviour of Al / SiC_p Composites 637
H AHLATCI, E CANDAN and H CIMENOGLU

Friction and Wear of Spheroidal Graphite Cast Iron Under Severe Sliding Conditions 643
K HIRASATA, K HAYASHI and H MATSUNAMI

Wear Particles Produced in Repeated Contact of Two Asperities 653
J SUGIMURA, J W KANG, S TANAKA and Y YAMAMOTO

SESSION XVIII Elastohydrodynamic Lubrication (3) - Experimental and Theoretical

In Situ Pressure Measurements in an Elastohydrodynamically Lubricated Point Contact Using Raman Microspectrometry. Comparison with Numerical Calculations 663
I JUBAULT, J L MANSOT, P VERGNE, A A LUBRECHT and J MOLIMARD

Numerical Modelling of Thermal Effects in Elastohydrodynamic Lubrication Solvers 675
R FAIRLIE, C E GOODYER, M BERZINS and L E SCALES

Analysis of the Oil Distribution in Starved Elliptical E.H.D. Contacts 685
B DAMIENS, P M E CANN, C H VENNER and A A LUBRECHT

Ordinary Shear Thinning Behaviour and its Effect Upon EHL Film Thickness 693
S BAIR and F QURESHI

SESSION XIX Tribological Studies Of Dry And Lubricated Surfaces (3) – Lubricated

Invited Paper : A Design Tool for Selection and Optimization of Surface Finish in Mixed Lubrication 703
D ZHU

Starved Lubrication of Isoviscous Rigid Circular Contacts 713
M-P NOUTARY and A A LUBRECHT

Gaseous Evolution of Hydrogen from Hydrocarbon Oil and Grease Lubricated Contacts 719
C NEWLANDS, A V OLVER and N BRANDON

The Effect of Superimposed Low-frequency Modulation on Lubrication in Machining 727
E OLGUN, W D COMPTON and S CHANDRASEKAR

SESSION XX Engine Tribology (2)

Invited Paper : Optimisation of Piston Assembly Tribology for Automotive Applications 739
M PRIEST

Optimised Additive - Surface Material Combinations for Reduced Friction and Wear in Internal Combustion Engines 749
M VOONG, A NEVILLE and R CASTLE

Prediction of Total Engine Friction Power Loss from Detailed Component Models 761
Q ZHOU, I SHILLING and S H RICHARDSON

Experimental Evaluation of Engine Valve Train Friction Under Motored and Fired Conditions 767
R A MUFTI and M PRIEST

Inertia, Shear-Thinning and Thermal Effects on Connecting Rod Bearing Behaviour 779
B FANTINO and B BOU-SAÏD

SESSION XXI Physical and Chemical Actions in Tribology

Experimental Study on the Effects of Sliding Velocity and Roughness Orientation of Metal Surfaces on the Function of Lubricant Additives in Controlling Friction in a Boundary Lubrication Regime 791
M MASUKO, K TOMIZAWA, S AOKI and A SUZUKI

Extraordinarily Low Friction of the Diamond / H- and Ag- Terminated Si(111) Sliding System 799
H MASUDA and F HONDA

A Novel Experimental Analysis of the Rheology of ZDDP Tribofilms 807
C MINFRAY, J-M MARTIN, A A LUBRECHT, M BELIN and T Le MOGNE

SESSION XXII Tribological Studies of Dry and Lubricated Surfaces (4)

Invited Paper : Friction from Fluid-Film to Boundary Lubricated Conditions 821
R BASSANI and E CIULLI

A Macro-Micro- Fatigue Approach for Predicting Fatigue Life in Mixed - EHL Contact 835
D EPSTEIN, Q J WANG, L M KEER, H S CHENG, D ZHU, S HARRIS and A GANGOPADHYAY

Progression of the Stick / Slip Zones in a Dry Wheel-Rail Contact : Updating Theories on the Basis of Tribological Reality 845
E NICCOLINI and Y BERTHIER

Tribology of Electron Beam Excited Plasma Nitrided Tool Steels with Narrow Gap Features 855
P ABRAHA, H SHOYAMA, K TANIGUCHI and T HARA

WRITTEN DISCUSSION 861

LIST OF DELEGATES 911

SESSION I
KEYNOTE ADDRESS

Chairman : Professor D. Dowson

Paper I (i) **Tribology in Design**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 M. Neale. Published by Elsevier B.V. All rights reserved.

TRIBOLOGY IN DESIGN

Michael Neale OBE. FREng. FIMechE.

Neale Consulting Engineers Ltd.
Highfield, Pilcot Hill,
Dogmersfield, Hampshire. RG27 8SX.

ABSTRACT

This paper reviews the contribution that tribology can make to machine design. This is focussed on the three most common types of contact involving machine components. These are area contacts, as in plain bearings, concentrated contacts, as in rolling bearings, and distributed contacts with hard materials, as in abrasive wear. It indicates for each of these how basic understanding can lead to some innovative design solutions.

Tribological research itself is also an area where there are considerable opportunities for new methods to be applied, particularly in the rational analysis of extensive field experience data, which can lead to major advances in the understanding of the phenomena involved.

1. TECHNOLOGIES USED IN MACHINE DESIGN

Machines consist of an assembly of component parts, and these parts may be connected rigidly together, or may be arranged to allow movement relative to each other. The technologies which underlie this basic structure are therefore, stress analysis, for the strength and rigidity of the parts, zygology for the fixed joints, and tribology for the components which experience relative motion.

In addition, since the complete machine is required to perform some particular function, other technologies influence its layout. For example, fluid mechanics is applied to the design of pumps and turbines, heat transfer and combustion technology to engines, and electromagnetism to electric motors and generators.
In terms of reliability, tribology is the most critical, because the surfaces of machine components, which operate with relative motion, tend to carry a combination of high loads and sliding speeds. They therefore have the potential, in the event of any malfunction, for the release of large amounts of destructive energy, capable of inducing component failure. Also there is always the possibility of wear of the working surfaces, which can either lead to leakage, as in seals and piston rings, or to the reduced performance of load carrying lubricating films, as in plain bearings. Surface fatigue can also occur on components with concentrated contacts, such as rolling bearings and gears.

This paper attempts to review tribology, in terms of its potential impact on machine design, and the guidance that it can provide. The main areas of this impact are considered to be :-

a) An understanding of the contact conditions at the moving surfaces, the materials from which they can be made, and the action of any separating fluid films.

b) An understanding of methods for collecting information on the tribological performance of machines, such as the wear rates of their components in service, and their need for maintenance.

2. CONTACT CONDITIONS BETWEEN MOVING COMPONENTS

In most machines there are three types of contact with components that are subject to relative movement :-

a) A contact of substantial area, as in plain bearings and seals.
b) A concentrated contact of limited area, as in rolling bearings, gears, and cams and followers.
c) A distributed number of contacts with a hard material as in most abrasive wear situations.

2.1. Area contacts

In any interaction between two surfaces, which operate in direct sliding contact, there is always some risk of seizure between the two. This is particularly likely if the two materials in contact are similar in terms of hardness and melting point, in that any small local surface melting, at a high spot between the contacts, can more easily spread to the whole contact area. While similar materials are sometimes used for manufacturing convenience, as in drive chains between two sprockets, these will only work if the sliding speed between the components is very slow and intermittent, as it often is in that application. The surfaces of the pins and bushes also need to be kept well lubricated.

This type of technology does work adequately in applications such as bicycles, where the chain can be replaced fairly easily at low cost. However if chains are used in continuous service, as in the step drives of public service escalators, the use of steel pins in steel bushes is unsatisfactory. This is because, even if the chains are regularly lubricated, they are still prone to local wear and the resulting debris inhibits the access of new lubricant to the surfaces. This can then develop into major wear and possible seizure on any one bush. The real problem in this application, however, is that each escalator has several hundred chain bushes, any one of which can move into a problem mode at any time. This makes it virtually impossible to carry out effective maintenance at extended regular intervals, and forces the operation into a breakdown maintenance mode, with failures at random intervals. The tribological solution to this problem is to use two different materials in the rubbing contact. The use of reinforced plastic bushes containing solid lubricants eliminates the random failures and results in slow steady and predictable wear, giving a chain life of about 40 years. While it might be expected that this technology would be applied universally, it tends not to happen if the purchase team on a new construction is given the target of achieving the lowest construction cost, as distinct from the lowest total life cost. This is because a chain with plastic bushes tends to cost about 40% more than one without.

There are many other applications of plain bearings which operate at relatively low sliding speeds, and often with reciprocating motion, such as in vehicle suspension and steering, and in mechanical handling equipment. In all these applications, unlubricated non-metallic rubbing bearing materials can provide an optimum design solution. Typically these are composed of a relatively soft non-adhesive core material contained in a matrix of stronger material or reinforced by fibres of a strong material.They may also contain added solid lubricants such as graphite or molybdenum disulphide. An outline of their performance is given in

Figure 1. Since they do allow unlubricated operation, and the absence of oil can allow corrosion, it is important that the mating surface is made from a non-corrosive relatively hard material, such as a chromium plated or stainless steel shaft.

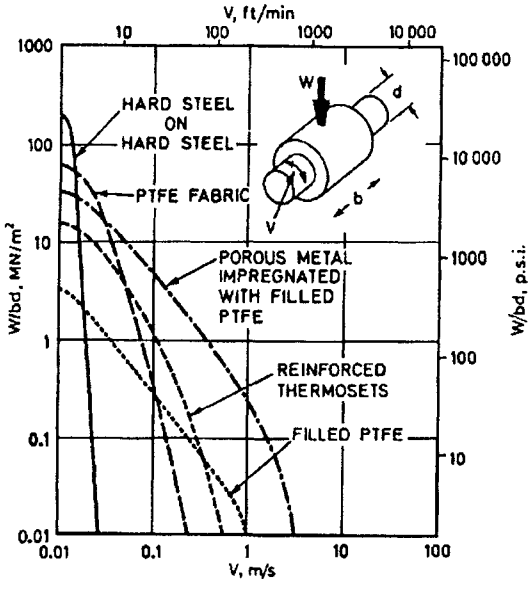


Figure 1. The application areas of rubbing bearing materials for oscillating movement

The major use of bearings with area contact is, however, in fully lubricated applications, such as in pumps, turbines, large gearboxes and electrical machines, and in internal combustion engines. In all these cases the tribological understanding of the operation of hydrodynamic films has made it possible to achieve an optimum fluid film plain bearing design. In the case of high speed rotating equipment this has included an understanding of rotor dynamics and its interaction with the bearing design, leading in many cases to the use of tilting pad journal bearings. In the case of engine bearings it is understood how their very high dynamic loads can be carried with adequate oil films, because their short duration is less than the time it takes to squeeze the oil from the bearing clearance. For the same reason, it is recognised that the process of refilling the bearing clearance with oil, at times of load reversal, is also critical and determines where oil feed drillings and grooves need to be positioned.

An understanding of the operating conditions of these bearings also indicates the kind of materials from which they should be made. For the steadily loaded bearings in large machines, the film thickness is small relative to the bearing width and the amount of misalignment that can occur due to shaft or housing deflections. This therefore requires a relatively soft bearing material to allow the shaft to bed into it locally without surface damage. However if it is too soft it could extrude sideways under load and it therefore requires some form of fibre reinforcement. These properties are provided by whitemetal, which consists of a soft tin core reinforced by a matrix of copper-tin needles. It is also common to add antimony to this material to create distributed tin-antimony cuboids, which by providing local harder particles in the surface, can increase the wear resistance of the material. The structure of whitemetal is shown in Figure 2.

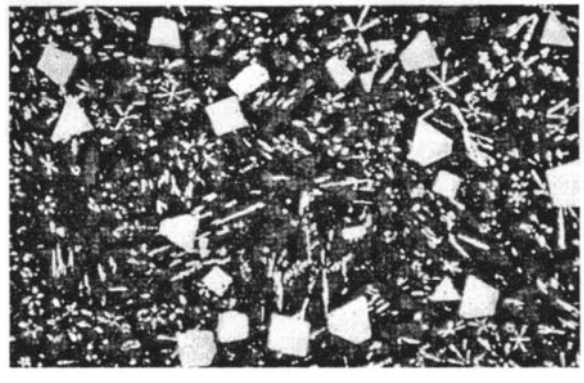


Figure 2. The structure of whitemetal

For high performance engine crankshaft bearings, with much higher loads, it is necessary to use a different type of material consisting of a strong matrix permeated with a softer material which can smear over the surface in the event of any shaft contact. Materials such as copper-lead, lead-bronze or aluminium-tin are typical materials of this kind. They are harder than whitemetal and

therefore usually require a hardened shaft in order to reduce shaft wear, or any damage in the event of a rubbing contact between the shaft and the bearing.

While the development of plain bearings has mainly related to the use of oil lubrication, the fact that the operation of the fluid films is largely understood, makes it possible to lubricate machines with liquids other than oil. This is particularly significant for machines such as pumps and turbines which already operate with some kind of fluid, and if this fluid is used to lubricate the bearings, some significant design improvements can be achieved. A good example of this is in boiler high pressure water circulation pumps where a novel and effective design can be achieved by using the water to lubricate the bearings. This not only eliminates the need to have shaft seals but allows the whole pump, and its driving motor, to be contained in a static pressure vessel, to which the only connections are the input and delivery pipes, and electrical power leads. With no exposed shafts there is no possibility of leakage.
There are, therefore some exciting possibilities for producing better and revolutionary machine designs and encouraging some free thinking by designers about component dimensions and layouts. In machine design, the shaft diameter is usually determined by its required torque capacity or by its required lateral stiffness, to avoid excessive deflection or too low a dynamic flexibility. With oil lubrication and a typical operating viscosity of 10 to 20 centipoises, the required bearing length is slightly less than the shaft diameter. This has resulted in established machine layouts, which provide the starting point for most designs. However, when designing for the use of different fluids for lubrication, it is important to remember that the bearing design equations indicate a link between the lubricant viscosity and the required shaft diameter.

In hydrodynamic bearings the simplest non-dimensional factor for the operating conditions, which relates to the geometry of the lubricant film, ie. the eccentricity ratio is proportional to :-

$\frac{ZN}{P}$ Where Z = lubricant viscosity
N = rotational speed
P = bearing load pressure
$= \frac{W}{DL}$ W=bearing load
D=bearing dia.
L=bearing length

When designing a bearing with a different lubricant, the rotational speed and the load do not change. ie. N and W are constant. So the lubricant film geometry is proportional to ZDL. For the same film geometry, DL needs to increase by the same proportion that Z is decreased. If the bearing geometry, in terms of D/L remains the same, then:-

ZD^2 should be constant, ie. D should be proportional to 1/ square root of Z

Figure 3. shows what this means in terms of the possible redesign of two machines, a gas turbine and a fuel pump, which with normal lubricating oils operate with shaft diameters of 100mm and 25mm. An interesting conclusion from this, is that it might be possible to design a gas turbine with air bearings outside the rotor so that the turbine blades were always in compression and therefore able to operate at higher temperatures and higher efficiencies. In fact this diagram does tend to slightly exaggerate the required shaft diameter, because it assumes a constant eccentricity ratio, which corresponds to an increase in minimum film thickness in proportion to the shaft diameter. In practice although the allowable minimum film thickness will need to increase slightly at higher shaft surface speeds, it will not be as much as this, so the optimum shaft diameters are slightly exaggerated.

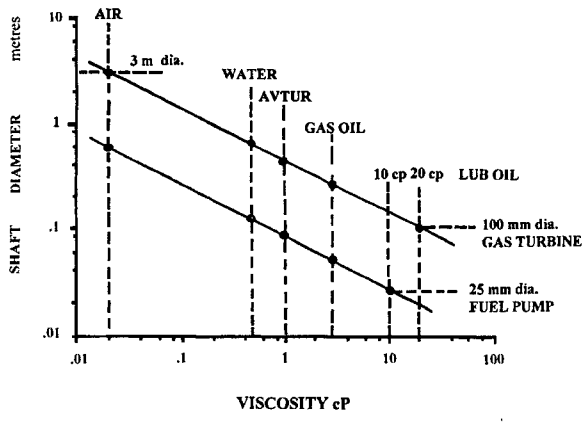


Figure 3. The variation of optimum shaft diameter with lubricant viscosity

The bearing materials commonly used in machines lubricated with fluids other than oil are often similar to the materials used for dry rubbing bearings or in the case of water lubrication may be ceramics which tend to operate well with water. Rubber-like materials can also be used as these tend to deflect in the higher pressure parts of the hydrodynamic film and provide closer clearance sealing zones around the edges.

2.2 Concentrated contacts

The other common type of tribological contact is where the load is transmitted across a small concentrated contact area, as in rolling bearings and gears. The surfaces therefore need to be hard to carry the loads without permanent deformation, and the material needs to have the minimum number of internal defects to reduce the risk of fatigue failure. The lubrication mechanism is elastohydrodynamic which means that there is some elastic flattening, and local concavity, of the surfaces, which traps the lubricant in this zone of very high pressure. Under these conditions typical lubricating oils increase their viscosity to the point of becoming semi-solid, and produce an effective separating film between the surfaces. It is important, however, for designers to remember that this mechanism only works with fluids, such as lubricating oils which have a high pressure coefficient of viscosity. Other fluids such as water do not have these characteristics, so there is a more restricted choice of working fluids that can be used with component contacts of this kind.

Tribology has also provided an understanding of some of the wider aspects of the methods of operation of concentrated contacts. Gear teeth, for example, operate with patterns of combined rolling and sliding, as shown in Figure 4.

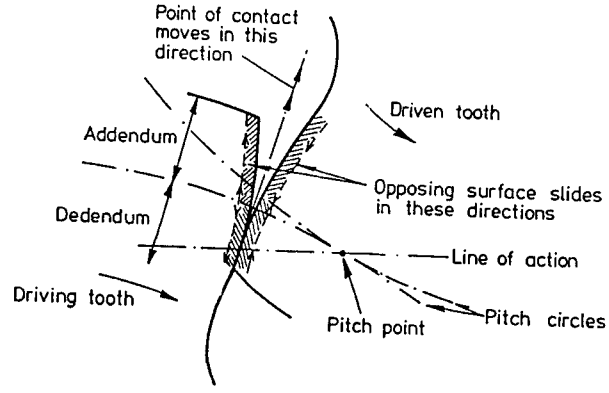


Figure 4. The contact conditions between gear teeth

The sliding tends to drag the surfaces to form inclined cracks, and if the rolling component of the movement is in the same direction, it will tend to push out any oil that gets into these cracks. If, however, the rolling is in the other direction, the oil will be pushed into the crack and will tend to extend it and generate surface failure by lifting out portions of the surface. It can be seen from Figure 4. that this situation arises in the dedendum of the teeth, and explains why pitting of the teeth usually occurs in these areas.

Another rolling contact application where tribological knowledge can help in understanding the detailed nature of the contact, is in guide wheels and rollers used to support large rolling structures. For example, large astronomical telescopes, are often

mounted on rotary tables with wheels about 1 metre diameter. If these wheels do not track properly, large sideways forces can be generated, which can damage the structure, or can cause shocks, which disturb the telescope settings, when the wheels jump sideways to correct their position. This problem arises because cylindrical wheels made from a relatively rigid material, such as metal, will always roll along a line exactly at right angles to their axis of rotation, as shown in Figure 5. It is very difficult, if not impossible, to set them to the required standard of tracking accuracy.

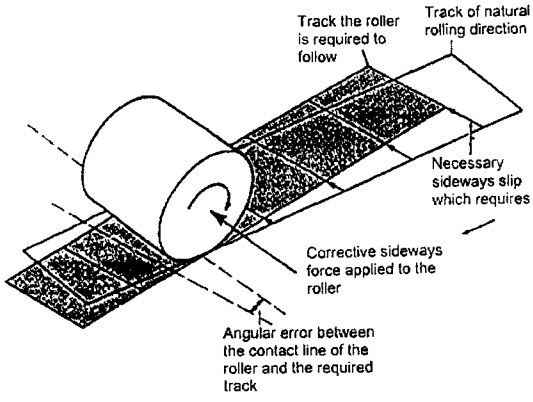


Figure 5. The track patterns of rolling wheels

Once this problem is understood, design solutions can be devised, such as the use of a central location bearing, and wheels, at a larger track diameter, which have a spherical outer profile, and bearings which allow the wheels to tilt and thus steer themselves back to the correct track, under the influence of any sideways forces that begin to be generated.

2.3. Distributed contacts with hard material.

A major cause of wear of the parts of machines arises from their contact with abrasive materials. These materials are typically very hard and may also be sharp. The general tribological solution to this component wear problem is to make the components as hard as possible. However, this can result in them being brittle and therefore liable to fracture. It is therefore generally better to try to obtain a hard surface on a tough core This can be achieved by conventional surface hardening processes or by the use of materials which work harden when subject to surface impacts and deformation. These solutions are particularly effective when the abrasive material is primarily moving along the component surface as in ploughing tools, or the inside of pipes carrying abrasive slurries. If however the movement is more at right angles to the surface due to the direction of the local flow, it is generally better to use an elastic, rubber-like, material which can cause the abrasive to bounce off the surface with the minimum damage.

Another common type of problem with abrasive wear occurs when abrasive material gets into the working surfaces of plain bearings. The type of bearings which commonly suffer from this problem are pivot pins and bushes operating in an abrasive environment, as on construction equipment. Both the pin and its mating surface are generally made from iron or steel to obtain adequate strength and the abrasive tends to embed in both surfaces and result in their mutual abrasion and damage. This type of wear is particularly sensitive to the relative hardness of the two surfaces, as shown in Figure 6. It can be seen from this that it is essential to avoid the use of materials of a similar hardness for the pin and its mating surface. The wear resistance is greatly increased if one surface, usually the pin is made as hard as possible, and the other left relatively soft. The abrasive dirt then embeds in the softer material, protecting it from wear, and attempts to wear the very hard mating surface. This results in much lower total wear rates.

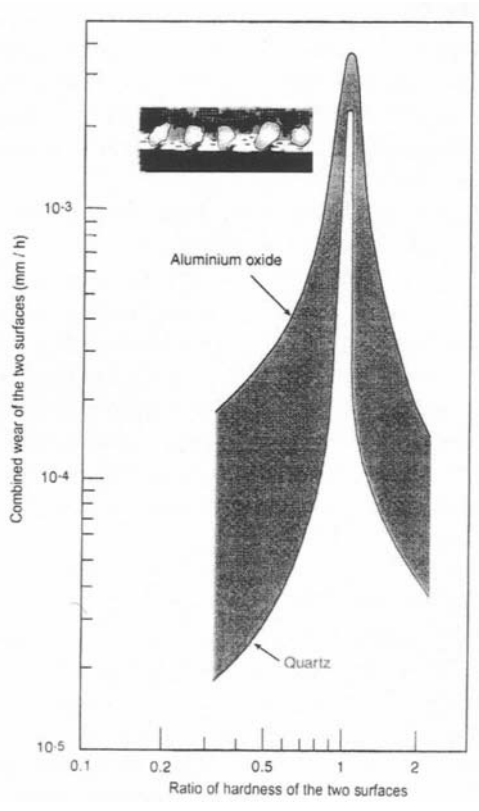


Figure 6. The wear rate of two surfaces with abrasive particles trapped between them

3. THE TRIBOLOGICAL PERFORMANCE OF MACHINES

When machines are in operation most of the time related features, such as maintenance intervals and machine life, are determined by the tribological components. The main deterioration processes tend to be wear and fatigue, which usually take hundreds or thousands of hours to produce critical levels of deterioration. Because of this situation, there are two areas where tribologists can make original and useful contributions:-

a) To develop new methods of research which concentrate on the collection of data points from the field, as the only real source of a lot of this information

b) To analyse machine performance from a tribological point of view in order to determine patterns of machine deterioration.

3.1 New research methods

There is a tendency for tribological research to become concentrated on detailed work in laboratories and to regard the collection of data from the field as low grade and empirical because the data does not derive from a controlled experiment. This probably arises from peer review pressure from scientifically oriented reviewers of research projects. It is, however, worth noting that in the social sciences, the analysis of data from the field is quite normal.

In research on wear, in a laboratory, it is difficult to carry out tests which relate to real situations, because of the need to accelerate the tests in order to match practical time scales. It can be argued that such work is really a waste of time, other than to obtain a few reference points, when there is a large amount of real data out in the field waiting to be collected. It is of course true that any one data point, from the field, or even a small group of them can be unreliable, particularly because the test conditions are not precisely defined. But if enough data is collected, the overall trends can be seen clearly, and a range of useful numerical values can be determined.

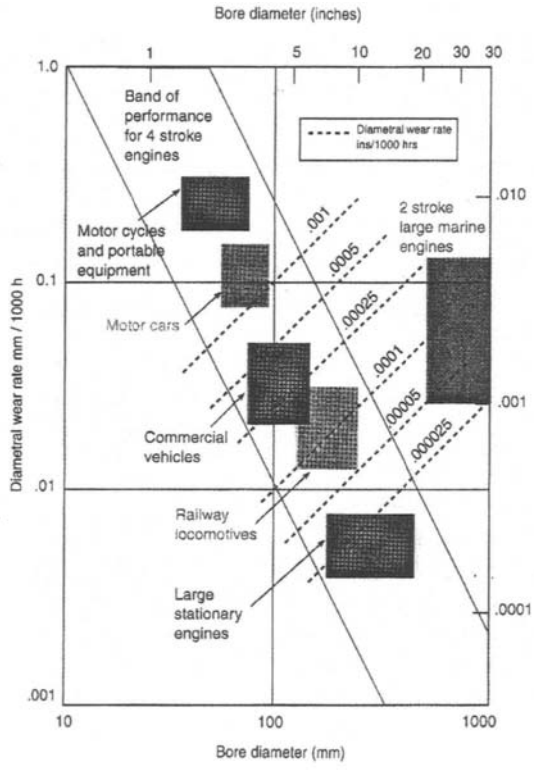


Figure 7. The wear rate of the cylinders of internal combustion engines

An example of some work of this kind is shown in Figure 7. which gives a clear indication of the patterns of cylinder liner wear in internal combustion engines, in normal service. It is particularly interesting that the lowest rates of wear occur on the larger engines, which indicates that it is primarily determined by the number of times that the piston rings visit the top dead centre position, and this is less on larger engines because they run more slowly. It is also clear why modern very high speed motorcycle engines tend to have a limited life between overhauls. This type of work can be supported by the examination of specimens of the liners and rings collected in the field to note any significant material or detail design effects, and link in to the metallurgical factors involved. Apart from providing an insight into the wear process in these engines the results also give a useful reference to the performance to be expected from the range of engines. This study involved the collection of over 600 data points from engine manufacturers and users all over the world. They provided the data willingly in exchange for a promise that they would receive a copy of the final result, in which only they would be able to recognise their own performance from the large mass of data points. It may be worth pointing out that any research student doing this type of work can find the process very interesting and instructive and can include a reasonable amount of foreign travel. The student will also be extremely employable as soon as the work is over.

There are plenty of opportunities for work of this kind and in 2002 perhaps the most obvious would be a similar study of wheel/rail interactions on the worlds railways. The main variables would probably be the radius of the curves, the axle loads, the wheel profiles, the wheel and rail materials, the degree of steering damping on the bogies, and the frequency of traffic. There are hundreds of data points out there waiting to be collected, and the patterns could then be matched to the results of more fundamental work on the tribological conditions in contacts of this kind, and could give them full validation.

3.2 Machine deterioration analysis

There is another major contribution waiting to be made by tribologists, and this is to understand the trend patterns of progressive deterioration of machines. This is a tribological study because most of the deterioration mechanisms are tribological. At first sight this might sound rather boring, but it has tremendous significance. This is because, while some machines deteriorate immediately if their maintenance attention is reduced, there are many others with longer response times, which may be up to four years. A high quality study and analytical

report on this subject could become a major international reference document.

The reason for this is that current trends in company management place the main emphasis of performance review on the results obtained over one year, to match the standard period for financial reports. If a company owns and operates a large amount of engineering equipment (such as a railway) an examination of its accounts by the management will show that a large amount of the company's expenditure is being spent on maintenance. It may be concluded that this expenditure is excessive and the almost standard way of dealing with this is to reduce the maintenance budget by about 15% and see whether this has any effect. If the equipment is robust, which in many industries it is, there will predictably be no effect. A further reduction will therefore be tried the following year with probably only minor effects. This is the point where the board members should move on using their excellent record of increased profitability.

Then after a further one or two years, disaster will strike, possibly with fatalities, and with a very large recovery cost, greatly exceeding the total of any savings that had been made. In any official enquiry, the board will claim that they acted sensibly in the interest of the shareholders, and no reference data was available to prove otherwise. An effective way of preventing this situation in the future is to produce the reference data, and the threat of corporate liability will then ensure that sensible decisions are made, with engineering advice.

4. CONCLUSION

This paper has attempted to review the contribution that tribology can make to machine design and operation. It is an extremely important contribution and can involve very interesting and challenging work for those involved. To achieve it, however, tribologists must lift their horizons above work such as the study of two high spots in contact, however important that may be. They should apply their skills to the large amount of internationally important and intellectually challenging work, on practical problems, that is out there waiting to be done. It is to be hoped that this paper may have given some leads on how to do this, and there are a number of excellent papers being presented at this conference, which are moving in this direction.

This Page Intentionally Left Blank

SESSION II
CHEMICAL ACTIONS IN TRIBOLOGY

Chairman :	Professor L Flamand
Paper II (i)	**Tribochemistry: Friction-induced Lamellar Solids from Lubricant Additives**
Paper II (ii)	**Assessing Friction, Wear and Film Formation Characteristics in Formulated Lubricants in Severe to Moderate Boundary Lubrication Conditions**
Paper II (iii)	**Approaches to Sensitising Engine Valve Train Friction Models to Lubricant Formulation Characteristics**
Paper II (iv)	**Electrochemical Control of Wear : A Third Body Approach**
Paper II (v)	**The Effect of Anti-Wear Additives, Detergents and Friction Modifiers in Boundary Lubrication of Traditional Fe-Base Materials**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribochemistry: Friction-induced Lamellar Solids from Lubricant Additives

R. Ladavière[a], J.M. Martin[a], T. Le Mogne[a], B. Vacher[a], B. Constans[b] and S. Iovine[b]

[a] Ecole Centrale de Lyon, UMR 5513, F-69134 Ecully (France)

[b] TOTALFINAELF, Centre de recherche de Solaize, BP 22, 69360 Solaize (France)

In this work, we have evidenced the formation of a lamellar solid from the tribochemical reaction of a borated additive and a succinimide additive. The result is the formation of h-BN in the tribofilm. The tribofilm is mainly composed of an amorphous borate matrix containing highly-dispersed h-BN nanoparticles in the form of sheets less than 10 nm wide and 5 nm thick. The originality of this h-BN tribochemical reaction lies in the nature of the intermolecular reaction between the two additives. At the opposite, MoS_2 formation from MoDTC is the result of intramolecular chemical reaction.

1. INTRODUCTION

MoS_2 (molybdenum disulfide) and h-BN (hexagonal boron nitride) are well known lamellar solids used in solid-state lubrication. Generally they are used as a powder or in the form of surface coatings. Recently, it has been shown that molybdenum organic compounds can generate MoS_2 material by a tribochemical reaction under boundary lubrication [1]. For example molybdenum dithiocarbamate (MoDTC) is currently used in engine oils as a friction reducing additive in the ultralow regime (friction coefficient well below 0.1). High Resolution Transmition Electronic Microscope (HRTEM) and surface analyses have shown that MoS_2 material formed by tribochemical reaction of MoDTC was dispersed at the nanometer scale in a phosphate matrix and that lubrication was ensured by MoS_2 single sheets [1]. Grossiord [2] has shown the formation of MoS_2 by tribochemical reaction of molybdenum dithiophosphate and molybdenum dithiocarbamate. Wahl *et al* also depicted the formation of MoS_2 by friction of Pb-Mo-S amorphous ternary phase [3].

Borated additives are commonly used in transmission oils for their excellent extreme-pressure action and good antiwear function. Many studies have shown that the action mechanism of borated additives is due to the formation of a protective film of borates by electrodeposition [4], [5]. On the other hand, nitrogenous additives (like succinimides) are often used as dispersants in most of lubricants.

Qiao [6] has evidenced an intramolecular reaction between boron and nitrogen forming boron nitride. However, in this study, only the chemical bond between boron and nitrogen has been shown and not the lamellar structure of h-BN. Friction-induced bidimensional films would be generalized if lamellar boron nitride is formed in presence of borated and nitrogenous additives. The synthesis of boron nitride by conventional method, for example by chemical vapour deposition (CVD) is very difficult and requires drastic conditions of temperature and pressure. These conditions could be obtained in the contact area during sliding for the tribochemical reaction.

In this work, we investigate the formation of hexagonal boron nitride (h-BN) from the tribochemical reaction between a borated additive and a nitrogenous compound. We developed a multi-technique approach including X-Ray Photoelectron Spectroscopy (XPS), X-ray Absorption Near Edge Structure Spectroscopy (XANES), Electron Energy Loss Spectroscopy (EELS) and high resolution TEM (HRTEM).

2. EXPERIMENTAL

2.1. Tribological conditions

Tribochemical films were obtained from bench tests evaluating the lubricating capacities of gear oils. The friction pairs used are a synchronizer ring made of brass (CuZn36Mn3Al2Si) and a gear cone made of steel (27MC5). In the test, the lubricant and friction pairs are heated to 90°C. The maximum

contact pressure is about 20 MPa and the sliding speed is 1.6 m/s. The duration of the friction test is approximately 20 hours.

The base oil is a mineral oil containing a little amount of synthetic base. Two different additives are used: a borated additive (potassium triborate) and a nitrogenous additive (succinimide). Three different formulations were compared: the base oil alone (test A), the base oil containing the borated additive (test B) and finally the base oil in presence of the two additives (test C). At the end of the tests, the surfaces of wear scars were analysed by XPS and XANES and wear debris from the worn lubricant were analysed by EELS and HRTEM, respectively.

2.2. Analyses

XPS analyses were performed in a VG ESCALAB 220i apparatus with the aluminium Kα source. Before XPS analysis, the surfaces were degreased by several immersions in pure heptane in order to eliminate the residual oil. Special attention has been paid here to the analysis of B1s and N1s photopeaks. In order to have well resolved peaks and a good signal-to-noise ratio, we only used the energy range between 189 eV and 200 eV for B1S and between 396 and 407 eV for N1s. Each region is scanned at least 10 times with an acquisition time of 1000 ms/eV. The energy resolution at the B1s and N1S peaks is 0.1 eV. Pure h-BN powder (supplied by MERCK) was analysed with the purpose to obtain standard values for quantitative analysis of boron and nitrogen and to measure chemical shifts of B1s and N1s photopeaks in their respective chemical bonding. The succinimide material was deposited on a polished AISI 52 100 steel surface after evaporation of the solvent. The deposit formed was analysed in order to identify the different contributions of nitrogen in the N1s peak.

XANES was performed at the Canadian Synchrotron Radiation Facility (CSRF) which is located at the Aladdin 1 GeV storage ring, University of Wisconsin, Madison, USA. The Grasshopper beam line was used in this investigation. The region studied is the B-K edge. The X-ray beam is monochromatized and covered the photon energy range of 70-900 eV. The energy resolution of this beamline is 0.2 eV at 200 eV. The B-K edge spectra were recorded by Total Electron Yield (TEY) that measured the sample current produced due to excitation of all electrons of different energies by incident X-Ray synchrotron radiation. Pure h-BN and pure B_2O_3 were analysed in order to study the B-K edge and to have standard spectra.

Wear debris were collected carefully from the used oil. After centrifugation, they were rinsed in pure heptane in an ultrasonic bath. We used a lacey carbon film mounted on a copper film, covered itself by a very thin carbon film (approximately 5 nm thick) which is particularly suitable for high-resolution TEM and EELS. Wear particles are deposited on the carbon film. They were observed in the analytical TEM (PHILIPS 420EM) equipped with Electron Energy Loss Spectroscopy (EELS) and in the analytical High Resolution TEM (JEOL 2010F HRTEM) operating at 200 kV accelerating voltage.

3. RESULT AND DISCUSSION

3.1. Friction results

The tests with the three formulations gave similar steady-state friction coefficient of about 0.12. So, the binary additive combination did not provide any friction reduction. However, friction fluctuations were significantly decreased when using the binary mixture. This smoothing effect on the friction has already been observed in the literature when h-BN powder was added in a base oil [7].

3.2. Analyses

In the XPS study, only the steel surfaces (cones) were analysed because the contact area was not large enough on the brass ring. The decomposition of the B1s and N1s photopeaks obtained with the three lubricants are compared to those of pure hexagonal boron nitride (h-BN). The results are shown in Figure 1 and Table 1. In order to avoid problems of charge correction, the difference of binding energy (δ) between B1s and N1s is given in the table 1 and will be considered as a reference for the existence of h-BN. Concerning the atomic B/N ratio, it has been calculated by using the equation of Wagner using Wagner sensitivity factors [8].

In the case of the base oil alone (test A), as expected, we did not observe any boron or nitrogen on the steel surface (not shown in Figure 1). In the test with only borate as an additive (test B), no nitrogen is found on the surface and the B1s peak can be well fitted with only one contribution corresponding well to the B-O bond (as in B_2O_3).

Concerning the test with the two additives (test C), our results show the presence of both nitrogen and boron in the tribofilm composition.

In order to correctly fit the experimental spectrum, three contributions are needed for the N1s peak and two contributions for the B1s peak. The comparison of binding energies of these peaks with those of standards shows that the tribofilm is mainly composed of borate material (B-O binding).

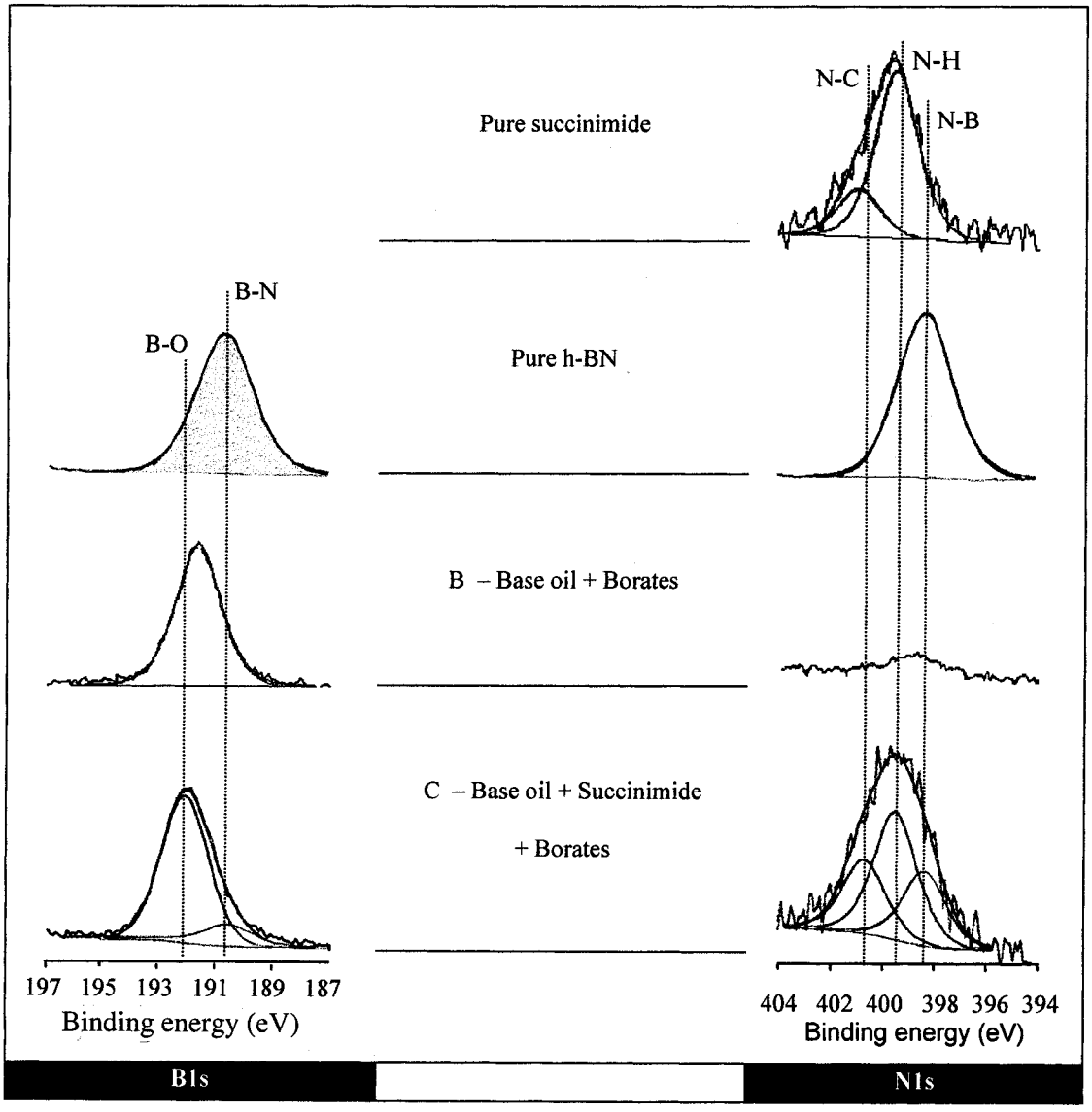


Figure 1 : XPS spectra of B1s and N1s peaks for Pure Succinimide, Pure h-BN, B - Base oil + Borates and C - Base oil + Succinimide + Borates

	B/N ratio	δ (eV)
Pure h-BN	B/N = 1,0	δ (B-N) = 207,6
Base oil + Succinimide + Borate	B/N = 1,2	δ (B-N) = 207,7

Table 1 ::B/N atomic ratio and difference of binding energy obtained by XPS for pure h-BN and for the oil (C)

Other expected contributions are observed, particularly the N-C bond (at 400.6 eV) and N-H bond (at 399.3 eV) due to residual nitrogenous additive present in the tribofilm material. However, the third peak in the N1s photopeak is attributed to N-B binding (399.3 eV) and can be well correlated with the second peak in the B1s XPS spectrum. Both the energy values and the energy difference between these peaks are characteristic of the presence of boron nitride chemical bonding [9]. From the calculation of atomic ratio B/N, the stoichiometry of the BN compound is very close to pure h-BN powder. (see Table 1).

From this XPS study of both B1s and N1s photopeaks in the tribofilm, B-N binding is detected in the tribofilm material but in a relatively small proportion (a few atomic percent).

(XANES) is known to be a sensitive technique to detect small changes in the chemical environment at low-Z elements. This capability is due to the use of the synchrotron radiation. This tool allows the sp^2 and sp^3 molecular hybridizations to be distinguished and therefore XANES is useful to identify the molecular structure of a chemical compound [10].

First we study the comparison between XANES analyses of the tribofilm of test C and two standard materials h-BN and B_2O_3. The spectra at B K-edge are displayed in Figure 2. The fine structures of these absorption edges show the presence of peaks due to transitions of the 1s electrons into π^* and σ^* unoccupied states. The presence of π^* fine structures is characteristic of the sp^2 hybridization of boron [11]. For pure B_2O_3 and pure h-BN, $1s \rightarrow \pi^*$ transition peaks (at energies of 194.4 eV and 192.4 eV respectively) are attributed to the trigonal hybridisation of boron in these compounds. The spectrum corresponding to analysis of the tribofilm (test C) shows effectively the presence of both peaks at 192.4 eV and at 194.4 eV. Therefore, XANES gives strong evidence that the tribofilm is composed of both borates (B-O) and boride (B-N) bound in sp^2 hybridization. However, no indication on the crystal structure or long-range order is given by XANES.

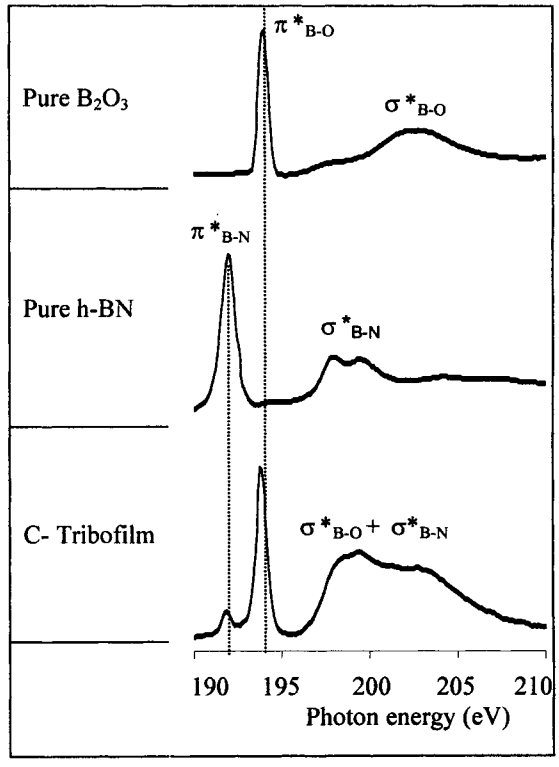


Figure 2 : XANES Spectra of B-K edges on :
a) Pure B_2O_3, b) Pure h-BN and c) Tribofilm of Base oil + Succinimide + Borate (C)

In order to show that nanoparticles of h-BN are present both in wear debris and in tribofilm, we used EELS spectroscopy associated to TEM investigation. In a similar way as XANES spectroscopy, the presence of π^* peak in the EELS spectrum is a fingerprint of sp^2 hybridized boron. EELS and XANES techniques give similar data except that the energy resolution is better for XANES than for EELS using a LaB_6 electron source. EELS spectra at boron K-edge for pure h-BN, pure B_2O_3 and wear particles (test C) are shown in Figure 3. As in XANES, EELS analysis of wear debris clearly shows the presence of π^* peaks corresponding to both h-BN and B_2O_3 molecular structures at 192.4 eV and 194.4 eV, respectively. Therefore, we give strong evidence for the presence of boron nitride in the wear debris. Once B-N binding detected by EELS, HRTEM is used here to determine the crystal structure of boron nitride species and the distribution of this compound in the matrix. HRTEM micrographs are shown in Figure 4-a. The fringes image and corresponding nanodiffraction pattern clearly show, in some place, the presence of the hexagonal lattice of h-BN basal plane by comparison with pure h-BN powder (Figure 4-b). The measured inter reticular distance (2.16 nm) corresponds well to the distance for (100) and (010) inter-reticular planes in hexagonal boron nitride. From HRTEM study, it can be concluded that h-BN lamellar nanoparticles are oriented with the banal plane parallel to the flaky debris and then probably to the sliding direction.

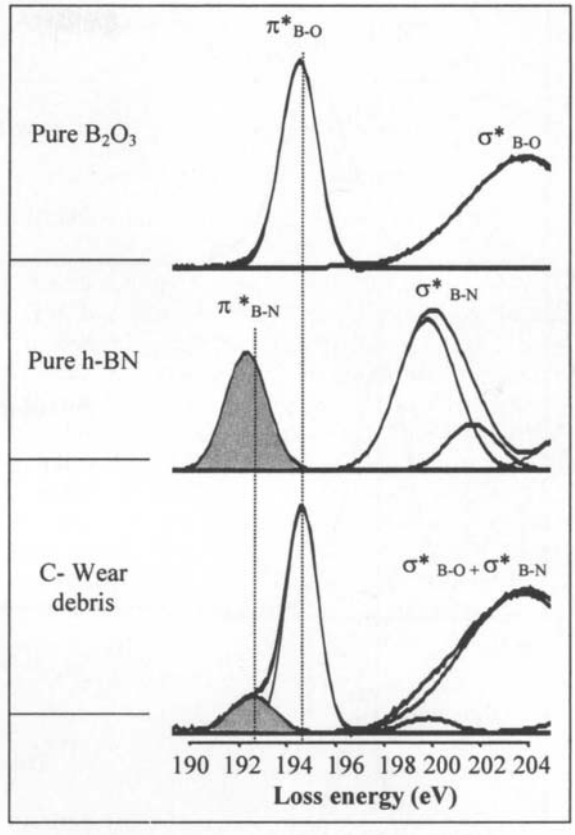
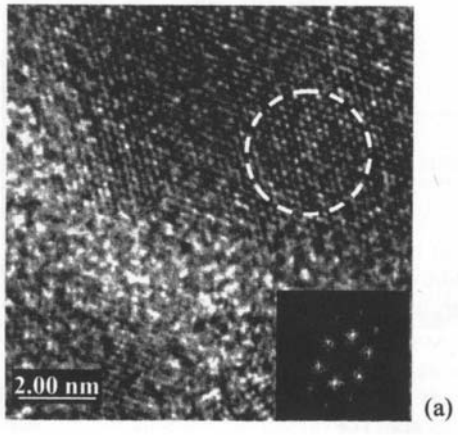


Figure 3: EELS Spectra of B-K edges on :
a) Pure B_2O_3, b) Pure h-BN and c) Wear debris of Base oil + Succinimide + Borates



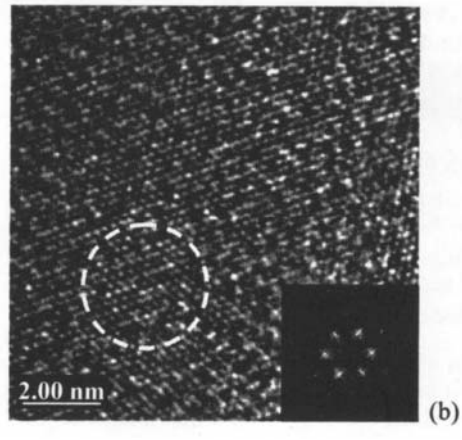


Figure 4 : HREM picture and nanodiffraction of :
a) wear debris of Base oil + Succinimide + Borates b) pure h-BN

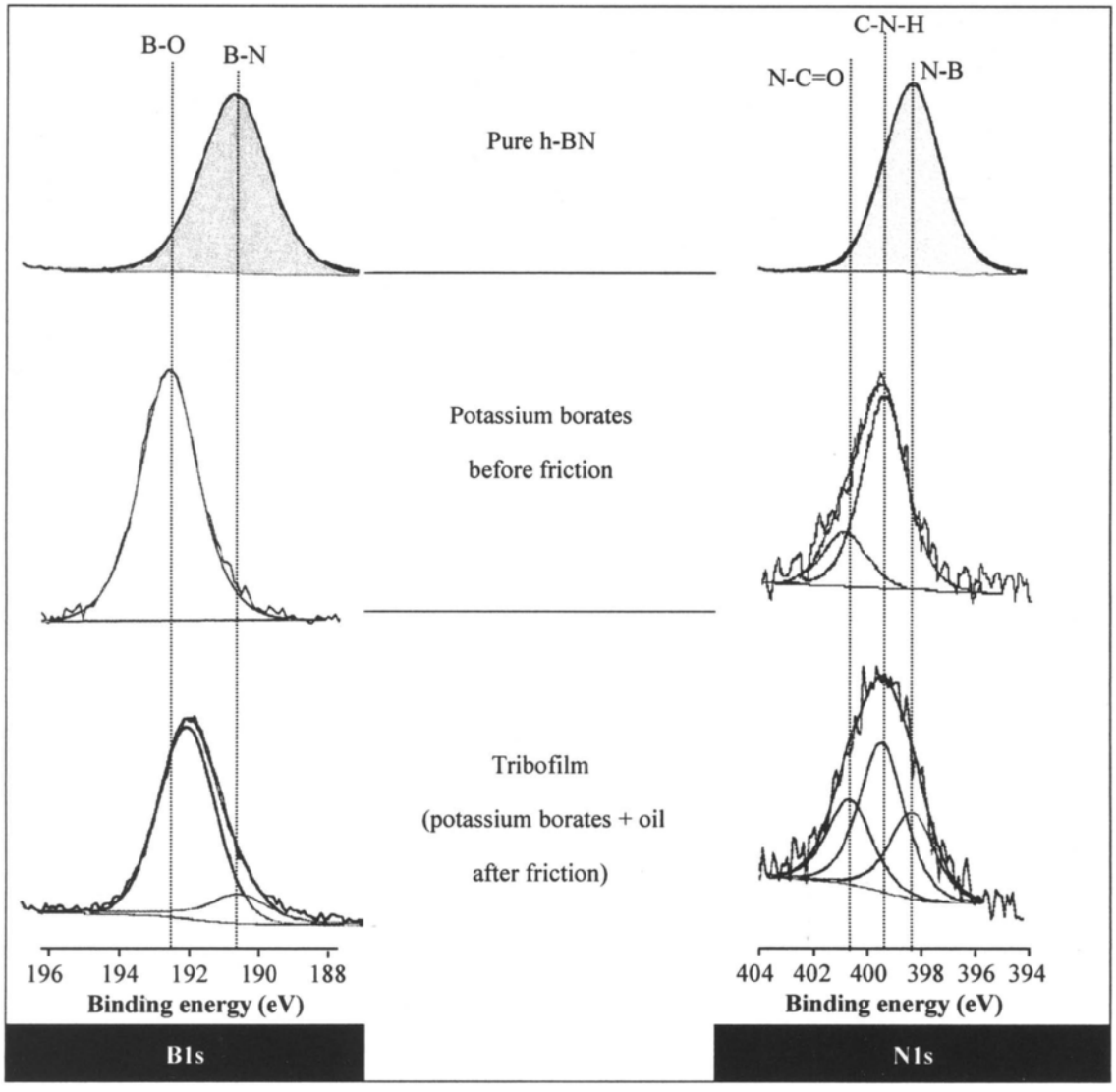


Figure 5 : XPS B1s and N1s peaks. Spectra "Potassium borate before friction" recorded on coating of this additive deposited on the steel 100C6.

According to the TEM image and the EELS analysis, h-BN appears as highly-dispersed nanoparticles embedded in a borate matrix.

3.3 Origin of hexagonal boron nitride

Before it can be concluded that h-BN compound is triboformed, we have to be sure that this material does not exist in the lubricant before the friction test. According to the additive manufacturer, no BN has been added during the synthesis of potassium borate additive. Moreover, the chemical synthesis of boron nitride is very difficult because it requires drastic

conditions that are never obtained during the storage period. Thus the chemical formation of BN inside the lubricant is improbable. We performed XPS and ToF-SIMS analyses of the original borate nanoparticles to confirm this point.

The comparison of B1s and N1s XPS peaks for pure potassium borate additive, tribofilm of test C and pure h-BN is shown in. No BN bonding is detected in the spectra of pure potassium borate, only B-O (borate), N-C=O and N-C-H (succinimide molecules) are detected. We deduce that h-BN is not present (or below 0,1 %) in the lubricant.

ToF-SIMS spectra for pure potassium borate additive and h-BN powder are compared in Figure 6. Two different mass regions were observed: the first between 35,96 and 36,06 amu and the second between 36.96 and 37,06 amu. These two regions correspond to B_2N^+ and B_2HN^+ ions, respectively. Results show that boron nitride is not present in the pure borate additive. This remark is valuable for the ppm quantity.

Thus, the h-BN evidenced in the tribofilm composition is effectively *in situ* generated by the friction process (so-called tribochemical reaction).

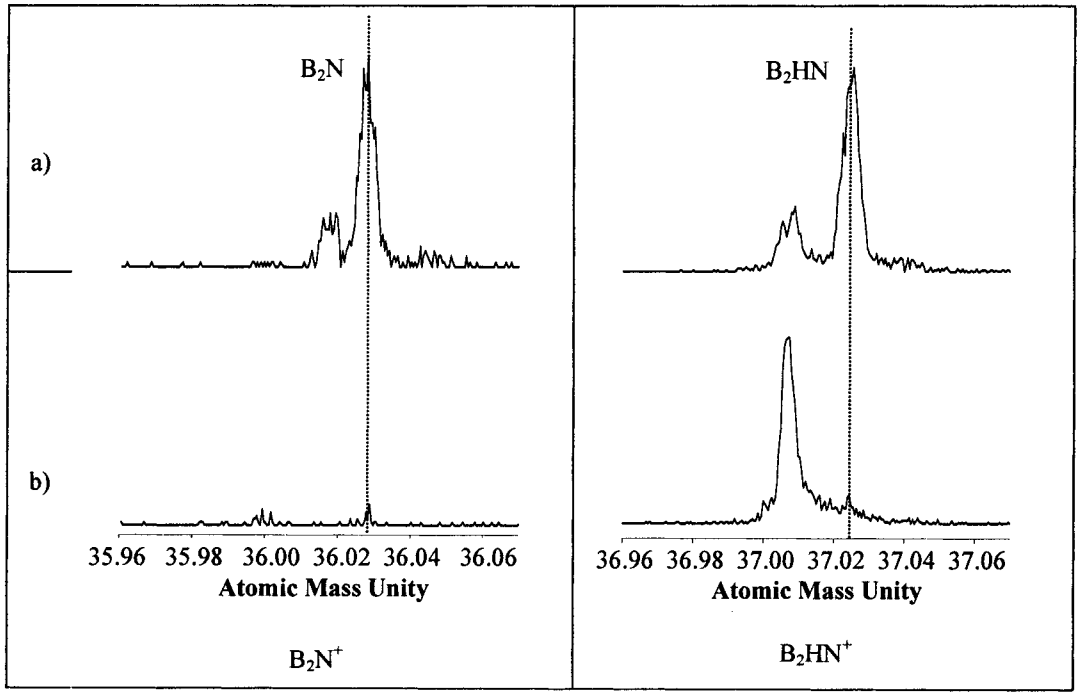


Figure 6 : Partial positive ion TOF-SIMS spectra of pure h-BN :
(a) and additive before friction (b) in the region of the B2N- and B2HN- ions. The peaks at 37.028 and 37.024 amu in the spectra b) are due to the noise and no to a response of the h-BN molecule.

4. CONCLUSION

In this work, we have evidenced the formation of a lamellar solid from the tribochemical reaction of a borated additive and a succinimide additive. The result is the formation of h-BN in the tribofilm. Results can be summarized as followed:

1- The tribofilm is mainly composed of an amorphous borate matrix containing highly-dispersed h-BN nanoparticles in the form of sheets less than 10 nm wide and 5 nm thick.

2- The quantity of triboformed h-BN is very low (a few atomic per cent) and only powerful and sensitive analytical and observation tools are able to identify this lamellar compound.

3- We developed a multi-technique approach for a clear identification of the lamellar solid. XPS analysis showed the presence of B-N binding with a stochiometry close to pure h-BN. XANES analysis evidenced the sp^2 hybridation of boron in these B-N bonds. On wear debris, these bonds are identified by EELS analysis. Finally, HRTEM depicted the crystal structure and the location of h-BN in the borate matrix.

4- An interesting feature has been observed from HRTEM pictures: a preferential orientation of h-BN nanocrystalline sheets parallel to the sliding direction. This orientation has certainly a favourable effect on friction.

5- The originality of this h-BN tribochemical reaction lies in the nature of the intermolecular reaction between the two additives. At the opposite, MoS_2 formation from MoDTC is the result of intramolecular chemical reaction.

ACKNOWLEDGEMENTS

The authors are grateful to M. Liraut (Renault) for financials supports and to Masoud. Kasrai (University of Western Ontarion, Canada) for his help in XANES analysis.

REFERENCES

1. C. Grossiord, K. Varlot, J.M. Martin, Th. Le Mogne, C. Esnouf and K. Inoue, Tribology international, 31, 12, (1998), 737-743.
2. C. Grossiord, J.M. Martin, Th. Le Mogne, Th. Palermo, Surface and coating technology, (1998), 108-109, 352-359.
3. K.J. Wahl, L.E. Seitzmann, R.N. Bolster, I.L. Singer, Surface and coating technology, 73, (1995), 152-159.
4. J.H. Adams, D. Godfrey, Lubrication engineering, 37 (1), (1981), 16-21.
5. M.F. Morizur, O. Teysset, Lubrication Science, 4 – 4, (July 1992), 277-299.
6. Y. Qiao, W. Liu, S. Qi, Q. Xue, B. XU and S. Ma, Wear, 215, (1998), 165-169.
7. Y. Kimura, T. Wakabayashi, K. Okada, T. Wada and H., Nishikawa, Wear, 232, (1999), 199-206.
8. C.D. Wagner, L.E. Davis, M.Z. Zeller, J.A.Taylor, M. Raymond, Surf. Interface Anal., 3, (1981), 211.
9. Handbook of XPS, Perkin Elmer Corporation, (1992).
10. D. Li, M. E. Fleet, Journal of Electron spectroscopy and related phenomena, 79, (1996), 71-73.
11. M. Jaouen, G. H., V. Gonnet, Microscopy Microanalyse Microstructure, 6, (1995), 127-139.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Assessing friction, wear and film formation characteristics in formulated lubricants in severe to moderate boundary lubrication conditions

A Morina*, A Neville*, J H Green**, M Priest**

*Corrosion and Surface Engineering Research Group, Department of Mechanical and Chemical Engineering, Heriot-Watt University, Edinburgh, UK

**Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK

The ZDDP additive has been the subject of much of tribological/tribochemical research over the last few decades. Most studies have investigated the effect of ZDDP in simple ZDDP/base oil systems and in some cases in ZDDP/MoDTC/detergent combinations. From these studies useful information has been gained to help in the understanding of the mechanism of ZDDP performance. However, in real lubricant formulations there can be 10-15 additives and the interactions between those additives should be taken into account.

In this paper the effects of eliminating the friction modifier and friction modifier plus anti-wear additive ZDDP from the additive package of fully formulated lubricants on friction, wear and wear film forming characteristics are examined. Tests have been conducted under lubricated wear conditions at relatively low (20°C and 50°C) and elevated (up to 100°C) bulk oil temperatures using a reciprocating pin-on-plate tribometer. The wear film has been examined by Energy Dispersive X-Ray analysis (EDX) and X-Ray Photoelectron Spectroscopy (XPS).

In this work two fully formulated lubricants containing the same dispersant/detergent system have been used. One lubricant contains no antiwear (ZDDP) additive and no friction modifier (referred to as no AW oil) and the other contains ZDDP but no friction modifier (ZDDP oil). It has been shown that with both oils the friction coefficient increases with increasing temperature. In the fully formulated lubricants with no ZDDP the friction coefficient is lower than with ZDDP when tested in conditions where the starting lambda is much smaller than 1. With higher starting lambda values of around 1 the trend reverses. In all cases, as expected, the wear when ZDDP is present is lower. The nature of the film formed in the presence and absence of ZDDP confirms the competition between detergents and Zn-containing additives in film formation.

1. INTRODUCTION AND BACKGROUND

In the cases where full film lubrication is difficult to maintain because of severe working conditions such as low speed or high Hertzian pressure, the asperities of rubbing surfaces will come in contact with each other. This is known as the boundary lubrication regime. Because there is a sparse lubricant film between the surfaces, the physical properties of the lubricant such as viscosity are not as important as the chemical properties of the lubricant. The chemical additives added to the lubricant help in forming the protective thin film between the rubbing surfaces. These films have low shear strength, so they are removed, partially or fully, during sliding [1]. The effectiveness of the additive package in protecting the rubbing surfaces depends on the rate of film formation and rate of film removal. Both depend on the working conditions. Thus, various additive packages are effective only in certain operational windows [2].

The anti-wear performance of the additive packages is often insured by ZDDP additive. How ZDDP additives form protective films on rubbing surfaces has been a concern for researchers in the field of tribology since the performance of the additive was first reported. The mechanism of wear film formation is important for two reasons – firstly to understand how current additives function and hence determine the limits of their usefulness and

secondly, to enable new additive packages to be developed. It is the second aspect of this that is currently important with regard to current legislation, which will lead to a reduction in the phosphorus in oil additives from 2000 and beyond [3].

The composition of these films has been analysed using a range of surface chemical techniques. Using these techniques very useful information regarding the chemical composition has been found [4-7]. The current understanding relating to the composition and structure of boundary wear films formed on steel surfaces was reviewed by Martin *et al.* [4] and it is appreciated that a layered structure exists, where the upper wear film is long chain zinc polyphosphate polymer-like material and mixed iron and zinc short chain polyphosphate are present in the bulk with a base of iron oxi-sulfides. However, it is still not clear what the exact mechanism of formation of these films is although some reaction steps involving digestion of Fe_2O_3 by phosphate glasses have been proposed [8]. However, Fuller *et al.*[5] state that Fe_2O_3 is not needed and the mechanism is purely thermal. Besides this, relatively little is known about the kinetics of formation as well as the mechanical properties of these films. Some attempts have been made using scanning probe microscopy to analyse the micromechanical properties of the wear films [9, 10, 11] and these studies have concluded that anti-wear films formed from wear tests are extremely heterogeneous in their mechanical properties. The indentation modulus varied from 25 GPa to 81 GPa, while the maximum shear stress varied from 2-3 GPa [10]. The ultimate efficiency of these films depends from the equilibrium of film forming rate and film wear.

For the performance of tribological contacts the important parameters in terms of achieving fuel economy are friction and wear. ZDDP is most commonly used as an antiwear additive plus it often has anti-oxidation properties [1]. The effect of ZDDP on friction performance is often complicated. In the literature, there have been reports where ZDDP formulated lubricants produced higher friction than lubricant formulations without ZDDP [12-16]. Holinski [12] used the Bartel lubrimeter which was designed to test gear oils under boundary conditions. He investigated the effect of boundary layers formed from different additives on friction and wear, and found that when the lubricant is changed from base oil to base oil with ZDDP friction and frictional contact temperature increased. This was also observed by Kennedy *et al.* [13], who used a motorized engine test for evaluation of different lubricant formulations. They also found that ZDDP additives reduce the effectiveness of friction modifiers. They, too, attributed this increase in friction to the anti-wear film formed by ZDDP. Kubo *et al.* [14] investigated the frictional behaviour of ZDDPs with and without an organic molybdenum compound under the boundary lubrication regime. They used Electron Probe Microanalysis (EPMA) to chemically analyse the surface film formed. From surface analyses made in the case when the friction was high and when friction was low, they observed that element intensities were higher when the friction was higher, suggesting that thicker films were formed. Tripaldi *et al.* [15] investigated the friction properties of ZDDPs in different testing conditions, from boundary to elastohydrodynamic lubrication regime. They found that in the EHD lubrication regime ZDDP did not affect the friction response compared with the base oil friction. However, in the boundary and mixed lubrication regime they observed an increased friction response, although in the boundary lubrication regime it was not a significant increase. They proposed three mechanisms by which the ZDDP film may cause higher friction in mixed lubrication regime. The first is formation of a high traction EHD film. They took into account that ZDDP, as well as forming a solid like film, also forms on top of this a more fluid-like phosphate-rich layer which is likely to have a higher traction coefficient than the base oil because of its three-dimensional network structure. The second mechanism is starvation, where it is supposed that because of the formation of the ZDDP boundary film, the EHD film thickness will be reduced and thereby the lambda ratio. The third cause, according to Tripaldi *et al.*, for increasing the friction by oils with ZDDP is the effect of increasing the roughness of the surfaces by the boundary film formed from secondary ZDDPs which in turn reduces the lambda ratio.

Taylor *et al.* [16] found that by increasing the temperature of the lubricant, the thickness of the film formed was increased. They suggested that ZDDPs increase friction by producing an effective surface roughening which promotes boundary lubrication conditions at the expense of EHD lubrication. Jahanmir [17], too, observed an increase

of friction when ZDDP was present in the lubricant. There are also works where a decrease and a neutral effect on friction of ZDDP are reported. Cann and Cameron [18] observed a decrease in friction when a surface film was formed using lubricants with ZDDP. In this work the tests were conducted using a High Frequency Reciprocating Rig facility with frequency of 50 Hz over a 0.5 mm stroke. Sheasby *et al.* [19] used a reciprocating wear test rig to evaluate boundary lubrication of ZDDPs and they did not register any difference on friction when using base oil alone or with ZDDP. In the reciprocating rig used in this work [19] line contact was used, where the upper sample loaded with 50 N load slides against the other sample, which was static, at a frequency of 25 Hz.

An increasing trend of friction as a function of time when ZDDP is used, was observed by So *et al.* [20]. In their tests, which were plane-on-plane contact in very severe testing conditions, the friction for all ZDDP containing lubricants falls in the range 0.08-0.16 depending on temperature. They suggested that response on friction was because of the severity of the testing conditions which did not allowed stable surface film to be formed. These tests were performed at a speed of 1.4 mm/s and load up to 250N

The main aim of the work presented in this paper is to assess the effectiveness of the ZDDP additive in a formulated lubricant over a range of operating conditions defined by the initial lambda ratio. The different approach in the study is that the lubricants tested were fully formulated with both antiwear additive and friction modifier missing and only friction modifier additives missing respectively. In a sense these can be regarded as model lubricants but in this case with selected additives missing. This is to complement the extensive literature described above which focuses on model oils containing certain additives such as ZDDP. The relationship between tribological response and the chemical nature of the wear film from surface analysis was investigated.

2. EXPERIMENTAL DETAILS

The tribotester used for this work was a reciprocating pin-on-plate apparatus with a moving plate and static pin with dead loading arrangement. Experiments were conducted at three temperatures (20°C, 50°C and 100°C) and typically lasted 8 hours. For most tests the load was 188N (corresponding to 0.98 GPa maximum Hertzian pressure) and the crank speed was 60rpm (corresponding to 0.02 m/s mean linear sliding speed) with a 10mm stroke. Based on the results from these preliminary tests a series of experiments was devised to probe the wear and friction characteristics further as a function of lubrication regime and involved variation of the sliding speed, and surface finish of the sample.

The pin-on-plate apparatus was instrumented to measure friction force via a bi-directional load cell. Readings of the frictional force were taken every ten minutes for 2s, which corresponds to two stroke cycles. The friction coefficient was then calculated and plotted as a function of time for the duration of the test. The average friction coefficient was also calculated for each test.

The conditions of the experiments were chosen in order to evaluate the lubricants under the boundary lubrication regime. The calculation of the dimensionless film parameter λ based on the initial composite rms surface roughness was used as an index of the lubrication regime. This parameter was calculated using the Dowson and Hamrock minimum film thickness equation in [1]. Lambda ratios for preliminary tests at 20°C, 50°C and 100°C are given in Table 1.

Table 1. Mean starting lambda ratios for a load of 188N and mean linear sliding speed of 0.02m/s

Temp. (°C)	Lambda ratio, λ
20	0.126
50	0.07
100	0.022

The materials used were bearing steel AISI 52100 (1.3-1.6%Cr, 1%C) as the plate material and Cast Iron BS1452 as the pin material. The microhardness of the plate is 220HV while for the pin the microhardness is 185HV. Pin geometry was cylinder with circled end so the contact was point contact. For preliminary tests the rms roughness (σ_{rms}) was 0.2μm.

The lubricants used contained the same dispersant and detergent packages but with the differences being that the first lubricant had no friction modifier or antiwear additive and the second one had no friction modifier but did contain ZDDP

anti-wear additive with 0.1%P level. These are referred to as 'no AW' oil and 'ZDDP' oil respectively. The viscosities of the lubricants are shown in Table 2.

Table 2. Measured viscosity of the oils

	Viscosity (Ns/m^2)	
	40°C	**90°C**
ZDDP oil	0.057	0.014
No AW oil	0.055	0.013

Before each test the samples were degreased with acetone, to remove any impurities from the surface. After each test the samples were dipped in heptane for 1-2 seconds, to remove excess oil from the surface. Samples were weighed to determine the mass loss from the tribological tests and to calculate the wear coefficient. The tests were replicated three times and a variation on friction results of 3% for tests at low temperatures (20°C and 50°C) and 5% for tests at high temperatures (100°C) was recorded.

Environmental Scanning Electron Microscopy (ESEM) with Energy Dispersive X-Ray Analysis (EDX) attachment and X-ray Photoelectron Spectroscopy (XPS) were used in the post-test analysis for this study. The ESEM is equipped with a LaB_6 gun, and is capable of operating as a conventional high-vacuum SEM, or under low vacuum in ESEM mode. The ESEM is fully equipped with a range of secondary electron (SE) and back-scattered electron (BSE) detectors. Corrections for atomic number, absorption and fluorescence (ZAF) are achieved through a virtual standard calibration routine. X-ray photoelectron spectroscopy data was obtained using the ESCA 300 facility at the CLRC laboratories, Daresbury. A monochromatic Al Kα (1486.6 eV) X-ray source was used. The survey scans were run focusing the X-rays to a line image 0.8mm x 0.2mm on the sample and the pass energy of 150 eV. For the high-resolution C 1s spectra, the pass energy was 300 eV. In order to assess the change of the elements as a function of depth, Ar^+ sputtering at 2 kV for different times was conducted.

3. RESULTS AND DISCUSSION

3.1. LOW STARTING LAMBDA RATIO, λ=0.022-0.126

3.1.1. Friction and Wear

Figure 1 shows the effect of temperature on friction using the oil containing ZDDP and Figure 2 shows the same data for the lubricant with no ZDDP.

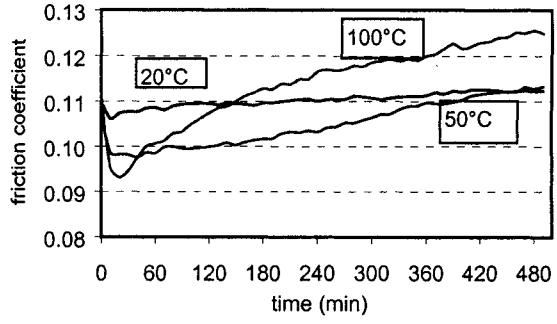


Figure 1. Friction coefficient using ZDDP oil at 20°C, 50°C and 100°C

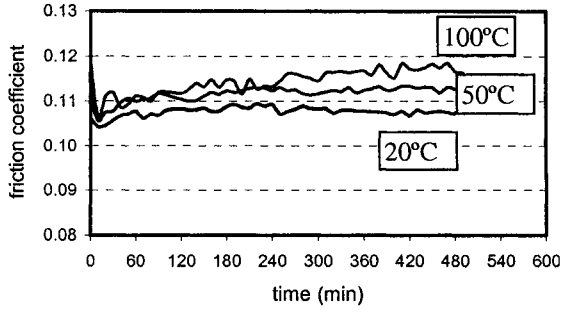


Figure 2. Friction coefficient using no AW oil at 20°C, 50°C and 100°C

It can be seen from these data that for the no AW oil the friction maintains a steady value for the 8 hours test at 20°C with a small drop on friction at the end of the test while at 50°C and 100°C it rises slightly. However, in the presence of ZDDP at 50°C and 100°C the friction coefficient rises after an initial drop in the first 60 minutes of the test, while the initial drop and subsequent rise of friction coefficient at 20°C is much smaller. This drop can also be seen (but to a much lesser extent) in the no AW tests. In Figures 3 and 4 the average friction

coefficient and the final friction coefficient values are plotted against the dimensionless wear coefficients (equation 1) for all the tests reported in Figures 1 and 2.

The dimensionless wear coefficients are calculated using the Archard equation [21]

$$K = \frac{V \cdot H}{W \cdot L} \quad \text{.....(1)}$$

K- Dimensionless wear coefficient
V- Volume loss (m^3)
H- Sample Hardness (Pa)
W- Load (N)
L- Sliding distance (m)

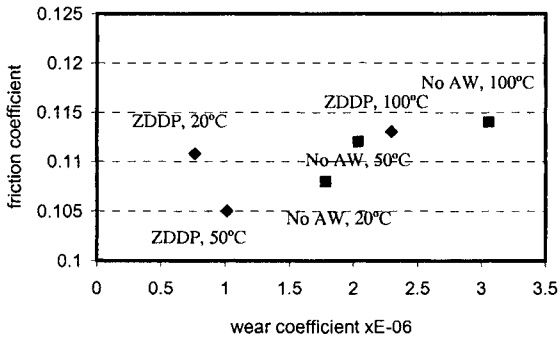


Figure 3. Average friction coefficient (over 8-hours test) versus wear with the no AW and the ZDDP oils at 20, 50 and 100°C

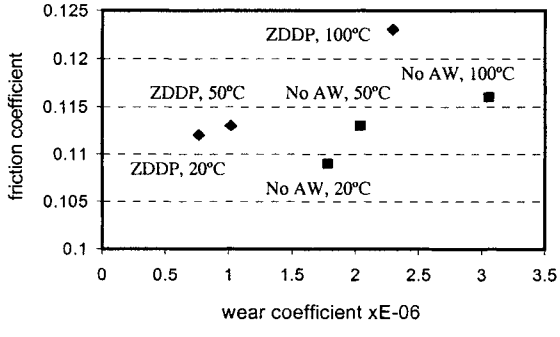


Figure 4. Final friction coefficient plotted against wear for tests at 20°C, 50°C and 100°C in the no AW and ZDDP oils

From Figures 3 and 4 the following observations can be made

- At each temperature the final friction coefficient is greatest when ZDDP is present even though at 50°C and 100°C the average friction is lower when ZDDP is present. Often friction trends are quoted as average values over an entire test. Conclusions from this can be misleading since the average friction coefficient does not give information about the antiwear film evolution throughout the test. The initial drop of friction coefficient in the first 20-30 minutes, when the ZDDP oil was used, is in accordance with observations made by Taylor *et al.* [16].

- At each temperature the wear is significantly reduced when ZDDP is present.

- For each oil, the final friction coefficient shows an increasing trend (almost linear) when plotted against the wear coefficient.

Dimensional wear factors calculated in this work under reciprocating conditions are 1.06×10^{-15}–1.18×10^{-16} m^3/Nm (2×10^{-13}–2.22×10^{-14} m^3/m) for the ZDDP containing lubricant which are similar to those reported from block-on-ring tests in the work of Jahanmir [9].

3.1.2. Post-test surface analysis

In Table 3 the EDX quantification in weight percentage of chemical elements found in the plates wear scars is shown. These results are taken from an area of 500μm x 400 μm. The EDX probes more than 1 μm in depth and so the response is from any wear film as well as the substrate metal. From Table 3 it can be seen that Zn, S and P are evident at each temperature in the wear scar produced with the lubricant containing ZDDP. Also Ca, from the overbasesd detergent package, is present on all the wear scars analysed no matter which lubricant is used. The level of Zn, S and P is higher at 100°C than at 50°C and 20°C. A series of EDX analyses in the middle and at the end of the wear scar were undertaken to assess the uniformity of the wear film. They show generally a higher amount of Zn, S and P at the end of the stroke but the difference is small. One interesting observation on the friction trace during the 8-hour tests with the ZDDP lubricant at 100°C was the decrease of friction force during the first testing hour.

The EDX quantification of the wear scar after 1-hour testing is shown on Table 4. Comparing the amounts of Zn, P and S from the 1-hour test and 8-hour test, it can be seen they are significantly higher after 8-hour testing which means that the antiwear film from the ZDDP additive is developed and most probably thickens as the test progresses. Also it can be seen from EDX analyses of the wear scars after

Table 3. EDX results in wt% from the tests with initial low starting λ (0.022-0.216) after 8-hours

	Temp	Zn	P	S	O	Fe	C	Ca
ZDDP oil	20°C	1.4	0.6	0.9	4.4	73.7	15.3	0.3
	50°C	1.4	0.5	1.0	2.1	53.8	39.1	0.3
	100°C	3.3	1.1	1.4	3.9	74.7	12.6	0.5
No AW oil	20°C	-	-	-	1.5	31.5	65.9	0.3
	50°C	-	-	-	3.0	72.5	20.9	0.4
	100°C	-	-	-	4.2	66.3	27.3	0.2

Table 4. EDX results given in wt. % from 1-hour tests

	Temp	Zn	P	S	O	Fe	C	Ca
ZDDP oil, 1h	100°C	0.8	0.6	0.9	0.8	48.1	45.3	0.2
No AW oil, 1h	100°C	-	-	-	1.6	34.9	62.3	0.1

1-hour tests that the film formed is more S and P rich (higher S/Zn and P/Zn ratios) then the 8-hour films.

Taylor *et al.* [16] suggested that the recorded initial drop in friction in the beginning of the test might indicate the formation of an initial very thin reaction film which has a low shear strength and so it is postulated that the different S/Zn and P/Zn ratios are linked to 'this initial tribofilm with contrasting friction properties. The friction coefficient has the highest value after 8-hour tests (and at 100°C) and at this point the surface has the largest concentration of additive elements.

One of the main surface techniques used for the analysis of thin surface films formed by ZDDP is XPS. It is found that the films formed from ZDDP alone are composed of inorganic polymer materials which contain zinc, phosphorus, oxygen and sulphur [1,4,5,6,7]. Martin *et al.* [4] suggested a two layer structure of these films: on the top surface the long chain polyphosphate (zinc phosphate) and in the bulk the short chain polyphosphate is a mixed Fe/Zn phosphate with a graded concentration, enriched in Fe at the base. This is supported also by the work of Fuller *et al.* [5] using the XANES technique, and mechanisms of conversion of long chain polyphosphates into short chain polyphosphates are given. The sulphur is found in tribofilm in sulphide form [4,5,6,7], and when ZDDP was no longer present in solution, sulphur was present in its oxidized form as sulphate [5]. Zinc was found engaged in both phosphate and sulphide [4].

From work where ZDDP was used in combination with detergents and dispersants [6,7], calcium was found in tribofilms and it has been reported that it can, to some extent, displace zinc.

Figure 5 shows the Zn 2p, P 2p, S 1s, Ca 2p and C 1s peaks on 4 surfaces after an etch period of 15 minutes.

From these traces it is apparent that:

- the highest peaks of Zn, S, P occur at 100°C after 8h. This confirms EDX results and also indicates that temperature and time are important for the growth of the film.
- Very small amounts of Ca are present in the no AW oil and in the ZDDP test after 1h – quantification is difficult due to the low values and the noise in the Ca 2p traces. This is surprising given that Ca is detected by EDX.
- Etching results show that Zn, S and P increase through etching and are therefore enriched as the distance from the outer

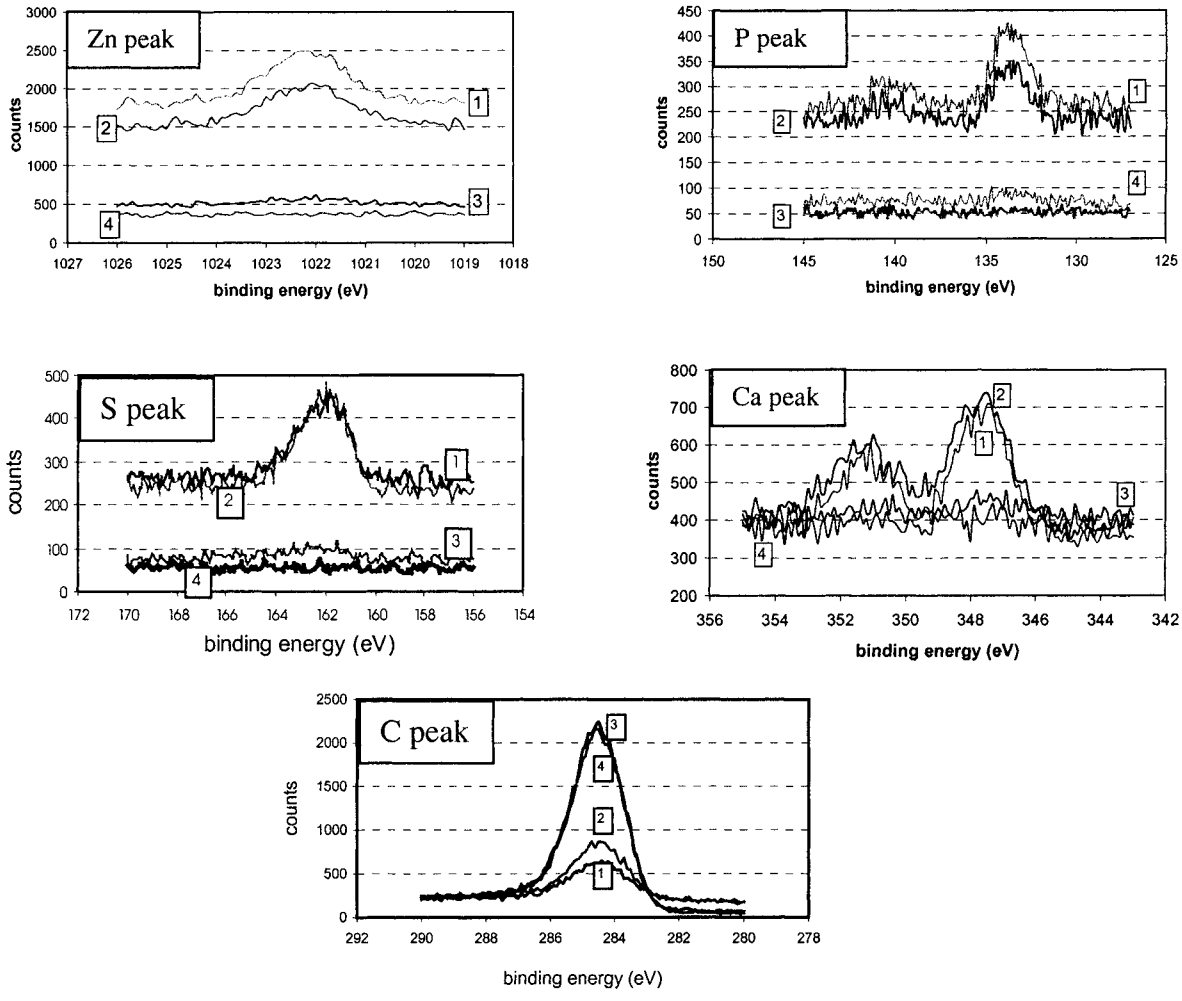


Figure 5. Zn, P, S, Ca and C peaks from tests 1) ZDDP oil, 100°C, 8h, 2) ZDDP oil, 50°C, 8h, 3) ZDDP oil, 100°C, 1h and 4) no AW oil, 100°C, 8h after 15 min of etching.

surface increase. In parallel, the concentration of C decreases.

The C 1s peak is large after short times (1h) with ZDDP and is generally higher on surfaces with no AW oil. The high C 1s peak appears to mask the detection of the Ca and other detergent elements on the no AW surfaces.

Quantification of the atomic % from XPS analyses in this work is shown in Table 5. Where the tribofilms are not fully developed (after 1 hour at 100°C and where no AW additive is present) there is a high concentration of C present at the surface and etching reduces this only slightly. Even after 15 minutes etching no Ca is quantified although it is detected by EDX. The tribofilm, formed in the presence of ZDDP contain less C in the outer layer and it is clear that etching decreases the C and the other elements as components of the wear film become predominant after 15 minutes. The increasing atomic percentages of Zn, P and S as etching proceeds confirms the enrichment of these elements at the base of the tribofilm.

Table 5. Quantification (at%) of the elements found on the wear scars after three etchings

	Etching time	Zn2p	P2p	S2p	O1s	Fe2p	C1s	Ca2p
ZDDP oil, 100°C, 8h	**30 sec**	0.2	2.2	0.9	5.2	0.9	89	1.3
	4 min	0.6	3.6	2.2	13.6	4.6	72.1	2.9
	15 min	1.5	5.7	5.9	29.1	19.1	29.7	5
No AW oil, 100°C, 8h	**30 sec**	0	0	0	0.4	0	99.6	0
	4 min	0	0	0	1.3	0	98.7	0
	15 min	0	0	0	2.4	1.6	95.8	0
ZDDP oil, 50°C, 8h	**30 sec**	0.2	2.9	1.6	6.4	0.6	86.4	1.
	4 min	0.6	4.9	2.8	12.2	3.0	73.6	2.8
	15 min	1.2	7.1	6.4	22.2	14.8	41.6	5.1

3.2. VARYING STARTING λ RATIO

3.2.1. Friction and Wear

One interesting observation from these preliminary results was the rising friction coefficient in the presence of ZDDP at 50°C and 100°C. It was postulated that this was due to the severity of the test conditions and in these boundary lubrication conditions the changing friction represents a running-in period.

A programme of tests was therefore devised to investigate the effect of the wear regime (defined from the starting λ) on friction, wear and the performance of the anti-wear additive. Table 6 shows the variation from the initial test parameters at 100°C (λ=0.022) where the intention was to vary the lambda ratio. The geometry of the contact (by changing the pin radius) was changed, sliding speed increased to 0.06 m/s and roughness decreased to 0.02 μm.

Table 6. Tests parameters to investigate different tribological regimes at 100°C

Radius	Lambda ratio
20 mm	0.24
100 mm	0.74
700 mm	1.25

In Figure 6 the friction traces as a function of time are shown for these revised test parameters. It can be seen that as the starting lambda ratio is increased the linearly increasing friction trend ceases at shorter times (shown on Figure 6) indicating that the steady state is reached at a shorter time. It may have been expected that this time would have been longer as the starting λ is increased if it represented running-in phenomena.

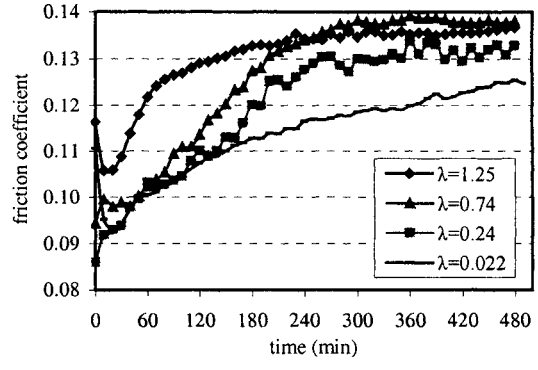


Figure 6. Friction plotted as a function of time for the tests where λ=0.022, λ=0.24, λ=0.74, λ=1.25 (defined in Table 6 and Table 1) at 100°C

It should be mentioned that the quoted λ is the value from the measured σ_{rms} value at the start of the test. As stated by Jahanmir [17] the regime will be determined from the extent of metal/metal contact and the resulting wear mechanism. Whether the surface roughens (either due to increased wear

or ZDDP film) or becomes smoother (due to polishing) is a determining factor for the lubrication regime.

Figures 7a and 7b show the wear coefficients and final friction coefficients plotted against the starting lambda ratio. For ZDDP oil as the λ increases, wear decreases. For the no AW oil there is a more complex relationship. It would appear that generally the difference in wear between the no AW and ZDDP oil becomes greater when the λ is smallest (suggesting that the additive functionality is exploited under these severe conditions).

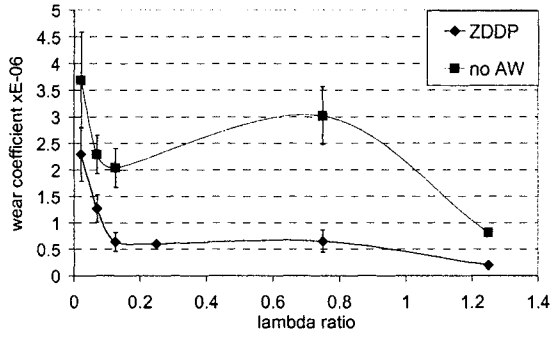


Figure 7a. Wear coefficient versus lambda ratio

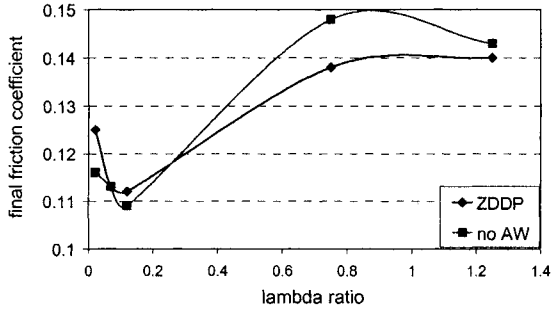


Figure 7b. Friction coefficient in the end of the test versus lambda ratio

In Figures 7a and 7b some similarities with the work of Jahanmir [9] can be seen where he linked friction coefficient and wear to film thickness. In his work a transition from boundary to full film EHD was observed at $\lambda \leq 1$ where as in theory it should be at $\lambda \approx 3$.

However, as in the present work the λ was calculated from the initial surface roughness and consideration should be given to the fact that in developing a different surface roughness and a concave wear scar, the promotion of fluid film lubrication at a lower λ is expected. In the present work it is fully appreciated that asperity contact exists throughout all of the test regimes since wear is evident. However, there may be evidence from these tests that partial EHL is achieved at the higher λ ratios.

3.2.2. Post-test surface analysis, High Starting λ

EDX results from the analysis made on the samples from the second set of tests where the initial testing conditions were less severe (i.e. higher starting lambda ratio) are shown on Table 7. It can be seen that amounts of P and S are much higher than the amounts recorded in tests under more severe testing conditions (lower λ). Importantly the Ca content significantly increased. This suggests that in these conditions the antiwear film is thicker and much more stable and that detergent adsorption and film formation can occur. This can perhaps also explain the more stable friction trace. Much smoother surfaces on the wear scar at higher λ were observed.

Table 7. The EDX results in wt% from the tests with higher lambda

	Zn	P	S	O	Fe	C	Ca
ZDDP oil, λ=1.25	1.9	1.6	2.1	3.5	62.0	26.1	0.9
ZDDP oil, λ=0.73	2.1	1.6	2.1	3.0	78.8	8.0	0.9
No AW oil, λ=1.25	-	-	-	5.3	53.0	38.9	0.5

Although tests where λ was higher are in less severe testing conditions, these tests are still in boundary lubrication regime but conditions are such that a stable film to be formed.

3.3. COMPARISON OF LOW AND HIGH STARTING λ

From the friction and surface analysis results some general characteristics of the films formed can be discussed. The ZDDP film is known to have different layers and for low starting λ it appears that the film, rich in Zn, contains very little contribution from the detergent and is much thinner than the film formed at higher λ. This suggests that the wear regime ensures that the rate of the removal of film layers is higher than the rate of formation and the overall film remains thin. Because of this friction when ZDDP oil was used was higher than when no AW oil was used ($\mu_{ZDDP}>\mu_{no\ AW}$). At higher λ the film incorporates Ca which can be stable in the less severe environment and in general a stable film is formed. As a consequence friction when ZDDP oil was used was lower than friction when no AW oil was used ($\mu_{ZDDP}<\mu_{no\ AW}$).

4. CONCLUSIONS

This paper has demonstrated that the role of ZDDP in affecting friction is complex and depends on the film formation rate and the nature of the film.

- ZDDP film formation increases the friction coefficient and stable friction values are reached more quickly when the conditions are less severe (higher starting λ).
- Fully developed films in ZDDP containing oil have a layered structure and contain much less C at the outer edge than equivalent film formed in lubricants containing no ZDDP where detergent films are activated.

ACKNOWLEDGEMENTS

The authors would like to thank EPSRC and Castrol Technology for financial support of this project.

REFERENCES

1. C H Bovington, "Chemistry and technology of lubricants", Chapter 12, Edited by R M Mortier, S T Orszulik, Second edition, 1997
2. S Korcek, R K Jensen, M D Johnson and J Sorab, "Fuel efficient engine oils, additive interactions, boundary friction and wear", Lubrication at the Frontier/D Dowson et al. Editors, pp 13-24, 1999
3. C H Bovington and R Castle, "Lubricant chemistry including the impact of legislation", Proceedings of the 28th Leeds-Lyon Symposium. Vienna, September 2001
4. J M Martin, C Grossiord, Th Le Mogne, S Bec, A Tonck, "The two-layer structure of Zndtp tribofilms Part I: AES, XPS and XANES analyses", Trib. Int. 34, pp 523-530, 2001
5. M L S Fuller, M Kasrai, G M Bancroft, K Fyfe and K H Tan, "Solution decomposition of zinc dialkyldithiophosphate and its effect on anti-wear and thermal film formation studied by X-ray absorption spectroscopy", Trib. Int. 31, No 10, pp 627-644, 1998
6. P A Willermet, D P Dailey, R O Carter III, P J Schmitz, W Zhu, J C Bell and D Park, "The composition of lubricant-derived surface layers formed in a lubricated cam/tappet contact II. Effects of adding overbased detergent and dispersant to a simple ZDTP solution", Trib. Int. 28, No 3, pp 163-175, 1995
7. G C Smith, J C Bell, "Multi-technique surface analytical studies of automotive anti-wear films", Applied Surf. Science, 144-145, pp 222-227, 1999
8. J M Martin, "Antiwear mechanism of zinc dithophosphate: a chemical hardness approach", Tribology Letters, 6, pp 1-8, 1999
9. A J Pidduck, G C Smith, "Scanning probe microscopy of automotive anti-wear films", Wear, 212, pp 254-264, 1997
10. O L Warren, J F Graham, P R Norton, J E Houston, T A Michalske, "Nanomechanical properties of films

derived from zinc dialkyldithiophosphate", Trib. Lett., 4, pp 189-198, 1998
11. S Bec, A Tonck, J M Georges, R C Coy, J C Bell, G W Roper, "Relationship between mechanical properties and structures of zinc dithiophosphate anti-wear films", Proc. R. Soc. Lond. A, 455, pp 4181-4203, 1999
12. R Holinski, "The influence of boundary layers on friction", Wear, 56, pp 147-154, 1979
13. S Kennedy and L D Moore, "Additive effects on lubricant fuel economy", SAE Paper 872121, 1987
14. K Kubo, M Kibukawa, Y Shimakawa, "Effect on friction of lubricants containing zinc dithiophosphate and organo-molybdenum compound" I Mech E Conference Publictions (Institution of Mechanical Engineers), 1985, pp 121-131
15. G Tripaldi, A Vettor, H A Spikes, "Friction behaviour of ZDDP films in the mixed, boundary/EHD regime" SAE Paper 962036, 1996
16. L Taylor, A Dratva and H A Spikes, "Friction and wear behaviour of zinc dialkyldithiophosphate additive", Trib. Trans. 43, pp 469-479, 2000
17. S Jahanmir, "Wear reduction and surface layer formation by a ZDDP additive", Journal of tribology, 1986
18. P Cann and A Cameron, "Studies of thick boundary lubrication-influence of ZDDP and oxidized hexadecane", Trib. Int., 17, No 4, pp 205-208, 1984
19. S Sheasby, T A Caughlin, A G Blahey and K F Laycock, "A reciprocating wear test for evaluating boundary lubrication", Trib. Int., 23, No 5, pp 301-307, 1990
20. H So, Y C Lin, Gibbs G S Huang, Terny S T Chang, "Antiwear mechanism of zinc dialkyldithiophosphates added to a paraffinic oil in the boundary lubrication condition", Wear, 166, pp 17-26, 1993
21. J. A. Williams, "Engineering Tribology", Oxford University Press, 2000

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Approaches to Sensitising Engine Valve Train Friction Models to Lubricant Formulation Characteristics

J H Green [a], M Priest [a], A Morina [b], A Neville [b]

[a] Institute of Tribology, School of Mechanical Engineering, The University of Leeds, Leeds, UK

[b] Corrosion and Surface Engineering Research Group, Department of Mechanical and Chemical Engineering, Heriot-Watt University, Edinburgh, UK

As a result of increased computing power and software development, the past twenty years has seen a growing trend amongst both academics and industry, to develop models for predicting total engine friction. The potential impact of such work centres around maintaining lubricant development and engine design costs to a minimum. Meanwhile the strongly led legislative demands on lubricant formulators, means that any indication of the performance of a new formulation under specific conditions, is a major advantage prior to commissioning expensive industry standard tests such as the Sequence VIB. This work reviews previous research on valve train friction modelling, with a focus on the approaches taken by the various authors to characterising the lubricant. A typical valve train model is described to provide a framework to evaluate methods of improving sensitivity to lubricant behaviour. Preliminary empirical work is described as a potential mechanism through which greater model refinement may be achieved.

Nomenclature

A	apparent area of contact
A_a	real area of contact
b	Hertzian contact half-width
$\frac{du}{dy}$	velocity gradient across the lubricant film thickness
e	eccentricity of point of contact from camshaft centre
E'	equivalent elastic modulus
F	frictional force
F_a	friction force due to asperity contact
F_s	friction force due to shear of lubricant
G	dimensionless materials parameter, $\alpha E'$
H	average cam/follower frictional power loss
h_{cen}	central lubricant film thickness
h_{min}	minimum lubricant film thickness
I_f	valve lift
k	valve spring stiffness
K	dimensional wear factor
L	cam lobe width
M	equivalent reciprocating mass of the valve system
p	pressure
p_{max}	maximum Hertzian pressure
R	equivalent radius of curvature
r	perpendicular distance from the cam centre of rotation to the frictional force vector
r_B	base circle radius of cam
U	dimensionless speed parameter, $\frac{\eta_o V_C}{E'R}$
V	wear volume
V_c	velocity of the point of contact relative to the cam

V_e	mean entrainment velocity, $\frac{1}{2}(V_C + V_f)$
V_f	velocity of the point of contact relative to the follower
V_s	sliding velocity, $(V_C - V_f)$
W	load
W_a	asperity load
W'	dimensionless load parameter, $\frac{W}{E'RL}$
x,y	cartesian co-ordinate directions
x_d	sliding distance
α	pressure-viscosity coefficient
γ	rate of change of shear stress with pressure
ϕ	cam angular rotation relative to maximum cam lift
ω	camshaft angular velocity
δ	initial compression of valve spring
η	dynamic viscosity
η_o	dynamic viscosity at ambient pressure
τ	shear stress
τ_0	eyring stress (varies with pressure and temperature)
τ_L	limiting shear stress of lubricant
μ	coefficient of friction
μ_a	asperity sliding coefficient of friction
μ_{lim}	limiting coefficient of friction

1. INTRODUCTION

Over the last three decades the importance of energy conservation has become an extensive issue. The developed nations have come to realise that environmental considerations are of importance to the future, whilst also identifying that natural resources of fossil fuels are finite with an estimated reserve of decades rather than centuries. This has therefore become a major variable for the automotive industry to encapsulate when considering the design of a new product. Prime examples during the last ten years include the mandatory exhaust catalytic converters and the development of direct injection gasoline and common-rail, high-pressure, direct injection diesel engines, in the search for increased performance and reduced emissions.

Purposeful steps have been taken to reduce energy consumption through optimisation of several areas, that can contribute to a meaningful overall efficiency gain [1]. Although some aspects rely upon disciplines such as thermodynamics and fluid mechanics for their optimisation, it is also true that realistic energy savings can be made through optimising operation of tribological components.

Newer engine designs commonly make use of DAT (Direct Acting Tappet) camshaft systems and it has been shown that these systems are inherently more able to tolerate lower levels of anti-wear additive elements, such as P and S, than traditional designs [2]. These levels are scheduled to be reduced over coming years due to legislative demands across the world.

As a result of the financial implications of testing lubricants, several mathematical models have been developed [3] that allow an insight into how a particular valve train design and lubricant system may perform under typical operating conditions.

This paper presents an overview of some of the main valve train friction models developed, with a particular emphasis on the workings of one typical model. The potential for refinements to models is evaluated through establishing tribological response data and surface chemistry analysis, in response to the use of different additive packages. This preliminary investigation is carried out through the use of a block on ring tribometer.

2. APPLICATION OF MODELS

Mathematical models have become more widely used by the automotive industry, as both computing power and engine development costs have risen. In order to design a new engine, it is critical that the designer is able to evaluate the impact on the efficiency and durability of the proposed design prior to manufacture. This is now increasingly carried out through the use of lubrication models which consider kinematic and dynamic analyses of the specified system and provide outputs such as lubricant film thickness, friction force and power loss.

However despite these lubrication models being able to analyse the influence of physical factors such as geometry and velocity, much less analysis has taken place with regard to lubricant additive parameters.

3. VALVE TRAIN LUBRICATION REQUIREMENTS

Several engine models have been developed over the last thirty years that aim to evaluate friction throughout the whole engine, including the three main tribological component systems: valve train, crankshaft bearings and piston assembly. As it is widely acknowledged that the frictional power loss of the latter two systems is mainly hydrodynamic, the natural response to optimise efficiency and reduce friction is to reduce lubricant viscosity. However taking this action alone without considering the lubrication of the valve train cam / follower components which typically operate in more severe boundary, mixed or EHL regimes, would lead to their premature wear and failure.

4. VALVE TRAIN MODELS

A number of mathematical models of valve train tribology are evident in the literature, such as those by Coy [4], Taylor [5], Sorab *et al* [6], Paranjpe *et al* and Goenka *et al* [7, 8, 9], Thring [10], Dowson *et al* [11], Dickenson [12], Zhu [13] and Yang *et al* [14]. These models rely on mathematical derivation from first principles based on acknowledged tribology theory. They enable a wide range of input parameters to be stipulated such as crankcase oil temperatures, speed, load and various choices of cam / follower geometry. The output (frictional force or coefficients) of the models are extremely sensitive to the quality of the input data and the number of input variables is large.

The alternative model types that have developed rely upon a much more empirical approach, pursuing a correlation between previous engine test data and experimentally measured lubricant parameters. These models include contributions by Moore [15], Devlin et al [16], Gangopadhyay *et al* [17], Tseregounis *et al* [18], Bovington *et al* [19] and Lacey [20] amongst others. This form of modelling is characteristically much more focused on specific engine types and case study conditions. This is a direct result of using previously collated engine test data that ultimately fixes the engine type, operational parameters and lubricant formulation studied. The fact that the lubricant industry, certainly on a commercial level, often revolves around performance requirements with regard to specific fuel efficiency tests such as the ASTM sequence VIB, means that this approach to modelling can be extremely useful in such focused testing programmes.

4.1. The Functioning of a typical model

The working principles of an existing computational valve train model developed by Dickenson [12] and originally constructed by Ball [21] are summarised in the following section. The kinematics are calculated based on the geometry present in Figure 1 prior to the Hertzian theory of elastic contacts being used.

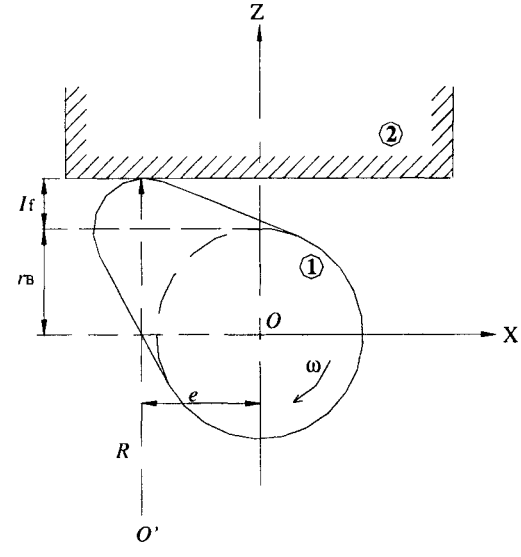


Figure 1: A Direct acting cam / follower with defined geometry

4.1.1. Calculating Film Thickness

As the oil film thickness (OFT) and hence lubrication at the cam / follower interface is variable between EHD, mixed and boundary regimes there are clearly a range of conditions present throughout a complete cycle. To all intents

and purposes however, OFT can be predicted / estimated using EHD lubrication theory [22].

The film thickness at the centre of such a contact can be predicted through the use of the Dowson and Toyoda formula [23]:

$$\frac{h_{cen}}{R} = 3.06U^{0.69}G^{0.56}W'^{-0.10} \quad (1)$$

This is applied to the cam / follower situation, assuming quasi-static conditions.

4.1.2. Friction Force

The substrate elastic deformation of both cam and follower can be considered to be large when compared to the OFT and therefore the contact pressures and dimensions can be approximated to a dry Hertzian contact. By neglecting any rolling friction and approximating the lubricant contact zone between the cam / follower as a region of constant film thickness (h_{cen}), the sliding friction force can be deduced.

The shear stress acting at the solid boundaries (assuming a Newtonian lubricant) is:

$$\tau = \eta\frac{du}{dy} \quad (2)$$

where du / dy is the velocity gradient across the lubricant film thickness.

Inputting the Barus relationship to take the place of dynamic viscosity at high pressures, and using the assumed geometry, the friction force is given by:

$$F = \int_{-b}^{+b} \frac{\eta_0 V_s e^{\alpha p}}{h_{cen}} dx \quad (3)$$

The Hertzian pressure distribution across the contact is given by:

$$p = p_{max}\left[1 - \frac{x^2}{b^2}\right]^{\frac{1}{2}} \quad (4)$$

where x is the distance along the width of the contact.

Finally substituting equation (4) into (3) realises the equation for friction force:

$$F = \frac{\eta_0 V_s e^{\alpha p_{max}}}{h_{cen}} \int_{-b}^{+b} e^{\left(1-x^2/b^2\right)^{\frac{1}{2}}} dx \quad (5)$$

The Barus relationship employs the assumption that the viscosity exponentially varies with pressure. Therefore at high pressures it follows that the viscosity magnitude and hence friction force can potentially become unrealistically large. Hence the model calculates the instantaneous coefficient of friction at each interval within the cycle (360 times per cycle) and compares the calculated value with that of an experimentally measured limiting value. If the limiting value is exceeded, the friction is calculated using the limiting coefficient using the equation:

$$F = \mu_{lim} W \quad (6)$$

where μ_{lim} is an experimentally measured boundary coefficient of friction that is dependant upon lubricant characteristics. The main problem with the functioning of the model is that as a result of the overestimated viscosity, the calculated friction coefficient tends to unrealistic values throughout a vast proportion of the cycle. Figure 2 shows work by Mufti *et al* [24] in evaluating the friction coefficent for experimental work carried out on a single cylinder test engine, plotted against that predicted by a modified version of the theoretical model by Ball [21]. The use of the limiting boundary value (in this case 0.12) between approximately –45 and 45 degrees in the cycle is evident.

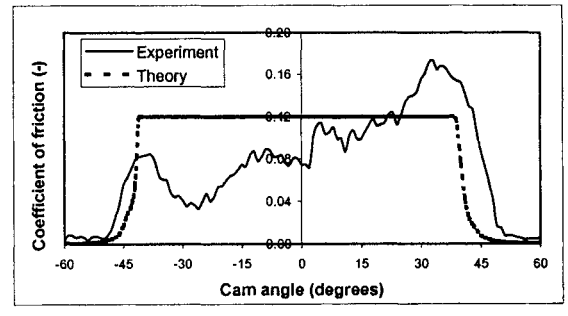


Figure 2: A comparison of experimental friction coefficient and that predicted by the theoretical model [24].

4.2. Alternative Models

A number of alternative approaches to modelling valve train friction have been taken by workers such as Coy [4], Zhu *et al* [13] and Yang *et al* [14].

Zhu *et al* [13] took the approach of refining the modelling of the cam / follower contact to reflect the acknowledged effect of the surface roughness of the components. The lubrication model employs the use of an asperity interaction model by Greenwood and Tripp [25]. Through the use of this contact model it is possible to predict both the asperity contact force and the real area of contact.

Through separating the friction at the cam / follower interface firstly into the shear of the lubricant and secondly the asperity contact, the following analysis can be employed.

For the shear of the lubricant the rheological model by Evans and Johnson [26] is adopted to take account of the the non-Newtonian behaviour of the lubricant that takes place throughout the cycle. Hence the shear stress between the two parallel surfaces is given by:

$$\tau = \begin{cases} \eta \dfrac{V_s}{h_{cen}} & \text{if} \quad \tau \le \tau_0 \\ \tau_0 + \gamma p & \text{if} \quad \tau_0 < \tau \le \tau_L \\ \tau_L & \text{if} \quad \tau > \tau_L \end{cases} \tag{7}$$

where the lubricant viscosity η is assumed to vary exponentially with pressure according to the Barus relationship:

$$\eta = \eta_0 e^{\alpha p} \tag{8}$$

Therefore the integration of the shear stresses over the whole Hertzian region yields the friction due to the shear of the lubricant within the conjunction of the cam and follower:

$$F_s = \int_{-a}^{a}\int_{-b}^{b} \tau dx dy \tag{9}$$

To consider the friction due to the asperity contacts it is assumed that the asperity sliding takes place under boundary lubricating conditions. Hence the friction force due to asperities can be expressed as:

$$F_a = \mu_a W_a \tag{10}$$

Finally, summing the two friction components due to the shear of the lubricant and asperity contact with regard to the asperity contact area gives the friction force F at the contact of the cam and follower as:

$$F = F_a + \left(1 - \frac{A_a}{A}\right) F_s \tag{11}$$

As the shearing of the lubricant is ultimately limited by the inputted value for the limit, it is this value that controls the hydrodynamic friction to a realistic value, similar to the limiting friction coefficient described earlier with the model by Ball [21].

Yang et al [14] also developed a model that functions in a very similar way to that by Zhu et al [13]. It again considers a mixed friction model with friction split into that developed by the process of shearing the lubricant and also that due to asperity contact. The same assumptions were taken as regards the use of the Greenwood and Tripp model [25] and the method used to sum the friction with respect to the asperity contact area is also familiar.

Finally a model produced by Coy [4] has taken a slightly different approach to modelling friction. The model is similar to that proposed by Staron and Willermet [27], and essentially assumes that boundary lubrication is the dominant regime within the cam/follower contact cycle. Through consideration of rheological aspects such as temperature / viscosity and load / speed inputs, the OFT (oil film thickness) is calculated through the use of EHL theory, as proposed by Dowson and Toyoda [23]. The model is able to accommodate the use of multi-grade oils, through the estimation that the viscosity can be derived from the base oil viscosity under high shear rates in the EHL contact. The functioning of the model relies on the comparison between the calculated OFT and the surface roughness of the components as stipulated by the user. Assuming the oil film thickness is less than the defined roughness value, then boundary lubrication is assumed to prevail. Through the

models main assumption that boundary frictional losses far exceed those due to viscous drag (fluid film lubrication), it is deduced that in the event of OFT being greater than surface roughness ($h > h_c$), then the coefficient of friction tends to a minimal value. Figure 3 shows a graphical representation of the transition model. The dashed line illustrates the above situation whilst the solid line gives the modelled response under EHL conditions, when the friction can be higher, hence the friction coefficient is taken as finite for $h > h_c$.

This model effectively has the approach therefore of utilising a boundary value at h=0 (experimentally derived from bench tests) as a limiting maximum. It is then assumed that a linear reduction in friction coefficient throughout the mixed regime exists, as film thickness gradually increases to the point of $h = h_c$ at which point the contribution is assumed to be minimal.

The model is found to possess a reasonable fit when compared to fuel efficiency engine test data, which is inherently undertaken at low crankshaft speeds [27]. However the model is found to be lacking in its ability to deal with higher speed running conditions and cold start situations, where lubricant viscosities and hence viscous losses can be much higher.

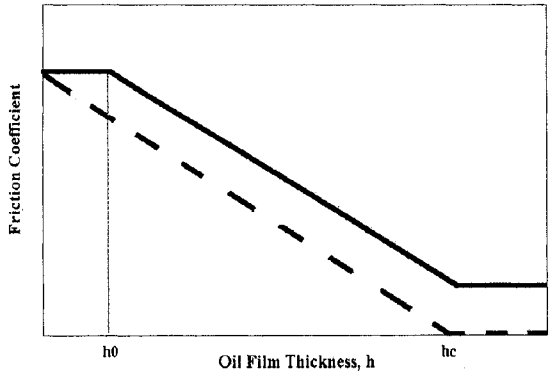


Figure 3: The transition model by Coy [4]

As discussed in section 4 a different approach can be taken to valve train friction modelling which is reliant upon observing a potential correlation between previous engine test data and experimentally measured lubricant parameters. This type of model is not reviewed in this paper due to its lack of adaptability and hence the approach being unsuitable for varying lubricant and design parameters.

The models discussed in this paper clearly attempt to model the lubrication of the valve train as accurately as possible. However it is evident from the methods used that they only manage to take a very limited account of lubricant properties. Certainly no account of the blend characteristics such as the additives used, their type, their interaction with other additives / the surface material and their relevance at different points within the cam / follower cycle are considered. It therefore remains an area in which possible improvements could be made, through the use of empirical data derived from tribometer tests.

The following section aims to evaluate the potential for future model sensitivity improvements, through using lubricant chemistry and tribological performance data from tribometer tests.

5. PRELIMINARY INVESTIGATIONS

5.1. Apparatus

The use of an LFW1 block on ring tribometer was proposed for initial testing of a pair of lubricant blends. The operating principle of the rig is highlighted in Figure 4. An anticlockwise rotating ring is run under pure sliding against a loaded block. Lubricant is supplied from a heated bath via a nozzle and is entrained into the contact before then re-entering the bath. Friction force was measured via a load transducer, placed in line with a connecting arm that attaches to the block holder. K type thermocouples were used for temperature logging of the bath temperature and the supply jet tube temperature whilst "near contact" temperature was also recorded, approximately 10mm from the contact, after conducting through the steel block.

5.2. Materials

The blocks were manufactured from SAE01 tool steel with a hardness of 27 to 33 HRC and a surface roughness of 0.05 – 0.15 μm rms. The rings consisted of SAE 4620 carburised steel having a hardness of 58 to 63 HRC and a surface roughness of 0.65 – 0.75 μm rms.

The lubricant blends used consist of a fully formulated mineral oil and a fully formulated mineral oil minus friction modifier (MoDTC). This

approach enables specific emphasis on analysing the tribological response to both ZDDP and ZDDP + MoDTC (zinc dialkyldithiophosphate and molybdenum dithiocarbamate) combined, under otherwise fully formulated conditions when additive interactions may play a part in anti wear / friction reducing film formation.

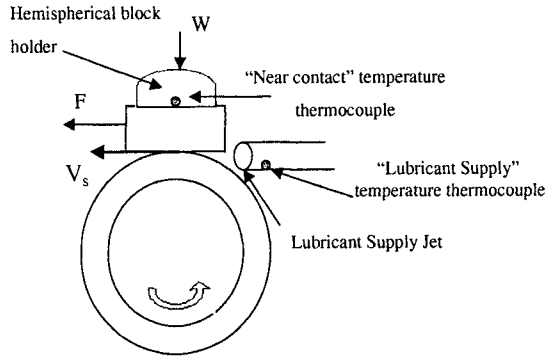


Figure 4: The contact geometry, load application, friction force measurement and two main thermocouple positions on the LFW1 tribometer.

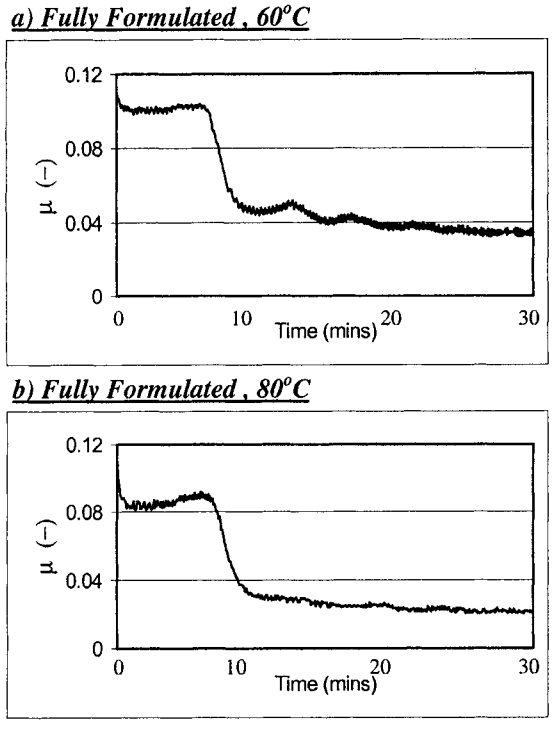
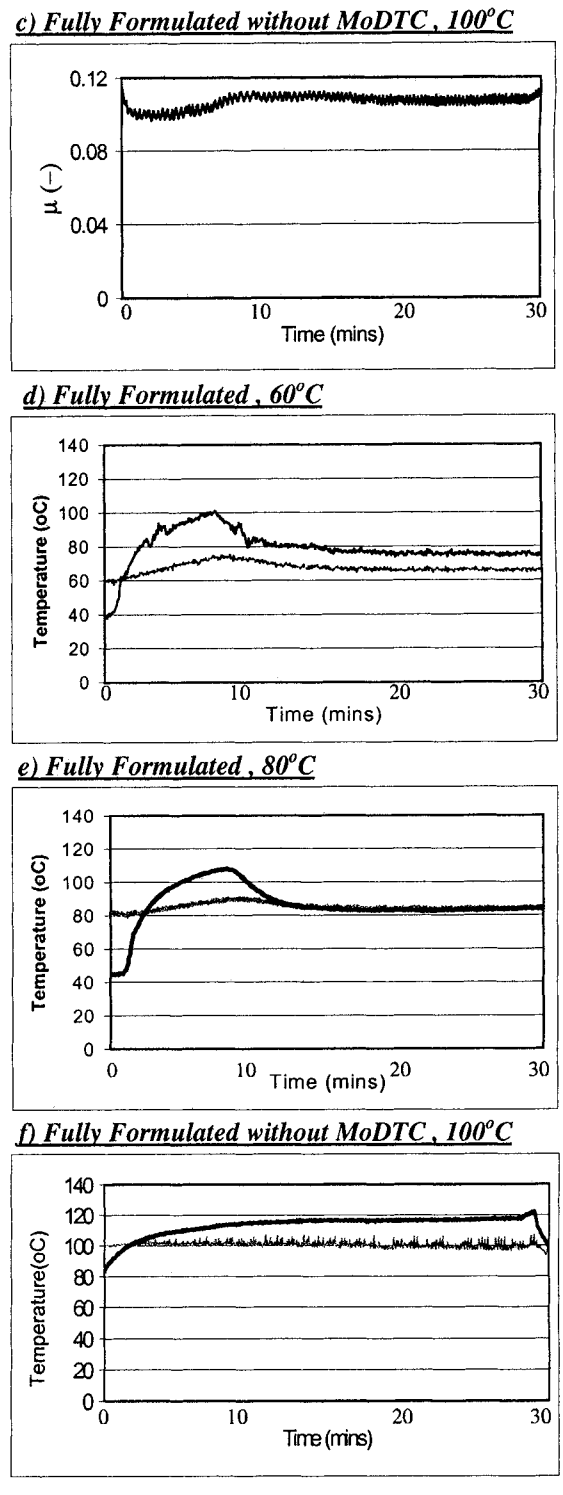


Figure 5: μ and temperature as a function of time.



5.3. Test method

Test conditions used were derived from analysis of previous engine test data, extracting details that enabled a specified point within the cam cycle (-40° from cam nose) of a BMW M40 finger-follower, inlet system at a speed of 400rpm to be replicated at 100 percent sliding. A sliding speed of 0.8m/s, and a maximum Hertzian contact pressure of 516 MPa were used as these were the relevant values at the specified cam angle.

Tests were run for thirty minutes, with data collected at 20 Hz. The data was averaged over each second.

5.4. Results

A selection of the most significant experimental results are displayed in figure 5. Graphs a, b and c provide the friction coefficient data for three tests, whilst graphs d, e and f have the near contact temperature which was measured through a thermocouple 10mm from the point of contact in the block holder. The other measured temperature was the lubricant supply temperature immediately prior to it being entrained into the contact.

Error magnitude in the data is difficult to quantify numerically, due to the high likelihood that temperature instability in the link arm (connecting the force transducer) appears to have the greatest bearing on data accuracy. Repeated tests under the same operating conditions showed this effect to cause fluctuations of up to $\mu = \pm 0.005$.

5.5. Analytical Work

Following the testing, a range of techniques were used to evaluate the wear of the components and the chemical properties of the surfaces.

5.5.1. Chemical Analysis

ESEM (Environmental Scanning Electron Microscopy) with an EDX (Energy Dispersive X-ray) attachment was used to provide surface chemistry data. The ESEM is equipped with a LaB_6 gun and is fully equiped with a range of secondary electron and back scattered electron detectors. Corrections for atomic number, absorbtion and fluorescence are achieved through a virtual standard calibration routine. The EDX results of a selected number of test components are shown in Tables 1 and 2.

The blocks appear to possess the typical characteristics of an antiwear /friction reducing film formation. Table 1 shows that it can be hypothesised, based on previous work, that significant levels of zinc phosphate, iron sulphide and molybdenum disulphide may be present on the block substrate. Future analysis with XPS (X-ray Photoelectron Spectroscopy) equipment will be used to confirm this. Analysis of the rings suggest a slightly different surface composition may be present with an obvious lack of any Phosphorous. This was seen throughout all test samples. The lack of any detectable level of Phosphorous on the rings cannot be accounted for at present.

5.5.2. Profilometry / Interferometry

The test surfaces were examined using both a Talyor Hobson Form Talysurf profilometer and a Veeko Wyko interferometer both pre and post test, for surface roughness evaluation. A typical pattern that was reproduced consistently over most tests was a small reduction in the surface roughness of the ring components. A typical example for a ring component, operating in a fully formulated blend at 60°C was a reduction (measuring across the ring) from R_q=0.68μm pre test to R_q=0.64μm post test. This is most likely due to the "running in" of the component providing a smoother surface compared to that of its ground initial state and also the formation of surface films may have contributed in part to some changes. The blocks however displayed increases in the surface roughness (along the wear scar) pre to post test. An example for a fully formulated blend at 60°C was an increase from R_q=0.10μm pre test to R_q=0.61μm post test. The reasons for this increasing trend are highly likely to be related to the contact geometry of the tribometer components. The relatively high degree of wear taking place within a small area of the component as opposed to distributed around the whole circumference (with the ring) is an influential factor on final roughness measurements. The mass measurements accurate to $\pm 5X10^{-7}$ grams for the blocks and $\pm 5 \times 10^{-6}$ grams for the rings were used to determine the dimensional wear coefficient. This is calculated through using an adaptation of the Archard wear law by Lancaster [28]:

Table 1
EDX data in wt% for the blocks.

BLOCKS	Spot (X mag)	C wt%	O wt%	Mn wt%	Si wt%	P wt%	S+Mo wt%	Ca wt%	Fe wt%	Zn wt%
60°C,Fully Formulated with MoDTC	200	2.4	6.0	1.0	0.5	0.8	2.7	0.4	83.3	2.1
80°C, Fully Formulated with MoDTC	200	3.3	6.5	0.9	0.5	1.0	2.1	0.6	81.8	2.3
100°C, Fully Formulated minus MoDTC	200	2.0	6.3	0.8	0.5	1.0	2.1	0.4	83.5	2.7

Table 2
EDX data in wt % for the rings.

RINGS	Spot (X mag)	C wt%	O wt%	Mn wt%	Si wt%	P wt%	S+Mo wt%	Ca wt%	Fe wt%	Zn wt%
60°C, Fully Formulated with MoDTC	200	0.6	3.7	0.6	0.4	0	0.7	0.1	94.0	0.1
80°C, Fully Formulated with MoDTC	200	13.5	0	0.4	0.4	0	0.8	0.1	84.6	0.2
100°C, Fully Formulated minus MoDTC	200	3.5	4.4	0.5	0.4	0	0.6	0.1	90.4	0.3

$$V=KWx \qquad (12)$$

Where (K) is the dimensional wear factor, (V) the wear volume, (W) the normal load and (x) the sliding distance. The wear volume is calculated from the measured mass through assuming a density of 7.83 g/cc. The mean values observed were 4.65 $X10^{-16}$ $m^3 N^{-1} m^{-1}$ and 3.32 $X10^{-17}$ $m^3 N^{-1} m^{-1}$ for the blocks and rings respectively, using a friction modified blend.

However slightly higher wear factors were observed with the fully formulated but non friction modified lubricant, with means of 1.12 $X10^{-14}$ $m^3 N^{-1} m^{-1}$ for the blocks and 1.56 $X10^{-16}$ $m^3 N^{-1} m^{-1}$ seen for the rings across all the tests conducted. These results appear to support the theory of synergistic actions taking place between ZDDP and MoDTC in oil formulations, to reduce both friction and wear successfully [29].

6. DISCUSSION OF EXPERIMENTS

The small preliminary testing programme has shown that there is relatively good sensitivity to lubricant characteristics. There is consistent differentiation between the fully formulated oil and the fully formulated oil minus the MoDTC friction modifier. The data shows that under the contact conditions used throughout the testing, it was seen that the effect of MoDTC in reducing μ, started to take effect after a time period of around eight to nine minutes (400 metres sliding distance). Higher wear factors and generally higher quantities of adhesive wear appear to be present with the non friction modified blend for both components, hence suggesting that the synergistic action between MoDTC and ZDDP helps to prevent wear as well as maintain low friction levels.

The geometry of the test components suggest that as a test progresses, the wear scar that develops on the block surface modifies the contact conditions, which not only provides a more conformal contact, but also modified Hertzian

contact pressures. This is a common problem with many different variations of tribometer.

7. CONCLUSIONS

- The review of models that have been developed and published over the past ten to twenty years, has highlighted that there are two clear and distinct approaches usually taken towards the prediction of valve train friction.

- Both model types have specific advantages and disadvantages although one factor that has not been considered in adequate detail in either variation, is that of lubricant formulation sensitivity.

- A small preliminary testing matrix has shown that the test data collated is sensitive to the two different blends that have been used, which have been kept as realistic as possible through the use of a fully formulated additive package being present in both.

- The correlation of friction coefficient data captured from the test rig and post test analysis techniques, have provided good support for the differentiation of the characteristics and performances of the two blends.

ACKNOWLEDGEMENTS

The authors wish to thank Castrol and the EPSRC for their financial support and assistance in this research work.

REFERENCES

1. S Korcek , J Sorab, M.D Johnson, R.K Jenson, "Automotive Lubricants for the Next Millenium", Industrial Lubrication and Tribology, Volume 52, Issue 5 (2000), pp 209 - 220
2. J.C. Bell, "Gasoline engine valve train design evolution and the Antiwear requirements of motor oils", Proc Instn Mech Engrs, Vol 212 (1998), Part J, pp 243 - 257
3. R.I. Taylor, R.C. Coy, "Improved Fuel Efficiency by Lubricant Design: a review", Proc Instn Mech Engrs, Vol214, Part J,1999, pp 1-15
4. R.C. Coy, "Practical Applications of Lubrication Models in Engines", Tribology International, Vol 31, No. 10,1998, pp563-571
5. R.I. Taylor, "Engine Friction: The Influence of Lubricant Rheology", Proc Instn Mech Engrs, Vol 211, Part J,1997, pp235-246
6. J. Sorab, S. Korcek, C.B. McCollum, and K.W. Schriewer, "Sequence VIB engine test for evaluation of fuel efficiency of engine oils. Part II:stage selection and time factor determination", SAE paper 982624, 1998
7. R.S. Paranjpe, B.A. Gecim, "Comparative Friction Assessment of Different Valve-Train Types using the FLARE (Friction and Lubrication Analysis) code", SAE paper 920491, 1992
8. P.K. Goenka, R.S. Paranjpe, Y.R. Jeng, "FLARE:An Integrated Software Package for Friction and Lubrication Analysis of Automotive Engines - Part 1: Overview and Applications", SAE paper 920487, 1992
9. R.S. Paranjpe, A. Cusenza, "FLARE: An Integrated Software Package for Friction and Lubrication Analysis of Automotive Engines - Part II: Experimental Validation", SAE paper 920488, 1992
10. R.H. Thring, "Engine Friction Modelling", SAE paper 920482, 1992
11. D. Dowson, C.M. Taylor, L. Yang, "Friction Modelling for Internal Combustion Engines", The Third Body Concept, D. Dowson et al (Ed), Tribology Series, 1996,pp301 - 317
12. A.N. Dickenson, "Engine Friction Modelling Considering Lubricant Tribological Characteristics", Ph.D. Thesis, School of Mechanical Engineering, University of Leeds, UK,2000, pgs. 232
13. G. Zhu, C.M. Taylor, "Tribological Analysis and Design of a Modern Automobile Cam and Follower", Chapters 4,5,6 and 7, Engineering Research Series 7, Professional Engineering Publishing, 2001
14. L. Yang, A. Ito, H. Negishi, "A Valve Train Friction and Lubrication Analysis Model and its Application in a Cam/Tappet Wear Study", SAE paper 962030, 1996

15. A.J. Moore, "Influences of Lubricant Properties on ASTM Sequence VI and Sequence VI-A Fuel Efficiency Performance", SAE paper 961138, 1996
16. M.T. Devlin, W.Y. Lam, T.F. McDonnell, "Critical Oil Physical Properties that Control the Fuel Economy Performance of General Motors Vehicles", SAE paper 982503, 1998
17. A.K. Gangopadhyay, J. Sorab, P.A. Willermet ,K. Schriewer., K. Fyfe. and P.K.S. Lai, "Prediction of ASTM Sequence VI and VIA Fuel Economy Based on Laboratory Bench tests", SAE paper 961140, 1996
18. S.I. Tseregounis, M.L. McMillan, "Engine Oil Effects on Fuel Economy in GM Vehicles - Comparison with the ASTM Sequence VI-A Engine Dynamometer Test", SAE paper 952347, 1995
19. C.H. Bovington, S. Korcek, J. Sorab "The Importance of the Stribeck Curve in the Minimisation of Engine Friction.", Leeds-Lyon Tribology Symposium, Sep 1998
20. P.I. Lacey, "Predicting Sequence VI,VIA and VIB Engine Tests Using Laboratory Methods", Society of Automotive Engineers, 2001-01-1904, 2001
21. A.D. Ball, "A Tribological Study of the Design and Performance of Automotive Cams", Ph.D. Thesis, Dept of Mechanical Engineering, University of Leeds, Uk,1988, pgs 250
22. D. Dowson, "Elastohydrodynamics", Proc Instn Mech Engrs, 1968, 182 (3A), pp151-167
23. D. Dowson, and S. Toyoda, "A Central Film Thickness Formula for Elastohydrodynamic Line Contacts", Elastohydrodynamics and Related Topics, Proc 5th Leeds-Lyon Symposium on Tribology, Mechanical Engineering Publications, London,1979, pp60-65
24. R.A. Mufti, M Priest, "Experimental and Theoretical Study of Instantaneous Engine Valve Train Friction" ASME Journal of Tribology, Volume 125, pp 1 - 10
25. J.A. Greenwood,J.H. Tripp, "The Contact of Two Nominally Flat Rough Surfaces", Proc. Instn Mech Engrs, Vol 185,1970 – 1971,pp625-633
26. C.R. Evans, K.L. Johnson, "The rheological properties of elastohydrodynamic lubricant", Proc Instn Mech Engrs, Part C, 200 (5),1986, pp625 - 633
27. J.T. Staron , P.A. Willermet, "An Analysis of Valve Train friction in terms of Lubricant principles", SAE paper 830165, 1983
28. J.K. Lancaster , "Dry bearings: A survey of Materials and Factors Affecting their performance", Tribology, December, 1973, pp219-251
29. C.H. Bovington, "Chemistry and Technology of Lubricants", Chapter 12, Edited by R.M. Mortier and S.T. Orszulik , Second edition, 1997

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Published by Elsevier B.V.

Electrochemical control of wear: a third body approach

S. Mischler

Ecole Polytechnique Fédérale de Lausanne, IMX-LMCH
1015 Lausanne-EPFL, Switzerland

A phenomenological description for tribocorrosion is developed by including electrochemical corrosion aspects into a third body approach. The developed third body approach stresses the relevance of electrochemical phenomena such as dissolution or formation of stable oxides for the overall wear process. Examples of electrochemical control of wear in engineering and laboratory systems are presented and discussed.

1. INTRODUCTION

Tribological contacts exposed to corrosive environments are found in many industrial situations such as mineral processing, metal forming, orthopaedic implants and marine equipment. Under these conditions the contact materials are subject to both mechanical and chemical solicitations. The rate of material degradation cannot be predicted simply by adding the wear rate in absence of corrosion to the corrosion rate in absence of wear. The reason is that corrosion and wear do not proceed independently and synergistic effects usually result in accelerated degradation (tribocorrosion).

Corrosion processes are of electrochemical nature and therefore electrochemistry likely interferes with the tribological behaviour of tribocorrosion systems. Attempts were made in recent years to control wear by electrochemical methods in engineering systems. For example cathodic protection by sacrificial anodes was used by Kurimoto [1] to reduce wear of steel cutting tools during lathe cutting in cooling water emulsions. The authors used as sacrificial anode a zinc coating deposited on the tool area surrounding the cutting edge. The galvanic effect obtained by coupling steel with a less noble metal alters the electrochemical conditions so that the corrosion potential of steel is reduced while the corrosion potential of the sacrificial anode (and thus its corrosion rate) is increased. In this simple way, tool wear could be significantly reduced. However the effect was pronounced only when commercial coolants of sufficient electrolytic conductivity were used whilst no significant protection was observed with non-conductive liquids. This is not surprising since sufficient ionic conductivity of the electrolyte is generally required to establish an efficient coupling effect of galvanic elements.

Fukuda et al [2] used cathodic protection by imposed current to control wear of tools used for boring of wet wood. For this, they pressed a steel plate on one side of the wood workpiece. Care was taken to electrically insulated the steel plate from the frame of the boring machine. A DC power supply was used to pass a current ranging approximately from 1 to 30 mA from the tool (negative pole) to the plate (positive pole) through the wet wood. Tool evaluations after a constant number of drilled holes showed a reduction in wear up to a factor of 3 depending on the current passed. This was observed for cemented carbide, high-speed steel and alloyed steel tools.

These examples clearly illustrate the potential interest of cathodic protection to control wear in technical tribocorrosion systems. Note that in these systems the corrosion rate was already minimised to negligible values using appropriate

inhibitors and materials. Cathodic protection did therefore interfere with the overall wear behaviour of the systems, including mechanical aspects. This shows the relevant role played by interactions of electrochemical and mechanical parameters in the degradation of engineering systems. Such interactions are little understood at present because of the complexity of the involved phenomena and the large number of critical parameters. However, they offer the opportunity to control wear by electrochemical methods already established for engineering corrosion systems. There is a need to gain a better scientific understanding of tribocorrosion allowing for the development of protection methods and of materials selection criteria, with a special emphasis on the balance between electrochemical and mechanical properties.

Most of the investigations on tribocorrosion were carried out using triboelectrochemical techniques involving the measurement of wear under controlled electrochemical conditions. These techniques offer the possibility to simulate different corrosion conditions by varying the electrochemically-applied potential. In addition, they allow one to record the corrosion reaction rate by simply measuring an electrical current and converting it in mass flow using Faraday's law. The typical configuration of an electrochemical experiment involves an inert material (typically alumina) sliding against the investigated metal held at a selected potential. Although metal on metal configurations are possible [3], the use of an inert counter piece simplifies the interpretation of the electrochemical results. Reviews on triboelectrochemical techniques and experiments are available [4,5].

Models were proposed to describe the extent of wear accelerated corrosion of passive metals subject to abrasion in corrosive environment [4,6]. They are based on the idea that a hard asperity rubbing on a metal produces a track of clean metal that corrodes until passivity is established again (chemical wear). The amount of corroded metal was related to electrochemical (passivation charge), mechanical (normal force, sliding speed) and material (hardness) parameters. Good correlation with experimental observation was found on a series of tribocorrosion systems operating in acid and neutral environments involving stainless steels, a Ti6Al4V titanium alloy and chromium submitted to abrasion against alumina [6,7]. Only partial agreement was however observed with carbon steel operating in neutral borate solutions [8].

In addition to chemical wear, the metal degradation in tribocorrosion proceeds by mechanical wear that arises from particles detached by asperities digging below the surface. The extent of mechanical wear can be determined by subtracting the amount of chemical wear from the volume of the wear track. Mischler et al [9] investigated the role of passive film on the mechanical degradation of carbon steel in tribocorrosion situations implying different neutral to basic electrolytes. These authors observed that the presence and the nature of the passive film critically affected the surface mechanical behaviour of the steel but not significantly the extent of mechanical wear. This apparent discrepancy was explained by considering that mechanical wear corresponds to the metal volume that left the contact as metallic particle and not necessarily to the metal volume that was detached from the metal during rubbing. In fact, a detached metal particle may experience further transformations (oxidation, smearing) before leaving the contact.

The different proposed models describe well-defined tribocorrosion phenomena occurring in specific systems. In light of the variety of tribocorrosion situations encountered in engineering systems there is however a need for a more general theoretical tool to evaluate the relevance in new systems of already modelled phenomena. This need is not specific to tribocorrosion but is common to dry and lubricated wear where many wear laws exists but are specific for well-defined systems. In an attempt to develop a general approach towards wear, Godet et al [10] proposed the third body concept in the early eighties. The concept describes the wear process by taking into account system aspects through mass flow considerations.

According to this concept, the wear process starts with the detachment of particles from the sliding partners. The detached particles remain for a certain time trapped between the two contacting bodies where they form a separating third body that influences contact stress distribution and thus the overall tribological behaviour. The effective material loss associated with wear is given by the definitive ejection of third body particles from the contact. In case of tribocorrosion, the third body approach should be adapted to include the mass flow associated with corrosion processes.

Corrosion is an irreversible surface modification of a material due to chemical reaction with the environment that results in the formation of metal ions dissolved in the liquid (material loss) and, in the case of passive metals, of surface oxide films. A preliminary attempt to include particle flow in tribocorrosion was already proposed by Stemp [11] and Mischler et al [9] to explain the discrepancy mentioned above between first body degradation and mechanical wear. This paper is aimed at developing a phenomenological model of tribocorrosion by combining electrochemical corrosion effects with the third body concept of wear. The approach is applied to three electrochemically controlled wear situations, i.e. wear under cathodic protection (absence of corrosion), wear in presence of passive films and wear combined with metal dissolution. The proposed concepts are compared to already published results concerning carbon steel and stainless steels and their merits are discussed.

2. WEAR IN ABSENCE OF CORROSION

Figure 1 illustrates schematically the third body approach as it may apply to contacts in absence of chemical reactions. The wear process is described in terms of metal flow from the first contacting body (for simplification the second body is assumed to be inert as it usually observed when using alumina) into the third body followed by material ejection out of the contact. Three steps are involved: first particle detachment occurs by one or more of the known wear mechanisms (adhesion, abrasion or fatigue). These particles are re-circulated for a finite time within the third body where they experience rolling, deformation, cracking, fracture or compaction. The third body accommodates part of the velocity difference in the contact and determines the load transfer between the first bodies. Thus, it plays a determining role for the overall wear process. As a last step, the third body particles may get definitively ejected from the contact (wear particles) leading to effective material loss (wear). The three steps are illustrated in Figure 1 as flows Q including: Q_1^{metal} (particle detachment), Q_2^{metal} (particle re-circulation within the third body) and Q_3^{metal} (definitive ejection of particles). Additionally third body particles may reintegrate the first body for example by smearing of particles ($Q_{-1}^{metal\ flow}$).

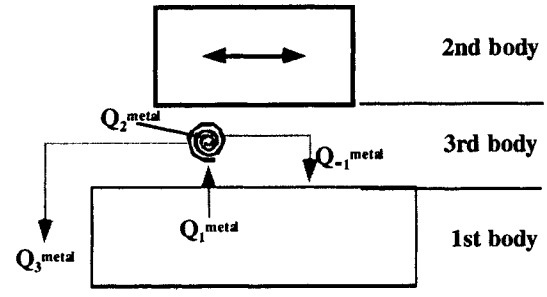


Figure 1. Schematic representation of the material flow occuring in a contact in absence of any chemical reaction. The second body is assumed to be perfectly inert.

Results of recent experiments on the tribocorrosion of carbon steel (DIN 34CrNiMo6, HV 395) in a pH 8.4 borate buffer solution [8] are used here to evaluate the proposed third body approach. The test configuration involved an alumina ball sliding on a horizontal steel plate in a reciprocating motion configuration (1 Hz oscillation frequency, 10 N normal force, and 5 mm sliding amplitude). The vertical position of the ball was continuously monitored using a laser photodetector to measure in real time the wear evolution. The metal plate was connected to a potentiostat as working electrode. A

standard mercury sulphate electrode (MSE) and a platinum coil were used as reference electrode and counter electrode, respectively. A cathodic potential of -1.5 V (MSE) was applied. According to the Pourbaix diagram [12] of iron (based on thermodynamics considerations), at this cathodic potential metallic iron is the stable phase in contact with the solution and no oxidation takes place. Instead the electrochemical reactivity is dominated by water reduction leading to H_2 evolution. The wear process is thus expected to be controlled by mechanical phenomena only and chemical components can be neglected.

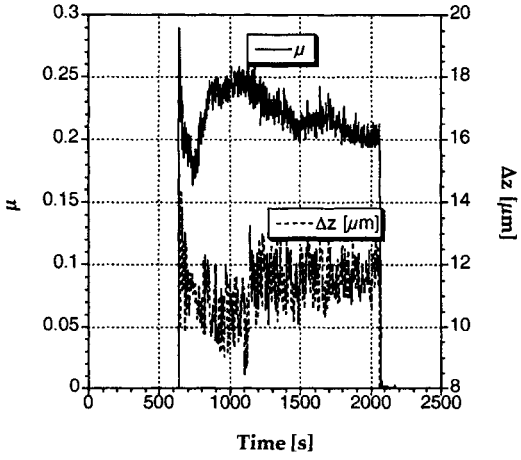


Figure 2. Evolution with time of the coefficient of friction ( ) and of the vertical position of the contact between an alumina ball sliding on a horizontal carbon steel plate held at a cathodic potential in a pH 8.4 borate solution. A decrease of the contact position implies metal indentation by the ball.

The evolution of the contact vertical position (Δz in Figure 2) is used here to characterise the wear process. No degradation or transfer phenomena were observed on the alumina ball after tests. Thus, the vertical displacement of the contact is related to the extent of metal wear only (by negligible third body thickness). Except for a rapid decrease at the beginning of rubbing, the ball position does not vary significantly during rubbing thus indicating negligible wear. This is confirmed by the fact that the wear track was barely visible by optical microscopy. SEM analysis (Figure 3) shows the formation of a thin surface layer of smeared metal in the wear track. Some isolated spallation pits were observed. The cracks formed by Vickers micro-indentation confirmed the layer sensitivity to delamination. The proposed flow model agrees well with the observed wear phenomenon. The fact that no significant material loss (Q_3 metal) exists does not exclude that particle detachment takes place (Q_1 metal) but simply that this flow was compensated by smearing (Q_{-1} metal). Accordingly, the small changes in Δz after about 1200 s (Figure 2) could be interpreted as wear debris entrapped in the contact.

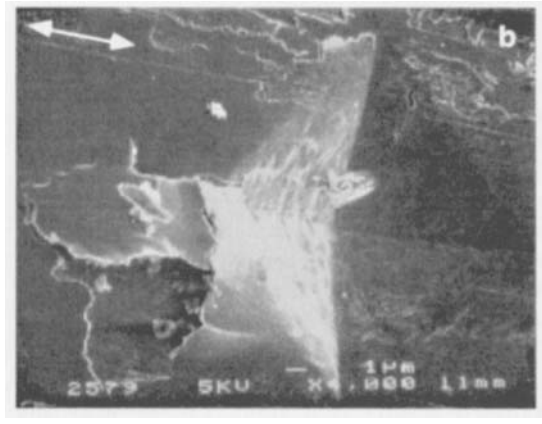


Figure 3. SEM image of the center of the wear track formed on a carbon steel plate held at cathodic potential in a pH8.4 solution by sliding an alumina ball. The pyramidal feature corresponds to a Vickers micro indentation.

3. WEAR OF PASSIVE METALS

Passivity may interfere with the material flow associated with wear in several ways illustrated in Figure 4. Here, first body degradation results not only from the detachment of metal particles but also from the detachment of oxide particles from the passive film (Q_1^{oxide}). The native metal exposed to

the solution by abrasion oxidises again to restore passivity. Oxide particles trapped within the contact experience the same solicitations as the metal third body particles in the contact and may therefore fracture or get compacted. The corresponding internal flow is designated Q_2^{oxide} . Elimination of third body oxide particles occurs either by definitive ejection (Q_3^{oxide}) or by possible incrustation into the metal (Q_{-1}^{oxide}). Third body metal particles are subject to cyclic depassivation/repassivation phenomena in a similar way as the first body. Therefore they generate third body oxide particles (flow $Q_2^{metal > oxide}$). Eventually they can completely transform in oxide before leaving the contact. Note that passivity may modify the mechanical behaviour of a metal as well as the abrasivity of third body particles so that Q_1^{metal} as well as Q_3^{metal} are not necessarily equivalent in a presence or absence of a passive film.

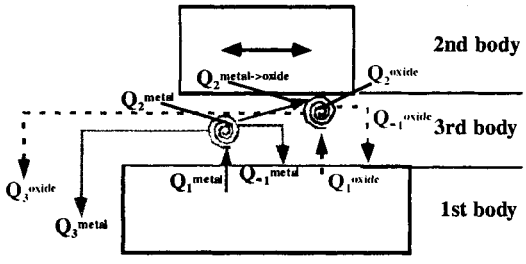


Figure 4. Schematic representation of the material flow occuring in a contact involving oxide film formation on the metal (1st body). The second body is assumed to be perfectly inert.

Experiments carried out under identical conditions as in the previous section but at a passive potential of 0 V MSE [8] are used to illustrate the role of electrochemical oxide formation on wear. At this selected potential, iron oxide is the thermodynamically stable phase and iron is in principle subject to complete oxidation. However, as soon as a passive oxide film forms the oxidation rate decrease because the kinetics becomes limited by the diffusion of reacting species across the oxide film.

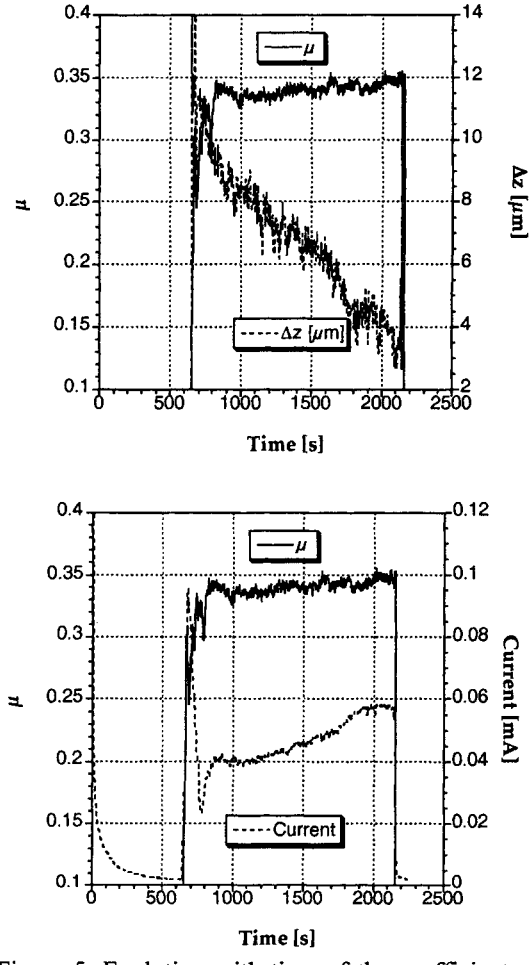


Figure 5. Evolution with time of the coefficient of friction ( ) and of the vertical position of the contact (top) and of the current (bottom) when sliding an alumina ball on a horizontal carbon steel plate held at a passive potential of —0.65 V MSE in a pH 8.4 borate solution. A decrease of the contact position implies metal indentation by the ball.

During anodic oxidation all oxidised metal is incorporated into the passive film and only negligible dissolution occurs [14]. The critical influence of the electrochemical conditions is evident when comparing the wear behaviour at the cathodic (Figure 2) and passive potential (Figure 5). The vertical displacement of the contact shows that the ball steadily indents the metal with a final impingement of 8 m approximately resulting in a large wear scar. Again, no modification of the alumina ball was observed. Using AES surface analysis no incrustation of oxide particles in the metal could be observed. Large amounts of wear particles were found accumulating at the edges of the wear track (Figure 6). High lateral resolution Auger Electron Spectroscopy revealed that these particles have the same composition as the passive film covering the metal [8]. Wear scars formed at the passive potential exhibited in the SEM images (Figure 6) a disrupted surface characterized by plastic deformation and cracks. Isolated grooves due to abrasion were also observed. The onset of rubbing at the positive potential corresponded to an increase in current indicating a wear enhanced anodic oxidation of the metal.

The presence of oxide debris indicates that oxide particles were generated by first body degradation (Q_1^{oxide}) and/or by transformation of metallic third body particles ($Q_2^{metal>oxide}$). In principle the electrochemical current measured during rubbing corresponds to the oxide generation rate $Q_1^{oxide} + Q_2^{metal>oxide}$ and, by negligible thickness of the third body, to the ejection flow Q_3^{oxide}. The anodic volume V_{an} was determined by integrating the current increase over the rubbing time (Figure 5) and converting it to metal volume using Faraday's law (valence of dissolution 2.8 according to the passive film stoichiometry). The obtained value of V_{an} ($1.5\ 10^{-3}$ mm^3) cannot account for the entire the wear track volume of $3.7\ 10^{-3}$ mm^3. The difference is due to the ejection of metal particles from the contact, i.e. to the Q_3^{metal} flow. The fact that at the passive potential wear and thus Q_3^{metal} values are higher than at the cathodic potential can be interpreted in different ways. It is possible that oxide debris present in the third body act as abrasive particles thus increasing particle detachment rate (Q_1^{metal}). On the other hand, the ductility of metals can be affected by the presence of surface films [15]

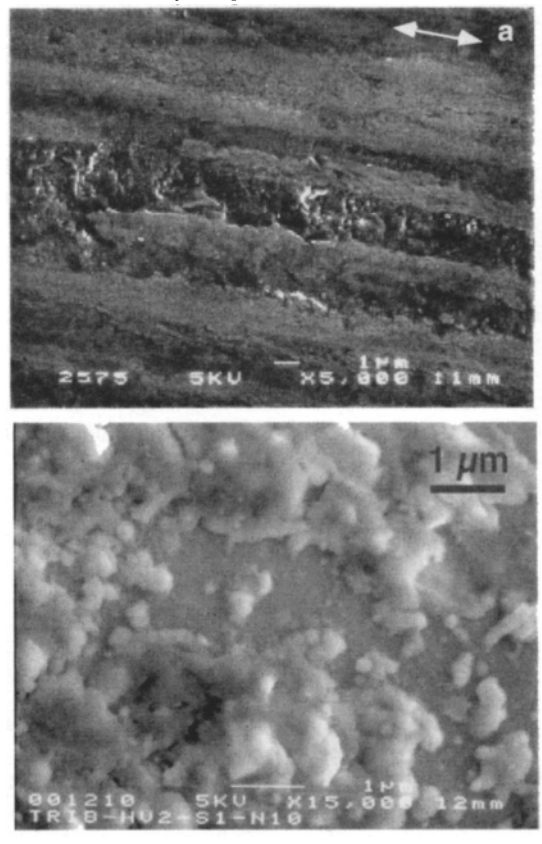


FIGURE 6. SEM images of the center of the wear track (top) and of debris particles (bottom) formed on a carbon steel plate held at a passive potential in a pH8.4 solution by sliding an alumina ball.

rendering the metal more brittle and prone to fracture by repeated passes. This would also contribute to increase the Q_1^{metal} flow and as consequence wear (Q_3^{metal}). On the other hand, the reduced ductility due to the oxide film could hinder the formation (Q_{-1}^{metal}) and the adhesion of the smeared metal film observed at the cathodic potential but not at the passive one. Thus a situation

where the metal particle detachment rate Q_1^{metal} is potential independent but where metal wear (Q_3^{metal}) is critically affected by the prevailing electrochemical conditions is theoretically possible when different smearing capacities subsist.

In Figure 7 the wear track volume V_{wt} values measured in a series of tribocorrosion tests on carbon steel were plotted as a function of the corresponding anodic volume V_{an}. The values were obtained by rubbing an alumina ball against carbon steel (DIN 34CrNiMo6) held at a constant passive potential under constant conditions (similar to the test mentioned above) but in different neutral to basic solutions [9]. In all these environments, steel spontaneously forms a passive oxide layer with only negligible dissolution. The experimental points of Figure 7 are well aligned to a straight line of slope 1 and intercepting the ordinate at a positive value. Thus two independent factors were contributing to material loss and one can write:

$$1) \qquad V_{wt} = V_{mech} + V_{an}$$

Where V_{mech} designates the chemical independent loss factor. According to the schema of Figure 4 V_{mech} results from the Q_3^{metal} flow while V_{an} from the Q_3^{oxide} flow.

Interestingly V_{mech} does not depend on test solution whilst large differences in steel surface degradation were observed depending on electrolyte composition. This can be explained by the third body schema (Figure 4) that assumes a distinction between particle detachment and particle ejection, two phenomena controlled by different factors. Particle detachment depends on the mechanical (shear and normal force, number of passes) and chemical (embrittlement by oxide films, lubricating properties of surfaces) conditions prevailing in the contact. Particle ejection depends in principle on particle size, repulsion forces acting between particles, openness of the contact (geometry, vibrations) and kinematics. A limitation of particle ejection by the contact openness would imply, in case of large Q_1^{metal} values, a longer permanence of metal particles in the contact where they transform continuously in oxide (flow $Q_2^{metal > oxide}$) before ejection. This suggests that particle ejection from the third body determines the wear process when chemically stable oxides form. However, further experimental evidence on the third body behaviour in tribocorrosion contacts is needed to confirm the proposed mechanism.

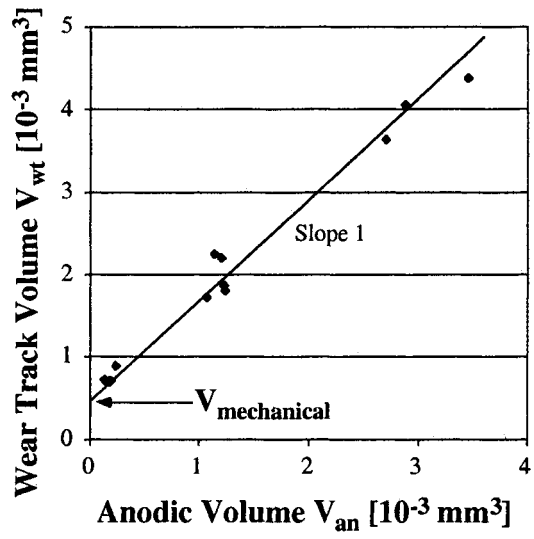


Figure 7. Plot of the wear track volume as a function of the corresponding anodic volumes determined on independent tribocorrison tests of carbon steel under applied passive potential in different solutions (data from reference[9]).

4. WEAR COMBINED WITH METAL DISSOLUTION

A tribocorrosion situation involving the formation of dissolved species differs considerably from the case of passive metals where corrosion products are solid and therefore can contribute to load carrying and velocity accommodation. The material flow in a tribocorrosion system where dissolution takes place is illustrated in Figure 8. Two additional flows are added with respect to Figure 1: Q_1^{ions} and Q_3^{ions} correspond to dissolution of the first body and of the third body, respectively.

The data plotted in Figure 9 were collected from experiments involving rubbing alumina against a ferritic stainless steel (Fe17Cr alloy) immersed in

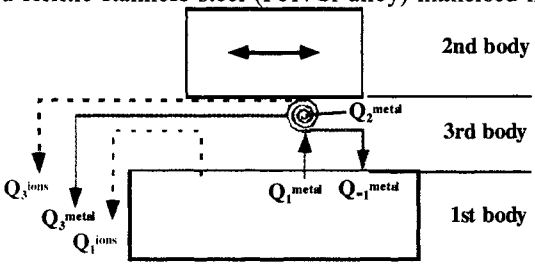


Figure 8. Schematic representation of the material flow occuring in a contact involving metal dissolution of the metal (1st body). The second body is assumed to be perfectly inert.

0.5 M sulphuric acid solutions [11]. The exact valence of dissolution n is not known exactly but is expected to lie between 2 and 3. Conversion of electrochemical charge in metal volume was therefore carried out using these two limit values and the corresponding results are plotted together in Figure 9. Different wear volumes were achieved by varying the applied passive potential, the alumina roughness, the normal force and the sliding velocity. The passivity of Fe-Cr alloys in acid solutions considerably differs from the situation of carbon steel illustrated in the previous section. First, during passivation of Fe-Cr alloys in acid solutions only a relative small fraction of the oxidised metal is incorporated into the oxide film, the rest being dissolved [16]. Secondly, according to the Pourbaix diagram of chromium and iron, no oxide species are thermodynamically stable at the considered potential and solution pH. The stability of the passive film relies on the kinetic equilibrium between oxide dissolution rate and metal oxidation rate. Thus isolated oxide particles are not stable and transform rapidly in dissolved species. The material flow associated with oxide particles was therefore neglected in the present discussion.

A linear relationship between V_{wt} and V_{an} is observed in Figure 9 but relevant differences appear here compared to Figure 7. First the fit line intercept the ordinate at zero and secondly its slope is larger than 1. This implies that a proportionality exists between Q_3^{metal} and the sum factor ($Q_1^{ions} + Q_3^{ions}$). In principle the model does not predict such proportionality because a particle ejection flow is in principle controlled by system properties (openness of the contact, kinematics) rather than by chemical factors. This discrepancy can be explained by considering that the condition for third body build up is $Q_1^{metal} > (Q_3^{metal} + Q_3^{ions})$. It is likely that in the present case fast dissolution of detached particles occurs and thus no significant third body forms. Consequently, ejection of metal particles occurs directly after detachment because there is no constriction by other third body particles. Thus $Q_1^{metal} = Q_3^{metal}$ and $Q_1^{ions} = Q_3^{ions}$ and the system can be considered as a two body system. Landolt [6] and Mischler [4] proposed models for corrosion in

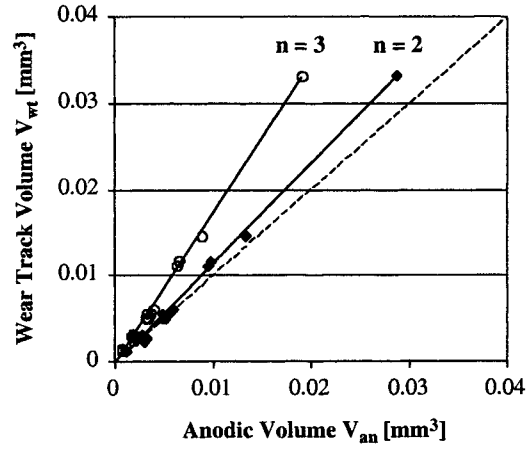


Figure 9. Plot of the wear track volume as a function of the corresponding anodic volumes determined on independent tribocorrison tests of ferritic stainlesss steel in sulphuric acid under different applied passive potentials, normal forces, sliding velocities and alumina counterpiece roughness (data from reference[11]). Two valences of oxidation (n) were considered for iron.

two bodies tribocorrosion systems. They assumed that the current during rubbing is given by the mechanical depassivation rate multiplied by the electrochemical charge necessary to repassivate the exposed bare metal. The depassivation was related to the real area of contact established at asperity contacts, essentially the same relation used in many models for particle detachment such as Archard s adhesive wear model. The observed proportionality between Q^{metal} and Q^{ions} is therefore not surprising because the same mechanical interaction generates these two flows. Interestingly, the two body models for tribocorrosion were recently found to correlate well with the behaviour of Fe-Cr / alumina contacts operating in sulphuric acid [17] while only partial agreement with experimental results was observed in the case of tribocorrosion of carbon steel as illustrated before [8,9]. This suggests that when dissolution dominates the corrosion process, first body degradation determines the overall wear behaviour while the third body effect is negligible.

5. CONCLUSIONS

An attempt was made to integrate electrochemical corrosion phenomena into a third body concept with the aim to develop a more general model for tribocorrosion. The proposed approach describes the wear process in terms of successive phenomena occurring in the contact: formation, ejection and smearing of particles, third body formation, oxidation and dissolution of material.

The comparison with experimental results has shown that the proposed approach allows one to compare and evaluate the relative role of factors such as chemical stability of oxides, particle abrasivity, metal dissolution rate, metal ductility, metal embrittlement by oxide films. In particular the approach suggests that when the metal is subject to dissolution third bodies are not likely to form and therefore the system behaves as a two bodies system. When stable metal oxide form, the system behaviour depends more on the third body behaviour (oxidation, abrasivity) and on the probability for a particle to get ejected from the contact (system dependent vibration and geometry). The phenomenological model systematically illustrates how the prevailing electrochemical conditions may control the wear process by interfering with specific phenomena and flows occurring within the contact.

Further comparison with well-documented experiments is needed to corroborate and elaborate this preliminary approach. In particular quantitative considerations should be introduced. Thus the proposed approach may constitute a promising way to systematically appraise the effect of different mechanical, chemical and system factors determining wear in tribocorrosion systems.

ACKNOWLEDGEMENT

The results presented here are the contribution of A. Spiegel and M. Stemp, previous members of our laboratory. The author thanks Professor D. Landolt and Dr F. Galliano for critically reading the manuscript.

REFERENCES

1) T. Kurimoto, Wear 127 (1988) 241.
2) H. Fukuda, K. Banshoya, Y. Murase, Journal of the Japanese Wood Research Society 38 (1992) 764.
3) M. Stemp, S. Mischler, D. Landolt, Corrosion Science (2002) in press.
4) D. Landolt, M. Stemp, S. Mischler, Electrochimica Acta (2001) 3913.
5) S. Mischler, P. Ponthiaux, Commission Tribocorrosion Cefracor Wear 248 (2001) 211.
6) S. Mischler, S. Debaud, D. Landolt, J. Electrochem. Soc. 145 (1998) 750.
7) P. Jemmely, S. Mischler, D. Landolt, Tribology International 32 (1999), 295-303.
8) S. Mischler, A. Spiegel, D. Landolt, Wear 225-229 (1999) 1078.
9) S. Mischler, A. Spiegel, M. Stemp, D. Landolt, Wear 251 (2001) 1295.
10) M. Godet, Wear 136 (1990) 29.
11) M. Stemp, Thesis EPFL 2292 (2001).
12) Pourbaix, Atlas of Electrochemical Equilibria in Aqueous Solution, M. Dekker Inc, New York (1985).

13) D. Landolt, Corrosion et chimie de surfaces des m taux, PPUR, Lausanne (1997).
14) A.J. Davenport, M. Sansone, J. Electrochem. Soc. 142 (1995) 725.
15) D. Buckley, Surface Effects in Adhesion, Friction, Wear and Lubrication, Elsevier, Amsterdam (1981).
16) D. Hamm, K. Ogle, C.-O. A. Olsson, S. Weber, D. Landolt, Corrosion Science, 44 (2002) 1443.
17) M. Stemp, S. Mischler, D. Landolt, submitted to Wear.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Published by Elsevier B.V.

The Effect of Anti-Wear Additives, Detergents and Friction Modifiers in Boundary Lubrication of Traditional Fe-Base Materials

M.Reyes, A.Neville

Corrosion and Surface Engineering Research Group
School of Engineering and Physical Sciences
Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, Scotland, UK

The action of lubricant additives has been analysed under boundary lubrication on metal couples of Cr-bearing steel plates and cast iron pins in reciprocating mode using model oil systems. Tribological parameters have been measured in addition to analyses of the wear films lubricated with different oils: base oil, oil+ZDDP (Zinc Dialkyldithiophosphate), oil+ZDDP plus addition of Ca-sulphonate detergent and oil+MoDTP (Molybdenum Dithiophosphate).

Isothermal, constant load tests have been carried out using a pin-on-reciprocating plate apparatus under lubricated wear conditions at a bulk oil temperature of 100°C. Tests where load has been progressively increased have also been conducted. The friction and wear analysis of one and eight hours tests were carried out. The chemical nature of the wear film formed by different oils at varying times and with varying loads has been examined by Energy Dispersive X-Ray analysis (EDAX) which penetrates a few μm into the sample. X-ray Photoelectron Spectroscopy (XPS) was used to analyse the top few nm of the surface, focussing mainly on the wear film chemistry. The chemical composition of the different films were analysed at different depths as the specimen was progressively etched with an argon ion beam at 2kV.

It has been shown that the nature of the wear film is directly affected by the nature of the lubricant additives and these additives are affected to a different degree by the load. A direct correlation between the nature of the boundary film and tribological response can be seen and both show the strong interaction between the additives.

1. INTRODUCTION

Friction and wear reduce the operating efficiency, increase power losses, oil consumption and the need for component replacement in engines. All of these, plus the desire to reduce pollutant emissions to the ambient, have promoted research and development of new materials, coatings and lubricants to work as an integrated system to optimise the overall operation.

Oil additives function primarily by their interaction with tribological surfaces. The addition of an appropriate additive to a lubricant increases its mechanical efficiency and reduces friction and wear, thus avoiding surface damage. Nevertheless, the lubricant additives can also interact with each other to provide beneficial or antagonistic effects. These effects need to be assessed and balanced such as in the case of adding detergents and dispersants to reduce deposits and provide anti-corrosion protection with the antagonistic effects found by Smith *et al.* [1] where the addition of detergents and dispersants create a more patchy tribofilm. Also, findings by Willermet *et al.* [2] suggest that detergent/dispersant packages partially replace Zn with the detergent cationic species and interfere with the formation of metaphosphates. The result of this is that addition of detergent can disrupt the anti-wear performance of the DDP additives.

This preliminary research focuses on the effects of adding Ca detergent to oil+ZDDP and comparing oil+ZDDP and oil+MoDTP as sources of wear protection. This work includes a compositional analysis of the tribo-film, complemented with wear and friction analysis on ferrous materials. This, and previous literature, will form the basis for further tests on advanced materials and coatings.

Table 1
EDX material chemical composition Wt%

	Fe	C	Si	Mn	Cr	O	Mo
BS EN 31	96.45	0.94	0.55	0.49	1.57	-	-
BS EN 1452	84.88	9.32	2.74	0.71	-	1.91	0.44

2. MATERIALS AND EXPERIMENTAL METHODS

2.1. Materials

The material couple used for all tests in this work was a Cr bearing steel (BS EN 31) of hardness 180HV (plates) and cast iron (BS EN 1452) of hardness 150HV (pins). Their chemical compositions from EDX measurements are shown in Table 1.

The plate dimensions are 17mm x 6mm x 3mm of roughness R_q=0.23μm. The pins were 20mm long and 6mm diameter with a hemispherical end of 20mm radius and a similar roughness finish of R_q=0.23μm.

2.2. Lubricants

Four different lubricant compositions were used for the tests. Mineral oil free from additives (base oil), mineral oil containing 1.2% ZDDP (ZDDP), mineral oil containing 1.2 mass% ZDDP and 1.2 mass% Ca-sulphonate detergent (ZDDP+Ca) and mineral oil containing 1.5 mass% MoDTP (MoDTP). The dynamic viscosity was measured for the oil additives giving the values as shown in Table 2.

Table 2.
Viscosities in Pa·s measured for tested oils at 20°C and 100°C

	20°C	100°C
ZDDP	0.097	0.006
ZDDP+Ca	0.090	0.006
MoDTP	0.097	0.006
Base oil	0.092	0.006

2.3. Experimental

Wear tests were carried out using a pin-on-reciprocating-plate apparatus instrument in which the load, speed and oil temperature can be controlled. In this study, most of the tests were carried out at 100°C with a constant rotational speed of 60 rpm creating a 10mm stroke. Loads of 12.5, 15, 17.5 and 20 kg were applied at the pin through a lever arm arrangement. The apparatus measures the friction force via a bi-directional load cell that is connected to a computer programme that controls the time elapsed and number of points in each data set. One and eight hours tests were conducted taking friction force measurements every 10 or 15 minutes for 2 seconds (120 points) which corresponds to two full cycles. An average of the 120 points gives the friction force, which is used to calculate the friction coefficient

Wear profiles analysis and measurements were conducted using a Talysurf profilometer taking five measurements of the cross section and along the wear scar.

New pins and plates were used for each test which were cleaned with acetone to degrease and remove all impurities. After the test, the specimens were cleaned with heptane for 2 seconds to remove excess oil from the surface. The specimens were weighed with an accuracy of ±0.1mg after test and heptane cleaning to calculate the material volume loss for each test.

Samples were analysed using the Environmental Scanning Electron Microscope, ESEM (Philips XL30 at Heriot-Watt University). The elemental composition of the film in the wear scar was analysed by an EDX attachment on the ESEM. The EDX penetrates a depth of approximate 1μm and so there is some signal from the steel substrate. X-ray Photoelectron Spectroscopy (XPS) is a surface sensitive technique and is used to probe the top 10-20nm of the surface. In this work, an Al Kα source was used in the SCIENTA ESCA 300 XPS at Daresbury Laboratories. The XPS is equipped with an Argon ion gun to clean the samples and etch the surface to analyse the tribofilm at different depths.

3. RESULTS AND DISCUSSION

For the tribocouple presented in this work and under the conditions described in the previous paragraph the lubrication regime is boundary lubrication as confirmed by the lambda ratio values calculated using the Dowson and Hamrock equation [3]. Table 3 presents lambda ratios for a point contact and Hertzian pressures at the different loads.

Table 3
Material couple parameters under test conditions at 100°C and 60 rpm crank speed.

Load (kg)	Load (N)	Hertzian Pressure (G Pa)	Lambda ratio, Λ
12.5	122.5	0.89	0.0165
15	147.5	0.93	0.0162
17.5	172.1	0.96	0.0160
20	196.2	1.03	0.0159

3.1. Friction response in 1 hour tests

A comparison of the average friction coefficient (μ) collected over a one-hour period is shown in Figure 1. As expected, the base oil shows significantly higher values of friction coefficient at every load. This was followed by ZDDP at higher loads, with friction remaining quite stable with increasing load. ZDDP+Ca shows the largest effect on friction as load was increased but interestingly at high loads show very similar low values to MoDTP.

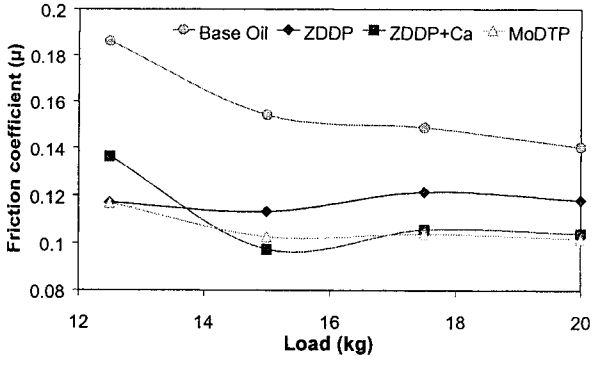


Figure 1. Average friction coefficients for 1 hour test at increasing loads (average of last 40 minutes).

Friction coefficients and ranking of the oil additives differ from long running tests due to the larger instabilities in the first 30 minutes and the use of average values as shown in the next section.

3.2. Friction response in 8 hours tests

Tests carried out for 8 hours showed interesting trends in friction coefficients. With the exception of base oil, the friction stabilised after the first hour, giving essentially constant values. For each of the experiments at least two replicate tests were carried out and showed a good reproducibility on the friction coefficient values with a variation of 0.004.

The base oil shows the highest and least stable friction coefficient. Average values of the friction considering the last 7 hours are 0.15 for the base oil, 0.12 for ZDDP+Ca, 0.11 for ZDDP and 0.10 for MoDTP. These friction results agree with results presented by Unnikrishnan *et al.* [4] on four-ball friction apparatus were MoDTP friction coefficients where are lower than ZDDP.

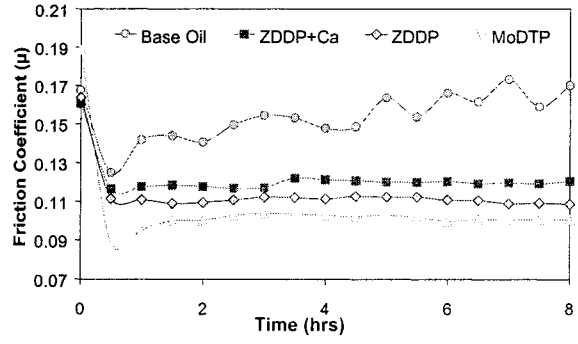


Figure 2. Friction coefficients on BS EN 31 tests for 8 hour test, 20kg load, 100°C and 60 rpm.

To further investigate the effect of load on friction in the four oils, 8-hour tests were run increasing the load progressively every 2 hours from 5 to 10 to 15 to 20kg. Increasing load results show the same ranking as results using constant loads. The base oil friction response is more stable at higher loads and gives friction coefficient values of 0.18, 0.19, 0.18 and 0.16 respectively. The friction results for the oils with additives presented in Figure 3 show a slight increment in friction with ZDDP+Ca as load increases. Interestingly the friction is higher at 5kg and reduces from 0.12 to 0.10 as load in increased for the ZDDP. Similarly, MoDTP shows a reduction of the friction coefficients as load is increased but interestingly a

peak develops immediately after the load is increased. It gives average values of 0.11-0.10 for the first 4 hours and 0.09 for the last 4 hours.

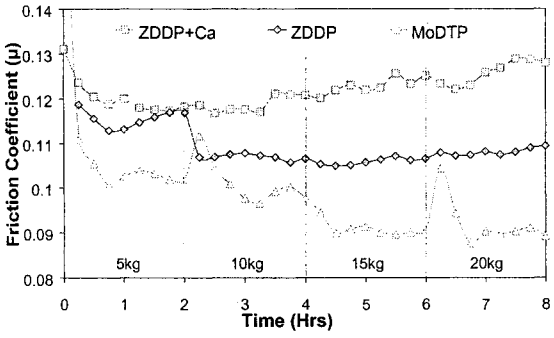


Figure 3. Friction coefficients on BS EN 31 as load is increased from 5 to 20 kg at 100°C and 60 rpm.

The friction increments developed just after the load is increased on MoDTP can be related to findings by Yamamoto *et al.* [5] that suggests that excessive reaction between the additives and the iron surface gives rise to an increase in friction coefficients. Yet, it would appear that this is transitional, most probably ceases when the nascent Fe from the base metal has reacted with the Sulphur compounds to develop the tribo film.

3.3. Wear results

Weight losses in 1 hour tests were between 0.0mg and 0.3mg for all tests so weight measurements were not used for volume calculations due to the lack of sensitivity in the measurements. In some tests carried out for 8 hours, the samples were superficially cleaned with heptane as normal and then cleaned thoroughly in an ultrasonic acetone bath for 20 minutes. This produced weight losses of 0.7mg for the oil+ZDDP and 1.6mg for the base oil.

The wear profiles showed that the severity of the wear for 1 hour tests at 12.5kg follows the same trend for 8 hours, 20kg tests as shown in Figure 4. For these tests the depth of the wear scar on the base oil > oil+ZDDP+Ca > oil+ZDDP > oil+MoDTP. The wear profile gives an indication of depth and roughness of the wear but cannot be used for volume calculations. The profiles show the narrowing and reduction in depth of the wear scar as additives are added. Also, the profiles show some peaks of material, some of which exceed the height of the mean surface. This is an indication of material being displaced or transferred from the pin or an indication of a tribo film being formed, so more sophisticated analyses are needed. A profile was obtained following the wear tracks, and similar results and ranking were obtained as depths are clearly shown in Figure 5.

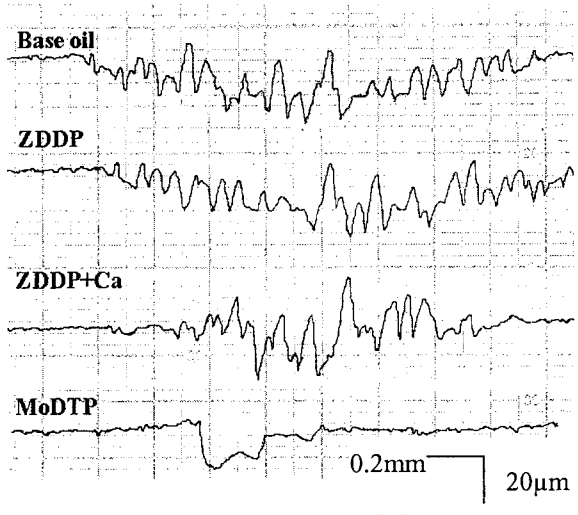


Figure 4. Cross section profiles of the wear scars at 100°C, 20kg for 8 hours

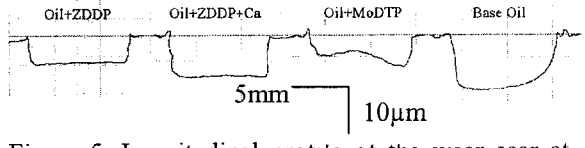


Figure 5. Longitudinal profile of the wear scar at 100°C, 20kg for 8

Using the analysis from a white light interferometer (Zygo New View 5000-3D), it was possible to obtain volume losses and calculate the dimensionless wear coefficient values of K using the Archard equation 1 [3] where V is volume loss, H hardness, W load and L sliding distance.

$$K = \frac{V \cdot H}{W \cdot L} \qquad (1)$$

Volume losses and wear coefficient results are shown in Figure 4 for tests carried out at 20kg for 8 hours, which are similar to those reported by other authors [6, 7]. The values are considered severe for base oil and ZDDP+Ca and the values for ZDDP and MoDTP are considered to be in the mild wear regime.

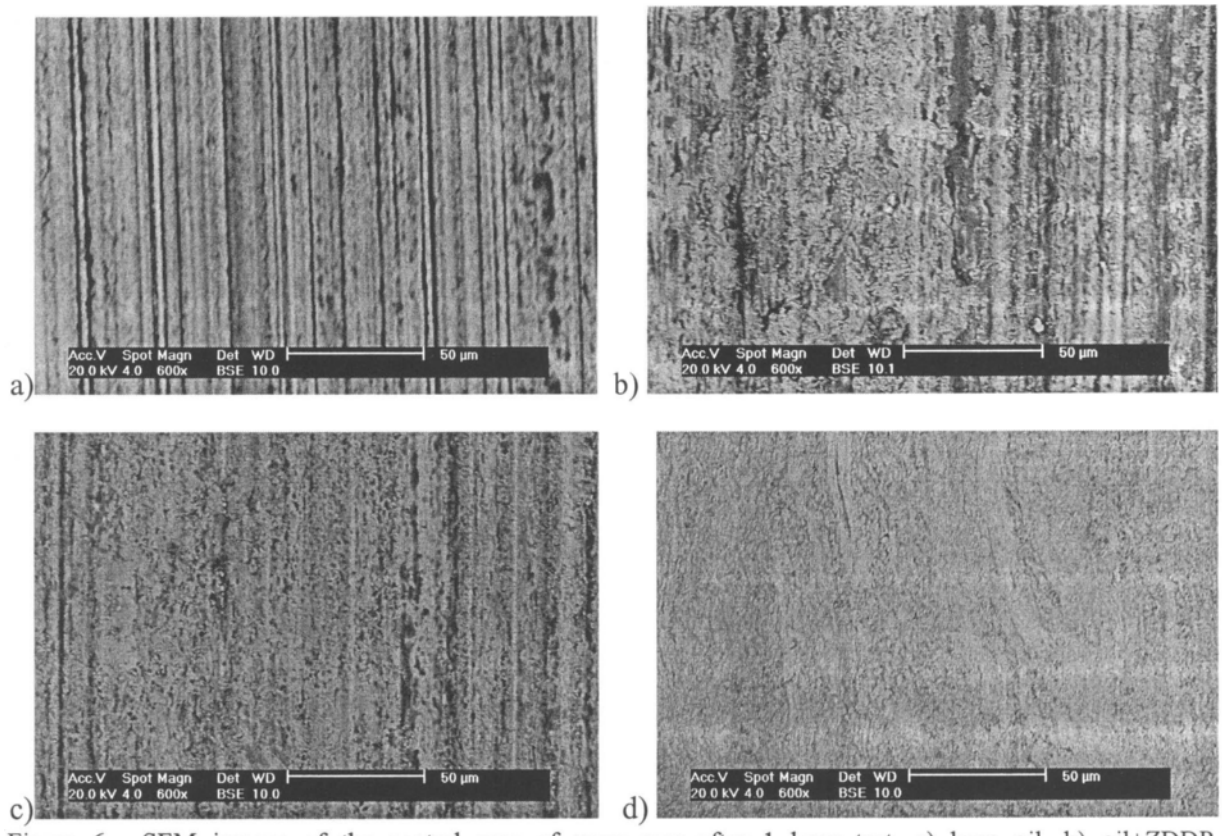


Figure 6. SEM images of the central area of wear scar after 1 hour test. a) base oil, b) oil+ZDDP, c) oil+ZDDP+Ca, d) oil+MoDTP

Table 4
Volume losses and wear coefficients

	Volume loss mm³	K
oil+ZDDP	0.0187	3.0×10^{-7}
oil+ZDDP+Ca	0.0432	6.93×10^{-7}
oil+MoDTP	0.0058	9.31×10^{-8}
base oil	0.1582	2.54×10^{-6}

To get a qualitative idea of the wear present using each of the oils SEM analysis can be used. The following images show the morphology of the wear scars. Using base oil with no additives (Figure 6 a) the deeply scratched surface is visible, implying occurrence of adhesive wear which was diminished by the addition of the anti-wear additive ZDDP (Figure 6 b). There is little difference on addition of Ca detergent (Figure 6 c).

Using oil containing friction modifier MoDTP it is clear that, apart from reducing friction, it performs as a good wear-reducing additive. The wear scar looks smooth and no traces of the surface being ploughed by the pin trajectory are visible as shown in Figure 6 d. Smoothing of rubbed surfaces has also been reported by Graham *et al.* [9] on soluble molybdenum additives, which may produce a reduction of friction by promoting hydrodynamic or elastohydrodynamic film at the expense of boundary lubrication.

Plotting average friction coefficients versus the wear in volume loss shows a linear trend where higher μ generally leads to higher volume loss for the oils tested as showed in Figure 7. These results are similar to finding by Yamamoto *et al.* [5] where lowest friction gives lowest wear on molybdenum additives containing P-O and P-S compounds. However, this is not always true and some ZDDP additives can reduce wear but increase friction.

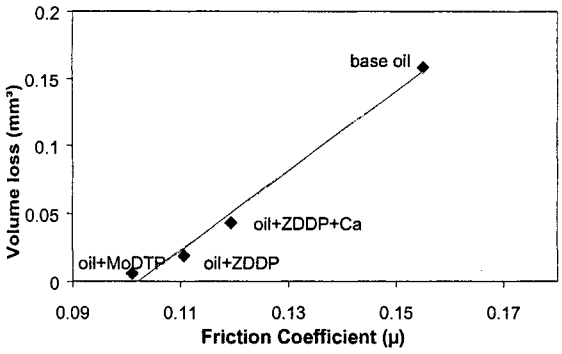


Figure 7. Volume loss versus average friction coefficients.

4. EDX/XPS ANALYSIS

4.1. Oil-additive behaviour as a function of load and time

Analysis of the surface elements through EDX was carried out primarily to detect the film and the main elements on the surface prior to XPS analysis.

Figure 8 shows key elements found on the films as the load is increased from 12kg to 20kg and time increases from 1 hour to 8 hours for the three oil additives. More Zn, S and P are detected on the film on the ZDDP as the load is increased for 1 hour (but high C content was detected at 12.5kg film). Load increments increase contact temperature so ZDDP decomposes more rapidly, but findings by Yin *et al.* [10] do not show increases in S concentration in the film as load is increased, but mention that the rate of film formation is affected by the load. This differs from findings by other authors [7] were the S species increased in the film with increasing load. In this particular case less amounts of Zn, P and S were found as time increased to 8 hours. Yin *et al.* [10] also mention that running time affects the chemical composition of the film but they did not mention specific changes in concentration.

On the film developed by the ZDDP+Ca, the Zn remains quite constant as load and time increased. P and Ca remain constant as load is increased, but are reduced with time and S is the most sensitive to load and time in the presence of Ca detergent.

In the MoDTP blend, the P content is slightly reduced with increasing load but increased with test time. Interestingly, the S+Mo content increases with load as findings by Yamamoto *et al.* [5] showing that in engine oils an increase in load increases P, S and O. Also in engine oils containing MoDTP the Mo content was considerably increased with load. Under this particular conditions S/Mo content also increases with time.

For S and Mo it is difficult to separate the peaks using EDX and so the Wt% on MoDTP is a combined S+Mo. Identification and values of these elements is clarified by the XPS analysis.

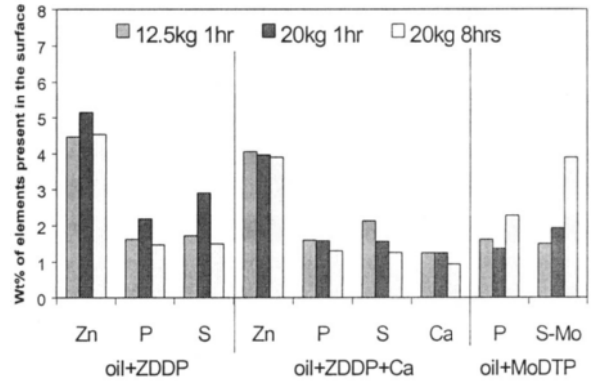


Figure 8. EDX surface analysis

The amount of Fe detected from the substrate plate during EDX surface analysis can be used as a qualitative indication of the film thickness. In Figure 9 as expected, the base oil shows more Fe than the rest of the additives indicating a thinner film, not showing changes as load or time is increased. Results consistently show less Fe on the ZDDP than on the ZDDP+Ca at all tested conditions, indicative of a thicker film.

The MoDTP show contrasting results, whereby in 1 hour tests more Fe is present than for the rest of the oils, indicating a thinner film. However, results at 8 hours, 20kg load show the lowest Fe content which is an indication that the Mo blend is more sensitive to load and running time and the film thickness is developed.

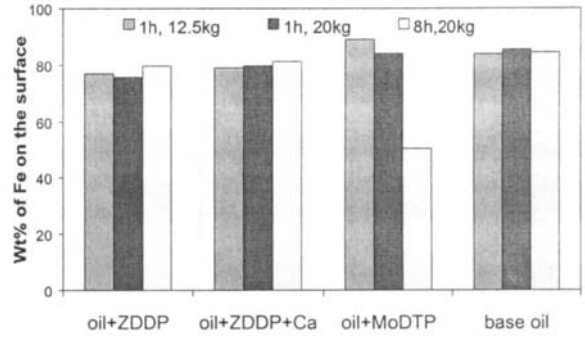


Figure 9. EDX Wt% of Fe on the film

4.2. ZDDP and ZDDP+Ca comparison

From the EDX results, it is possible to detect a reduction of Zn present in the film when Ca detergent is added to the system as shown in Figure 10a for 3 different conditions, which is in agreement with other authors [2]. The addition of Ca detergent to the ZDDP also tends to reduce the amount of S and P on the surface at all conditions as shown in Figure 10b and c which is confirmed by findings by Smith *et al.* [1]. A patchy film was found and lower surface coverage on ZDDP+Ca than on ZDDP. This is also verified by the slightly higher amounts of Fe detected on the ZDDP+Ca.

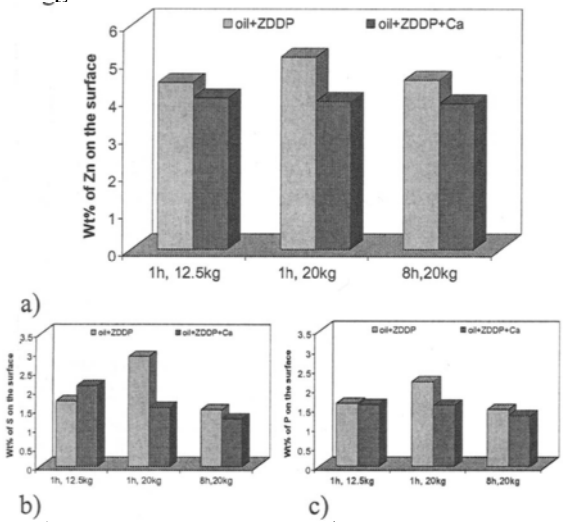


Figure 10. EDX analysis Wt% of a) Zn, b) S and c) P present at the surface.

XPS analysis demonstrates that the S 2p and P 2p species are sulphides at 162eV and polyphosphates at 133.64eV binding energies through the film thickness for both ZDDP and when Ca detergent is added to the ZDDP system. Ca is detected trough the film thickness as $CaCO_3$ at 347.1eV, only showing reduced quantities at the inner film. $CaCO_3$ is known to be abrasive so this is confirmed by the increases in friction and more wear probably when Ca is added to ZDDP.

XPS analyses show that the Zn peak position also do not change through the film as it is etched for both ZDDP and ZDDP+Ca as shown in Figure 11. Both films contain ZnO at 1022eV binding energy. Quantities of ZnO at the surface are higher for the oil+ZDDP but they seem to be equal at the inner film (45 min etch) for both additives.

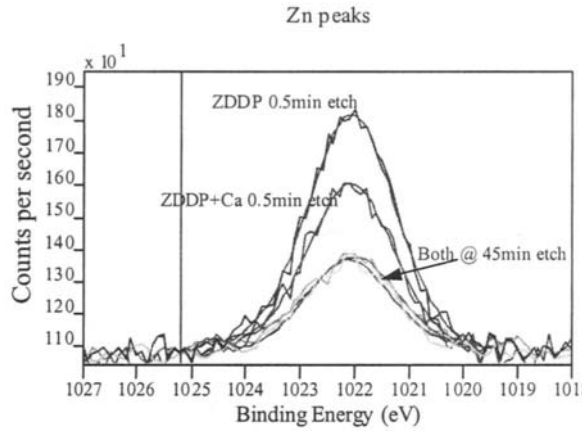


Figure 11. Zn 2p XPS peaks on the oil+ZDDP and oil+ZDDP+Ca films at 0.5 and 45 minutes etch

4.3. Main differences between ZDDP and MoDTP additives

From EDX analysis on 1 hour test samples it was found that ZDDP consistently shows higher percentages of S(S/Mo) and P than MoDTP even when S+Mo is considered in the MoDTP. In contrast, in the 8 hours tests at higher load, much higher P and S/Mo Wt% concentrations are found on the MoDTP than on ZDDP samples. This indicates the formation of a richer or thicker film, confirmed by the lesser amounts of Fe, detected in the film.

Fe peaks for the bearing steel plate, ZDDP and MoDTP films are compared at different XPS etching times in Figure 12. It is possible to detect a Fe_2O_3 (doublet at 710eV and 723.2eV) at the surface of both films and as the etching progresses, the peak moves towards the Fe 2p (707eV- 720eV) of the reference plate which was thoroughly cleaned and was not exposed to oil.

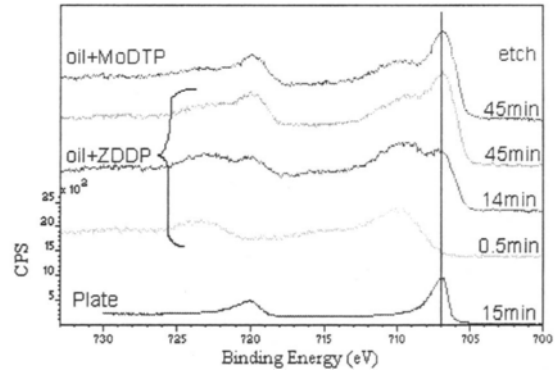


Figure 12. Comparison of XPS Fe 2p peaks

XPS analysis gives an insight of the film composition as it is etched for longer periods as shown in Figure 13. At the surface of the ZDDP film (0.5min etch) it is possible to detect a high percentage of C and O and low Fe content. There are changes in the concentration of P, S and Zn on the surface of the film at 0.5 and 1 min etch which are slightly reduced but remain quite constant up to 45 minutes etching time. As the etching time increases the amounts of O is reduced and the Fe increases, detecting more of the metal plate.

As Zn, S and P was detected up to 45 min etch, an extra 1 hour 20kg tests specimen on oil+ZDDP additive was further Ar^+ etched for 80, 100 and 120min. At each stage, a XPS survey scan was carried out to detect at which point the S and P was completely removed. Results show that the film is very tenacious, after 60min etch the P peak was still detected and the S peak still visible until 100 minutes Ar^+ etch.

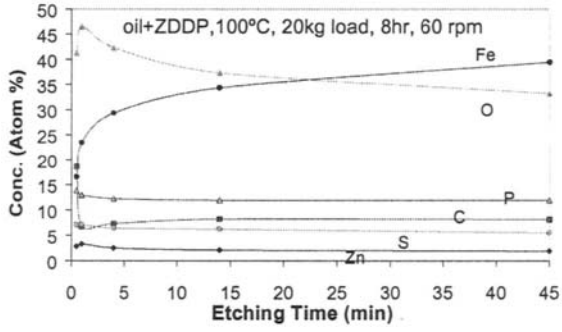


Figure 13. Oil+ZDDP film composition.

A distinction between Mo and S concentrations was possible using XPS and gave a ratio of Mo/S 1.83 showing that the Mo and S species are not present solely as MoS_2. S is likely present as iron sulphide in addition to MoS_2. Apart of having MoS_2 in the film Mo is present in the form of MoO_3. The S concentration remains quite constant throughout the film up to 45min etch. P decreases and O and Mo increase in the inner film as shown in Figure 14 at which point the Fe concentration detected from the plate is high, which agrees with work by Martin *et al.* on MoDTC [13] and Unnikrishnan *et al.* [4].

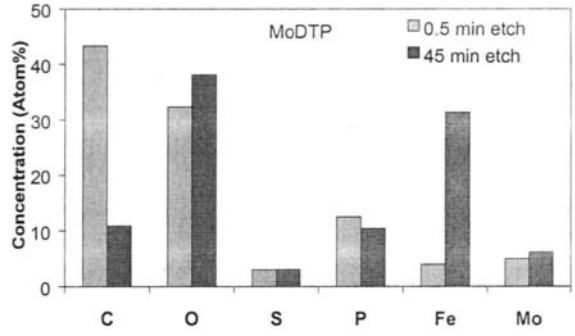


Figure 14. MoDTP film composition for 20kg, 8 hours

A comparison of the elemental composition of the 45 minute etched film at 20kg load for 8-hour tests for ZDDP and MoDTP film is presented in Figure 15. For MoDTP, lower concentration of Fe is detected indicating less signal from the substrate and hence a thicker film. Also, lower amounts of P and S in the inner film (45 min etch) and higher concentrations of Mo and O.

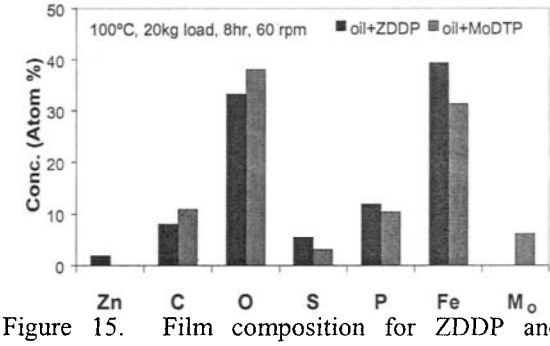


Figure 15. Film composition for ZDDP and MoDTP at 45 min etch

XPS analysis shows how the Mo changes from the outer to the inner film and how the concentration of the Mo slightly changes through the film thickness. It is highest at the surface then reduces as the film is etched and increases again close to the metal surface. A detailed analysis of the Mo 3d peak (231.7eV and 228.5eV) show that the outer layer (0.5min etch) consists of Mo oxides and sulphides. In the inner film (45min etch) some traces of oxides and sulphides still remain as shown in Figure 16. The low intensity S 2p (225.8eV) indicates metal sulphide and Mo 3d (235.3eV) indicates MoO_3.

The S and P species for both additives ZDDP and MoDTP, no changes in peak position were detected on the XPS analysis.

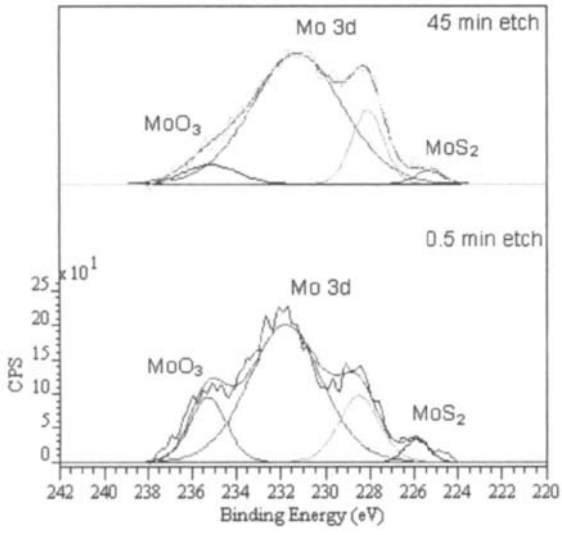


Figure 16. Mo peaks for oil+MoDTP

5. CONCLUSIONS

Under the conditions described in the previous paragraphs, the bearing steel tested (BS EN 31) shows that the friction and wear develop a linear relationship.

Interactions between the Ca-based detergent and ZDDP additives exist and lead to changes in friction responses and in film formation. They reduce the amount of S, P and Zn on the tribo film, hence increase friction coefficients due to the less rich film and the formation of $CaCO_3$.

Time is an important parameter in wear film (additives) studies and it has been shown that the reaction kinetics for ZDDP, ZDDP+Ca and MoDTP are contrasting.

MoDTP shows less S and P and higher O, C and Mo in the film. The lower Fe content is an indication of a thicker film, which agrees with the lowering of the friction coefficients with the formation of MoS_2.

In all conditions, the organo molybdenum additive (MoDTP) gave optimum friction and wear performance.

Constant load isothermal tests give useful information but need to be complemented with other tests.

6. REFERENCES

1. G.C.Smith, J.C.Bell, "Multi-technique surface analytical studies of automotive anti-wear films", Applied Surface Science 144-145 (1999), 222-227.
2. P.A.Willermet, D.P.Dailey, R.O.Carter III, et al, "The composition of lubricant-derived surface layers formed in a lubricated cam/tappet contact II" Tribology International, Vol. 28, No.3. 163-175, 1995.
3. J.A.Williams, "Engineering Tribology", Oxford Science Publications, 1994.
4. R.Unnikrishnan, M.C.jain, A.K.Harinarayan, A.K.Mehta, "Additive-additive interaction: an XPS study of the effect of ZDDP an the AW/EP characteristics of Molybdenum based additives", Wear 252, 240-249, 2002.
5. Y.Yamamoto, S.Gondo, T.Kamakura, N.Tanaka, "Frictional characteristics of molybdenum dithiophosphates", Wear, 112, 79-87, 1986.
6. A.Morina, A.Neville, J.Green, M.Priest, "Assessing friction, wear and film formation characteristics in formulated lubricants in severe to moderate boundary lubrication conditions", 29th Leeds-Lyon Symposium on Tribology, 2002
7. S Jahanmir, "Wear reduction and surface layer formation by a ZDDP additive", Journal of Tribology, 1986.
8. S Sheasby, T A Caughlin, A G Blahey and K F Laycock, "A reciprocating wear test for evaluating boundary lubrication", Tribology International, 23, No 5, pp 301-307, 1990
9. J.Graham, H.Spikes, "The behaviour of molybdenum dialkyldithiocarbamate friction modifier additives", Lubrication at the frontier / D.Dawson et al (Editors)
10. Z.Yin, M.Kasrai, M.Fuller et al, "Application of soft X-ray absorption spectroscopy in chemical characterizacion of anti-wear films generated by ZDDP", Wear, 202, 172-191, 1997.
11. Ph.Kapsa, J.M.Martin, C.Blanc, J.M.Georges. "Antiwear mechanism of ZDDP in the presence of Calcium Sulfonate detergent" Transactions of the ASME Vol. 103, October 1981.
12. C.H.Bovington, "Chemistry and technology of lubricants", Edited by R.M.Mortier, S.T.Orszulik, Second edition, 1997.
13. J-M.Martin, C.Grossiord, K.Varlot, B.Vacher, J.Igarashi, "Synergistic effects in binary systems of lubricant additives: a chemical hardness approach", Tribology Letters, 8, 193-201, 2000.

This Page Intentionally Left Blank

SESSION III
ELASTOHYDRODYNAMIC LUBRICATION (1) - DYNAMIC EFFECTS

Chairman : Professor H A Spikes

Paper III (i) **Dynamic Effects in EHL Contacts**

Paper III (ii) **Comparison of Transient EHL Calculations with Start-up Experiments**

Paper III (iii) **Comparison of Transient EHL Calculations with Shut-down Experiments**

Paper III (iv) **Optical Interferometric Observations of the Effect of Moving Transverse Asperities on Point Contact EHL Films**

Paper III (v) **Mechanism of Dimple Formation under Elastohydrodynamic Conditions**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Dynamic Effects in EHL contacts

C. J. Hooke

School of Engineering, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Virtually all EHL contacts operate under conditions where the load, surface velocities and contact radii change with time. In gears the load, slip and effective contact radius change over the tooth meshing cycle; in cams the surface velocities, load and contact radii change, often very rapidly, during cam rotation while in rolling element bearings the load changes continuously as the balls or rollers rotate around the track but may change more rapidly if the external load alters.

Predictions of film thickness (and pressure distribution) are relatively straightforward when conditions are constant but under dynamic conditions a separate, lengthy EHL analysis is required for each situation. This paper explores dynamic effects in EHL contacts and attempts to define criteria that can be used to decide where steady state results will give reasonable estimates of the contact's behaviour and where a more detailed analysis is required.

It examines the way in which clearances and pressures change when the contact conditions alter. Detailed results are presented showing the effect of changing the entrainment velocity and load. It is shown, for example, that sudden increases in load produce large, localised increases in film thickness and significant increases in the maximum hydrodynamic pressure and stress. It is also shown that high frequency load variations can produce very large changes in clearance, maximum pressure and contact stress. Use of a quasi steady state analysis in these types of situation could overestimate the minimum film thickness and underestimate the maximum stress.

1. INTRODUCTION

The behaviour of EHL contacts under steady operating conditions is well understood with theoretical predictions of clearances and pressures being readily available [1, 2]. The clearance predictions have, largely, been confirmed by experimental measurements using optical interferometry [3, 4]. However, many practical contacts operate under conditions where the load, entrainment velocity and contact radius change with time. For example in cams all three can change very rapidly and it is common for the direction of entrainment to reverse near the nose of the cam. In gears, the contact radius and entrainment velocity change continuously over the tooth contact while there are rapid changes in load as the number of teeth in mesh alters. Even in rolling element bearings the load on any element changes continuously as the roller or ball moves around the track. In addition to this, all contacts may be subjected to external oscillatory forces that can produce large, high frequency variations in contact load and entrainment velocity.

A number of analytical studies of non steady EHL are available. For example, Dowson et al. [5] examined the lubrication of a typical cam over a full operating cycle. Hooke [6] looked at the behaviour of contacts where the entrainment velocity reversed while Winjnant [7] carried out a detailed study of the stiffness and damping in point contacts. There have also been a limited number of attempts to obtain experimental results under dynamic conditions [8] but these have had mixed success largely due to the extreme difficulty of controlling the dynamic behaviour of optical test rigs.

It remains the case, however, that most EHL contacts are treated as running under steady state conditions. The effects of the variation of the operating conditions are generally ignored. In addition, most of the dynamic analyses have concentrated on the thickness of the fluid film with little attention being paid to the contact pressures and the stresses in the two surfaces.

The aim of this paper is to examine a number of dynamic contacts and to calculate the variation of film thickness, contact pressure and subsurface stress with time. It is hoped to offer some guidance as to when a quasi-steady treatment is likely to be acceptable and where more accurate analyses are required.

To do this the effects of radius, entrainment velocity and load changes are isolated and each examined separately. For radius changes a step change in radius is analysed while for entrainment velocity both step and oscillatory variations, together with entrainment reversal, are considered. For load changes step and oscillatory load variations are examined.

All results are for line contacts since most practical contacts are either line contacts or have large ellipticity ratios enabling an equivalent line contact to be substituted [1].

In order to enable the results to be compared with each other the operating conditions have been kept at P = 20, S = 10 with a Hertz contact pressure of 1GPa. A Newtonian fluid characteristic has been assumed, while Dowson-Higginson [9] density characteristics and Roelands [10] pressure viscosity characteristics have been adopted.

2. CHANGES IN ENTRAINMENT VELOCITY

Three types of change in velocity were considered. The first was a step change in which the entrainment velocity was either instantaneously halved or doubled. The second was where a sinusoidal variation of entrainment velocity with a magnitude equal to 50% of the underlying velocity was superimposed on the steady motion. The third was a reversal of the entrainment direction – a type of motion for which the clearance change has been discussed in detail in [6].

2.1 Step changes

Fig. 1 shows the effect of instantaneously doubling the entrainment velocity. In these plots clearance and pressure are plotted against position, with the clearance being plotted downwards for clarity. Curves are displayed in isometric form for a number of time steps. The change of velocity occurs at t = 0 and entrainment is on the left of the figure.

It may be seen that doubling the entrainment velocity increases the film thickness being generated at the inlet to the contact. This increase is carried through the conjunction at the (new) entrainment velocity, while the original clearance is removed from the conjunction. Thus it takes a time approximately equal to b/u (the new velocity is 2u) for the clearance change to propagate through the contact. The change in film thickness is fairly sharp and the step change in clearance generates a small bump in the pressure distribution and this increase, because it is attached to the change in film thickness, is also carried through the contact at the entrainment velocity.

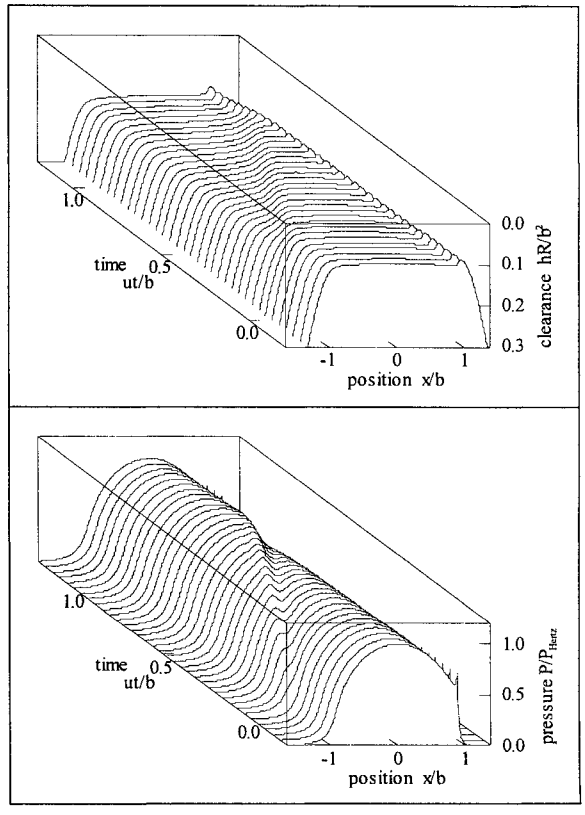


Fig 1 Variation of clearance and pressure with time, entrainment velocity doubled at t = 0.

This behaviour may be summarised by examining the variation of the minimum film thickness, the maximum pressure and the maximum sub-surface stress with time. This is shown for the present case in Fig 2a while the behaviour when the entrainment velocity is halved is given in Fig. 2b.

When the velocity is doubled there is a very small, almost instantaneous, increase in the minimum film thickness as the exit restriction adjusts itself to the new velocity. Thereafter the minimum film thickness remains almost constant

until the higher clearance generated at the inlet reaches the exit. The clearance then rises rapidly to its final value. In contrast, halving the velocity causes a rapid initial fall in clearance as the exit readjusts, followed by a slower continued fall as lower clearances are generated in the inlet. These clearances are transported through the contact almost unchanged giving a period of constant minimum film thickness before they reach the exit. As they do so the exit restriction again adjusts itself, this time to a lower clearance.

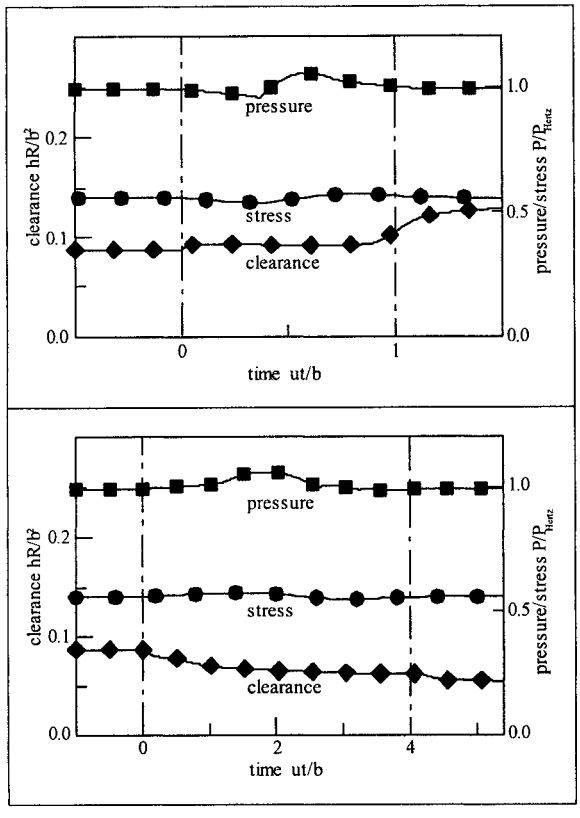


Fig 2 Variation of minimum film thickness, maximum pressure and maximum von Mises' stress with time.
a) Entrainment velocity doubled at t = 0.
b). Entrainment velocity halved at t = 0.

In both cases the step change in clearance produces localised disturbances to the pressure causing small increases in the maximum pressure as these pass the centre of the conjunction (where the undisturbed pressure is a maximum). The actual increase is small, in both cases being well below 5% of the Hertzian pressure. This perturbation in pressure, in turn, produces changes in the von Mises stress distribution in the underlying surface. However, since the spatial extent of the pressure changes is small compared with the Hertz width, the maximum changes in stress occur fairly close to the surface where the maximum Hertz stress is relatively small. The effect on the maximum stress, therefore, is minimal.

In addition to these results, extra cases where the change in velocity was more gradual were examined. In all these the clearance change was less abrupt and consequently the pressure and stress changes were reduced.

2.2 Sinusoidal variation

Sinusoidal variations in entraining velocity are likely to occur where the contact is subject to vibrational excitation as, for example, in a gear pair where there is some cyclic variation in torque.

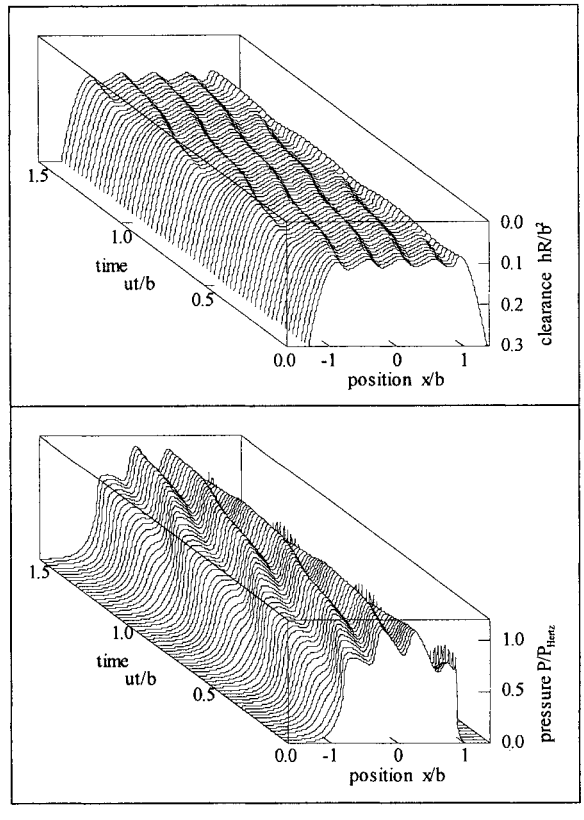


Fig. 3 Sinusoidal variation of entrainment velocity. $\delta u = 0.5u$. Period of oscillation, $T = 0.5b/u$.

In order to examine the contact under these conditions a sinusoidal variation of entrainment

velocity equal to 50% of the mean velocity was superimposed. A range of exciting frequencies was considered and Fig. 3 shows how the clearances and hydrodynamic pressures varied with time for the case where the period of oscillation, T, was given by uT/b = 0.5.

The variation in entrainment velocity affects the inlet region, generating clearance variations there. These are carried, almost unchanged, through the contact producing the clearance distribution shown. In order to maintain these clearance variations the pressure under the conjunction alters with a near sinusoidal pressure variation being generated. This variation moves with the same velocity as the clearance. These pressure variations, in turn, affect the stress, producing a moving stress field that is superimposed on the original Hertz distribution. For the particular case shown the wavelength of the clearance and pressure variation is 0.5b and the maximum change in stress lies around b/4π below the surface compared with the location of the maximum Hertz stress which lies around 0.8b below the centre of the conjunction. The net effect is that the maximum Hertz stress is almost unaffected by the sinusoidal variation in entrainment velocity.

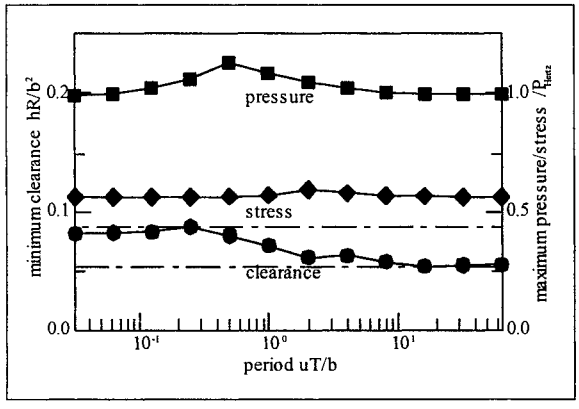


Fig. 4 Variation of minimum clearance, maximum pressure and maximum stress with period of oscillation. δu/u = 0.5.

The effect of changing the period of the oscillation is shown in Fig. 4 where the minimum film thickness, maximum pressure and maximum von Mises' stress are plotted against the period of oscillation. The horizontal chained lines represent the minimum film thickness under steady conditions. The upper line is for the mean entrainment velocity and the lower for the lowest entrainment velocity. As might be expected, for long periods of oscillation where the behaviour might be expected to be quasi-steady, the minimum film thickness does appear to follow the steady behaviour with the lowest minimum film thickness occurring just after the lowest velocity. The delay is due to the time taken to carry the clearance through the conjunction. For short periods the effect of the oscillation becomes small with the minimum film thickness tending to lie close to the value for steady motion. The slight drop in minimum film thickness does appear to be real and to arise because of the greater ease with which lower clearances are carried into the conjunction.

For both low and high frequencies the pressure distribution remains close to Hertzian with no significant pressure rippling. At low frequencies, where the wavelength is long compared to the contact width, this is because the contact is close to steady conditions at all times. At high frequencies, where the wavelength is short compared to the inlet length, the amplitude of the clearance variation transported into the contact reduces and with it the amplitude of the pressure ripples. It is only at the intermediate frequencies where the wavelength is longer than the inlet but shorter than the contact width that large amplitude variations in clearance form under the contact and it is here that the maximum pressure variations occur. Even at its maximum, however, the increase in maximum pressure is below 20%.

Because the largest pressure amplitudes are generated for wavelengths less than the contact width, the maximum variation in stress occurs near the surface. The overall effect of the oscillatory motion on the maximum stress is, therefore, small.

2.3 Entrainment reversal

In the two previous cases changes in the entrainment velocity produce little change in the maximum stress. However this is not always the case. Fig. 5 shows the clearance and pressure changes during a typical entrainment reversal. These results are for a value of P of 20 and S, defined using a reference velocity equal to $\sqrt{ab}$ in place of the usual entrainment velocity, u, of 10. a is the rate of change of entrainment velocity and b the semi-contact width.

Initially entrainment is on the left of the conjunction, changing to the right after the entrainment velocity passes through zero at t = 0. At first the reducing entrainment velocity generates

lower clearances at the inlet and these are carried into the contact. The effect is to produce a sharply sloping clearance profile under the conjunction. At the same time the exit restriction increases in magnitude producing a localised reduction in clearance there. Then, as the velocity falls close to zero, squeeze film effects take over maintaining a finite clearance under the conjunction.

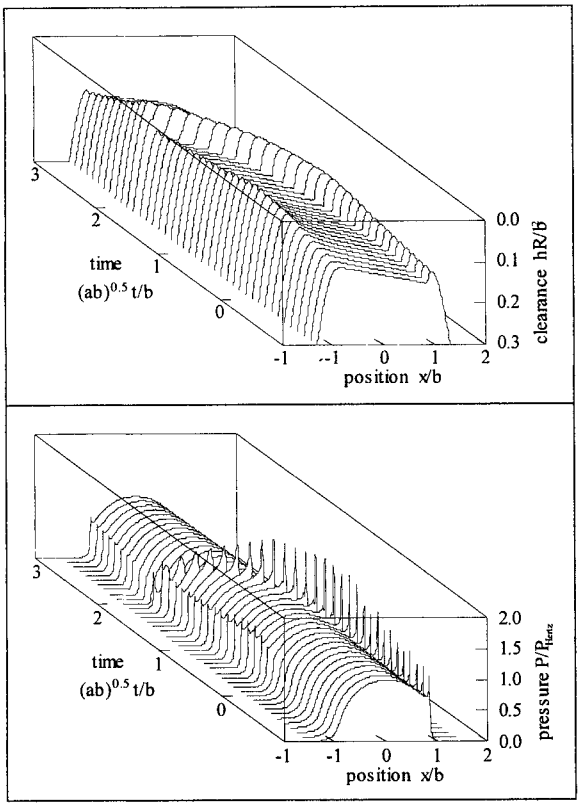


Fig. 5 Variation of clearance and pressure during entrainment reversal.

After reversal, as the velocity increases, the exit restriction is transported into the contact followed by larger clearances generated in the new inlet. The effect is to carry a localised region of low clearance through the conjunction. Eventually, as the old exit restriction passes out of the contact, the larger clearances generated by the increasing entrainment velocity cause the contact to revert to near steady state conditions.

In interpreting the time scale in the figures it may be noted that the entrained distance is given by $s/b = at^2/2b$, giving a time for the exit restriction to pass from the contact of $(ab)^{0.5}t/b$ of 2.

The increase in magnitude of the exit restriction initially increases the size of the exit pressure spike. Then, when the restriction is carried across the contact after entrainment reversal, the pressure spike is carried with it producing significant increases in the maximum contact pressure. It may be noted that the velocity at which the clearance moves is given by $u - \frac{h^2}{4\eta}\frac{dP}{dx}$. Thus the region of low clearance under the exit restriction tends to move more rapidly into the conjunction than the higher clearance region just in front of it. The effect is to sharpen the transition at the front of the restriction, increasing the height of the pressure spike.

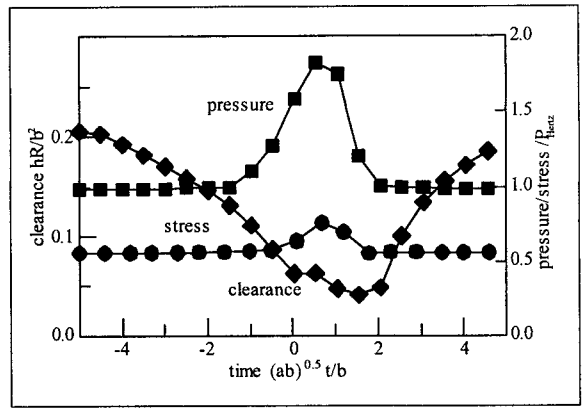


Fig. 6 Variation of minimum film thickness, maximum pressure and maximum stress during a reversal of entrainment. P = 20, S = 10. S is defined using a reference velocity equal to $(ab)^{0.5}$. Note that a is the rate of change of entrainment velocity with time. i.e. u = at.

These effects are summarized in Fig. 6 which shows how the minimum film thickness, maximum pressure and maximum stress change during the entrainment reversal. The convection of the original exit restriction through the contact is clearly visible as is the additional reduction in film thickness as it passes through the new exit. However the major feature is the very large increase in maximum contact pressure. This increase in pressure, in turn, generates extra stresses under the conjunction. Because the pressure spike is localised these stresses are close to the surface of the contact where the steady state Hertz stress is relatively low. However, because of the very large magnitude of the pressure spike the surface stress exceeds that of the maximum

Hertz stress and the consequent increase in maximum stress of around 34% is clearly visible

It may be noted that, in calculating the von Mises' stress, plain strain conditions have been assumed and the transverse stress, assessed using a value of Poisson's ratio of 0.28, has been included in the calculation.

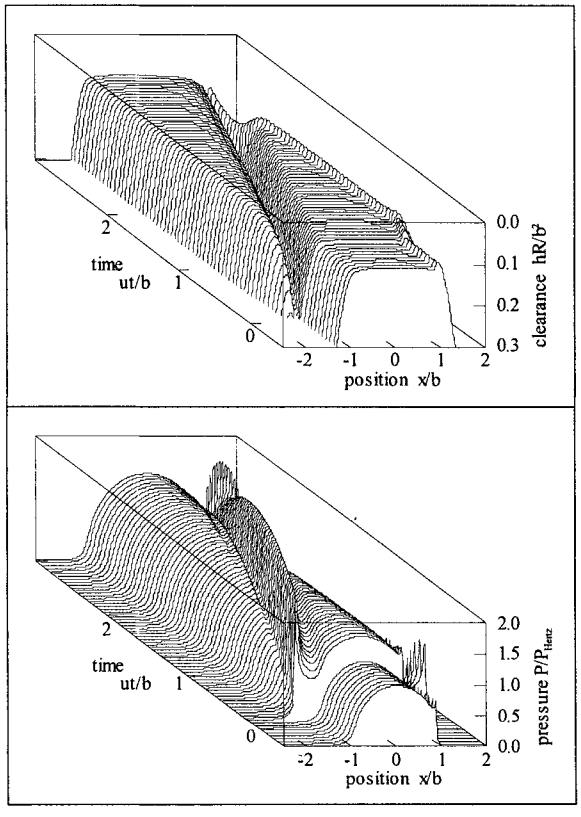


Fig. 7 Clearance and pressure variations following a doubling of the contact load

3. CHANGES IN CONTACT RADIUS

Two cases where there was a step change in contact radius at a point on the surface were examined. In one the radius was doubled; in the other it was halved. In both cases there was a gradual change in behaviour from one steady state condition to the next as the change in radius passed under the contact. No large changes in pressure or stress were observed and the minimum clearances always lay between their initial and final values.

4. CHANGES IN LOAD

Two cases of load change were examined. The first involved step changes in load with the load either being halved or doubled. The second involved sinusoidal variations in load of ±50% superimposed on the steady loading.

4.1 Step changes

Fig. 7 shows the clearance and pressure variations produced by an instantaneous doubling of the contact load. Since load variations produce little change in steady state film thickness it might be thought that changing the contact load would have a relatively small effect on the conjunction. However the effect is dramatic. As the load is doubled the contact width increases by 40% and the edges of the larger contact move down rapidly, trapping areas of high clearance each side of the conjunction. The high clearance region at the front of the contact is carried through the conjunction by the entrainment while that at the exit is removed.

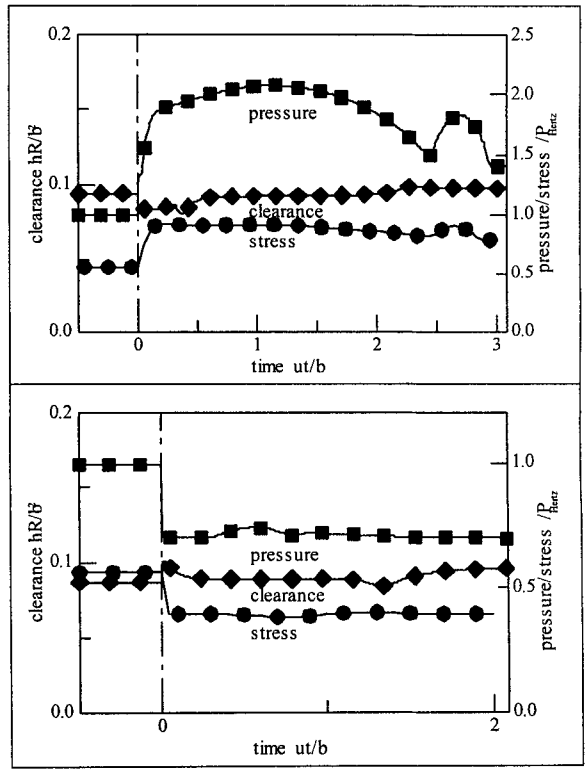


Fig 8 Variation of minimum film thickness, maximum pressure and maximum stress with time.
a) Load doubled b) Load halved

The effect is to produce a deep valley under the contact. In order to maintain this localised deformation of the surfaces, a local high pressure region must form above it and this peak, which can be seen clearly in Fig.7, is carried through the conjunction with the valley.

The behaviour is summarized in Fig. 8a where the minimum film thickness, maximum pressure and maximum stress are plotted against time. The increase in pressure between ut/b = 2.5 and 3 is due to the interaction of the transported valley with the exit restriction. Overall, doubling the load produces a slight, temporary reduction in minimum film thickness, a very marked rise in maximum pressure and a small (15%) increase in maximum stress. In contrast, halving the load, Fig 8b, almost instantly reduces the maximum pressure and stress to their new, steady state value while generating a very slight reduction in minimum film thickness.

The results shown in Fig 8 are for instantaneous rises and falls in load. However, the values are affected by the rate at which the load is applied since this affects the way in which fluid at the front of the contact is trapped. In order to investigate this, results similar to those shown in Fig 8a were obtained for a linear increase in load with time instead of the instantaneous step and Fig 9 shows how the transition time affects the behaviour of the contact.

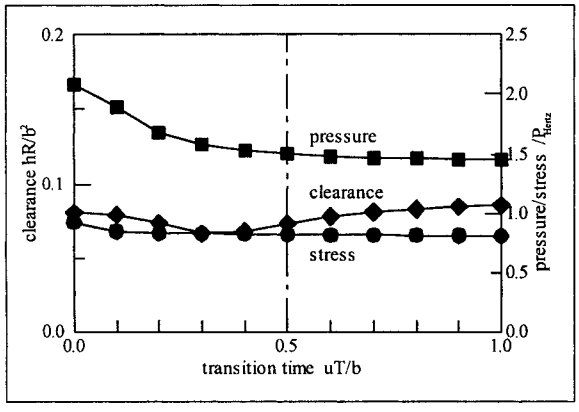


Fig. 9 Effect of transition time on minimum film thickness, maximum pressure and maximum stress. Load doubled.

It may be seen that increasing the transition time significantly reduces the maximum pressure while even a very short ramp effectively removes any increase in maximum stress. The vertical chained line corresponds to the transition time, T, for which the rate at which the contact width increases is equal to the steady entrainment velocity. At this ramp duration the effective entrainment velocity at the front of the contact doubles while the effective exit velocity vanishes. The reduction in minimum film thickness that occurs with transition times of around uT/b of 0.3 is due to clearance changes in the exit region. It appears to be associated with inwards entrainment in the exit region for short ramp times. For transition times below this minimum the inwards entrainment velocity is large enough to carry the clearances generated at the 'exit' into the contact. For transition times above the minimum the velocity is too low to do this.

Broadly similar effect occurs when the load is reduced. The reduction in load reduces the contact width and the effective entrainment velocity is reduced from its initial value by the rate of change of contact width, i.e. to u – db/dt. For transition times shorter than uT/b of 0.4 the entrainment velocity temporarily reverses while for ramp times greater than this the effective entrainment velocity remains positive but is reduced. Where the velocities match and entrainment disappears this produces a reduction in minimum film thickness of around 20%; for other ramp times the reduction is less.

4.2 Sinusoidal load variation

The effect of changes in contact width produced by load variations is, perhaps, seen more clearly when sinusoidal load variations are considered. Load variations of this type are likely to occur either where high frequency excitation is present or because of the system's response to sudden changes in external load.

The particular case shown in Fig. 10 is for a dynamic load equal to half the steady load with a period of oscillation given by uT/b = 1. For an entrainment velocity of 5m/s and a semi-contact width of 0.25mm this corresponds to a period, T of 0.05ms or a frequency of 20kHz.

The variation of the contact width is clearly visible and, when the width is increasing, the effective entrainment velocity rises, generating larger clearances at the inlet. Conversely, when the load is reducing and the contact width decreasing the effective entrainment velocity falls and lower clearances are produced. These clearances are carried through the conjunction producing the clearance variations seen in the figure.

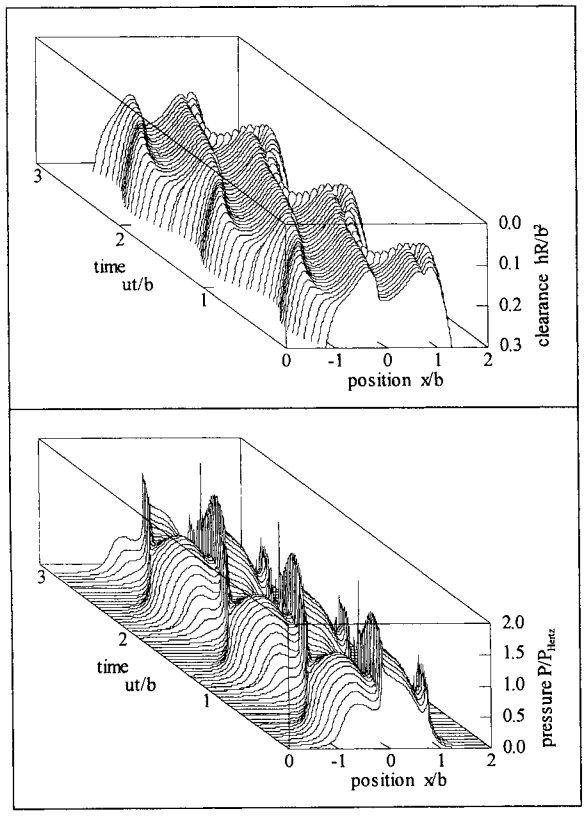


Fig. 10 Sinusoidal variation of load. $\delta W/W = 0.5$, $uT/b = 1$

In order to generate the surface distortions associated with these ripples the contact pressure must vary and the lower figure shows the resulting pressure distributions. It should be noted that the clearances do not vary sinusoidally across the contact. Because the convection velocity is given by $u - \frac{h^2}{4\eta}\frac{dP}{dx}$ smaller clearances tend to be carried into the conjunction more rapidly than higher clearances. The effect, as in entrainment reversal, is to tend to produce sharp changes in clearance which, in turn, generate pressure spikes. These may be seen in the lower figure as they move through the contact.

Pressure variations of this type generate substantial variations in the sub-surface stresses and Fig. 11 shows the distribution at one point in the load cycle. The upper curves in the figure give the pressure and clearance distribution in the contact at that time.

Because of the large magnitude of the pressure spikes the maximum stress occurs on the surface, under the spike, with a value of 0.75GPa, which may be compared with the maximum Hertzian stress of 0.69GPa - an 8.4% increase. Elsewhere the stress remains close to the steady state distribution. It may be noted that the maximum von Mises' stress varies with position in the load cycle with a minimum of 0.55GPa and a maximum of 0.84GPa

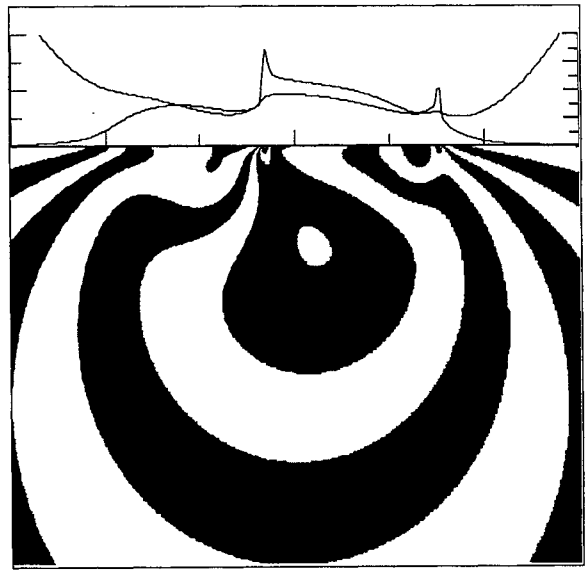


Fig 11 von Mises' stress distribution produced by sinusoidal load variation.

The contact's response is affected by the period of oscillation and Fig. 12 summarizes the behaviour. For long periods the minimum clearance, maximum pressure and maximum stress tend to values corresponding to steady loading at the highest load. The horizontal chained line in the figure corresponds to the minimum film thickness under these conditions. As the frequency is reduced there is a fall in the minimum film thickness as the reduction in effective entrainment velocity over the unloading part of the cycle becomes significant. This is associated with an increase in maximum pressure although the effect on maximum stress remains small. Then, as the period is reduced further, there is a steady increase in minimum film thickness and this appears to arise because the fluid trapped as the contact expands tends to override the reduction in entrained clearance on the reducing load. The oscillation of the inlet, therefore, appears to pump addition fluid into the conjunction.

There remain, however, large changes in clearance under the contact for moderate periods of oscillation and these generate significant pressures and stresses with the maximum increase in stress being found around uT/b = 0.5. At higher frequencies the clearance amplitude decreases rapidly and the pressures and stresses decrease to their steady state values.

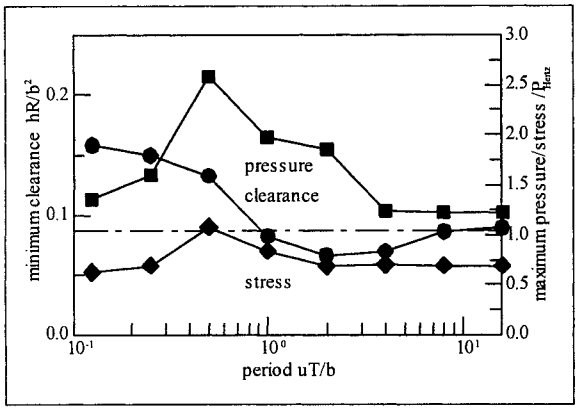


Fig. 12 Variation of minimum film thickness, maximum pressure and maximum von Mises' stress with period.

The transition from long to short period behaviour seems to be related both to the velocity with which the contact expands and contracts and to the wavelength of the clearance oscillation carried through the contact, uT. For periods greater than uT/b = $\pi/2$ the entrainment velocity is always positive and, apart from the reduction in clearance close to this point transition, conditions tend to be quasi-steady state. For shorter periods, entrainment reverses each cycle giving rise to complex dynamic behaviour. Initially the effect is to increase the pressures and stresses since, for a given magnitude of clearance variation, pressures are inversely proportional to wavelength. However as the period is reduced only part of the increased clearances generated as the load increases enter the contact before the load reduces. This appears to reduce the amplitude of the clearance variations and with them the pressure and stress changes. It is, perhaps, worth noting that it seems to be the ratio of the oscillation wavelength to the inlet length that controls the short period behaviour with oscillations longer than the inlet length entering the contact and with shorter wavelength oscillations being flattened.

5. DISCUSSION

Although only a limited number of cases have been examined some tentative conclusions may be drawn.

First, where the rate of change of the operating conditions is relatively slow, quasi-steady results give acceptable estimates of the clearances, pressures and stresses. This includes all changes in effective contact radius.

Second, where there are more rapid changes in entrainment velocity, quasi-steady results appear to give adequate estimates of the minimum film thickness under most circumstances. Where they do not, the lowest clearance is underestimated. Generally the maximum stress remains close to the Hertz value. The only case found where changes in entrainment velocity significantly affected the maximum stress was where there was a reversal of the entrainment direction. Here the increase in stress was produced by the very large pressure spike generated as the original exit moved across the contact after entrainment reversal. In practice since most entrainment reversal occurs in contacts where there is significant slip it is likely that thermal and non-Newtonian effects will drastically reduce the magnitude of the spike. It appears unlikely, therefore, that any changes in entrainment velocity will produce significant increases in stress above the Hertz value.

Third, changes in load can produce significant reductions in minimum film thickness and increases in maximum stress. As the load changes the contact region expands and contracts. This produces variations in the effective entrainment and exit velocities which, in turn, generate large variations in clearance at the inlet and exit. As the inlet variations are carried through the contact they can generate localised pressure variations that affect the stress distribution. Because these variations are far narrower than the contact width, the increase in stress occurs close to the surface.

It appears that the maximum rate of expansion (or contraction) determines whether quasi-steady results can be used. Reduced clearances and increased pressures are found where the expansion or contraction rate exceeds about 25% of the entrainment velocity. The maximum stress is less sensitive and appears to stay close to the Hertz value provided that the expansion rate does not exceed the entrainment velocity.

In addition to this it may be noted that, for both load and clearance variations, high frequency

oscillations did not affect the pressure or stress distributions (although high frequency load changes did seem to increase the clearance). Basically any oscillation with a wavelength (defined by the product of the period and the entrainment velocity, uT) that was shorter than the inlet length of the EHL contact was smoothed.

REFERENCES

1 Engineering Sciences Data Items 85027 and 89045, 'Film thicknesses in Lubricated Hertzian Contacts Part 1: Two dimensional contacts (line contacts)' and 'Part 2: Point contacts'. ESDU International 1985 and 1989.

2 Venner, C. H. 'Multilevel Solution of the EHL Line and Point Contact Problems', PhD thesis, University of Twente, Enschede, 1991.

3 Koye, K. A. and Winer, W. G., 'An experimental evaluation of the Hamrock and Dowson minimum film thickness equation for fully flooded EHD contacts', A.S.M.E. J. Lubric. Tech., 1980, 103, 284-294.

4 Moore, A. J., 'Nanometer Elastohydrodynamic Lubrication', Proc. Leeds-Lyon Symposium on Tribology, Sept 1996, 'Elastohydrodynamics-'96 Fundamentals and Applications in Lubrication and Traction', Elsevier, 1997, 57-67.

5 Dowson, D., Taylor, C. M. and Zhu, G. A. 'A transient elastohydrodynamic lubrication analysis of a cam and follower', International Conference on *Frontiers of Tribology*, 1991, (Institute of Physics, Stratford-upon-Avon).

6 Hooke, C. J., 'The minimum film thickness in lubricated line contacts during entrainment reversal - general solution and the development of a design chart', Proc. Inst. Mech. Engrs, 208(J), 1994, 53-64.

7 Wijnant, Y., 'Contact Dynamics in the field of Elastohydrodynamic Lubrication', Ph. D. thesis, University of Twente, 1998.

8 Sugimura, J. and Spikes, H. A., 'Technique for Measuring EHD Film Thickness in Non-Steady State Contact Conditions', Proc. Leeds-Lyon Symposium on Tribology, Sept 1996, 'Elastohydrodynamics-'96 Fundamentals and Applications in Lubrication and Traction', Elsevier, 1997, 91-100.

9 Dowson, D, Higginson G. R. and Whitaker, A. V., 'Elasto-hydrodynamic lubrication – a survey of isothermal solutions', J. Mech Engng Sci., 4, 1962, 121-126.

10 Roelands, C. J. A., 'Correlational aspects of the viscosity-temperature-pressure relationship of lubricating oils', Ph. D. thesis, Technische Hogeschool Delft, 1966.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Comparison of Transient EHL Calculations with Start-up Experiments

M. J. A. Holmes, H. P. Evans, and R. W. Snidle,
Cardiff School of Engineering, Queen's Buildings, The Parade, Cardiff CF24 3TA, UK, e-mail: evanshp@cf.ac.uk

Keywords: elastohydrodynamic, transient, non-Newtonian, start-up.

ABSTRACT

The paper presents numerical results for the elastohydrodynamic lubrication of point contacts where the entrainment velocity is increased with time from an initial value of zero. The cases considered correspond to experimental work presented by Glovnea and Spikes [1] for the case of a stationary dry contact that is brought into elastohydrodynamic operation by a controlled angular acceleration of one of the contacting bodies.

The experiments modelled involve simple sliding of a steel ball on a stationary glass disk, with the entrainment velocity increased uniformly, at different rates, to a steady state value. The steady state characteristics of the test case, which is reached by all experiments after different times, are used to select appropriate values of the lubricant parameters for incorporating in a transient numerical analysis of the start-up test conditions.

The paper compares the numerical results obtained using a differential deflection point contact transient analysis with the detailed published experimental film thickness measurements presented by Glovnea and Spikes. Differences in the calculated and experimental behaviour are observed, particularly at the highest acceleration rates. An explanatory factor to account for the differences is proposed.

1. INTRODUCTION

Transient analyses of the Elastohydrodynamic (EHL) problem are necessary in attempting to develop an understanding of the lubrication of rough surfaces. The authors have developed a new point contact solution technique [2,3] based on a novel differential deflection method [4]. The need to validate a new approach against existing published numerical solutions has been achieved [2,3] using the well known solutions of Venner and Lubrecht [5]. It is also advantageous to validate these solutions by comparison with experimental results and that is the objective of the current paper and its companion [6].

Glovnea and Spikes [1,7] have recently published detailed results of transient point contact EHL situations achieved experimentally using their ball on transparent plane apparatus. The transient conditions considered consist of the shut-down of EHL action caused by the reduction of entrainment velocity under constant deceleration [7] and the start-up of EHL under a constant acceleration of the contacting components [1]. The start-up experiments and their comparison with numerical results are the focus of the present paper.

The experimental work [1] consisted of two configurations. In both of these the entrainment velocity was increased linearly from zero to a final steady state value, and a series of experiments was carried out corresponding to different rates of increase of entrainment velocity. In one configuration the velocity of both contacting components was maintained so as to give pure rolling conditions throughout the experiment. In the other configuration the disk component remained stationary throughout and the contact operated in simple sliding, i.e. with $\xi = 2$, with the ball velocity increased linearly from zero to twice the final steady state entrainment velocity. The kinematics of this second configuration are stated to be controlled more precisely than the first due to difficulties caused by the high inertia of the disk component in comparison with that of the ball. The numerical analyses presented in the current paper have thus been concentrated upon the simple sliding results in order to avoid ambiguities concerning the time history of the component velocities. The simple

sliding configuration introduces the complication of uncertainty in the contact's operating temperature which, although known at the beginning of motion, will be influenced by heat input to the stationary glass disk during the start-up. This difficulty was addressed in the isothermal analyses carried out by adjusting the assumed operating temperature such that the calculated final steady state result corresponded to the film profile information given in the experiment [1] for that phase of the motion. A non-Newtonian rheological model was used and its parameter values were also chosen so as to obtain a good fit with the experimental, steady state film profile.

2. THEORY

The Reynolds equation for the fluid is

$$\frac{\partial}{\partial x}\left(\sigma_x \frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\sigma_y \frac{\partial p}{\partial y}\right)-\frac{\partial(\rho \bar{U} h)}{\partial x}-\frac{\partial(\rho h)}{\partial t}=0 \qquad (1)$$

where the flow factors σ_x and σ_y ($\neq\sigma_x$) are obtained using the hyperbolic sine model proposed by Johnson and Tevaarwerk [8] as described in [9].

The novel feature of the method is that the elastic deflection, u, which gives rise to the elastic film thickness according to

$$h(x,y) = \frac{x^2}{R_x} + \frac{y^2}{R_y} + u(x,y) + h_0 \qquad (2)$$

is specified in differential form [4]. The equation thus specifies the Laplacian of the deflection in terms of the influence of the pressure distribution according to

$$\frac{\partial^2 u(x_i,y_j)}{\partial x^2}+\frac{\partial^2 u(x_i,y_j)}{\partial y^2} = \sum_{all\ k,l} f_{k-i,l-j} p_{k,l} \qquad (3)$$

Equation (2) is then cast in the form

$$\frac{\partial^2 h}{\partial x^2}+\frac{\partial^2 h}{\partial y^2} = \frac{1}{R_x}+\frac{1}{R_y}+\sum_{all\ k,l} f_{k-i,l-j} p_{k,l} \qquad (4)$$

The advantage of the formulation is the extremely rapid decay of the influence coefficients $f_{i,j}$ as the indices i and j increase from zero. This enables equations (1) and (4) to be solved as a coupled pair as described by Holmes et al [2,3]. The corresponding approach for the line contact case has been found to be extremely useful in solving moving roughness problems where the composite roughness of the surfaces is an order of magnitude or more greater than the smooth surface film thickness [10,11].

A finite difference Reynolds equation with the central difference formulation of the first order spatial derivatives and a Crank-Nicolson scheme for incorporating the time dependent term [3] was used for the analyses presented in the current paper, and the results obtained with an alternative finite element formulation were indistinguishable.

For the start-up configuration the Hertzian region is initially at a film thickness value of zero. This is achieved by adopting the Hertzian pressure and using equation (2) to obtain a central film thickness $h_c = 0$ by suitable adjustment of the constant h_0. Subsequently equations (1) and (4) are solved simultaneously for each timestep with the boundary values for equation (4) obtained from use of equation (2). This solution is obtained using a coupled iterative scheme [3], and for points where the calculated film thickness is negative, h is set to zero and only equation (4) is applied to calculate the contact pressure. This process maintains the Hertzian pressure when applied to all the contacting points. As the transient solution develops the pressure developed on the 'dry' part of the contact is seen to remain unchanged.

To check that the procedure adopted does not influence the rate at which the film is calculated to develop upstream of the remaining dry contact area, a test was carried out where the initial condition was a steady state solution for the EHL problem with $\bar{U}$ = 0.0075 m/s. This results in a very thin film thickness distribution in the Hertzian contact area with a central film thickness, h_c = 0.011 μm. For this test, therefore, a fluid film satisfying the Reynolds equation was present at all mesh points during the calculation. The differences in the way that the film thickness developed in this test case during the transient analysis and for the corresponding standard analysis where contact was specified initially were trivial.

3. LUBRICANT PARAMETERS

The non-Newtonian rheological model briefly discussed above was used with the assumption of isothermal behaviour of the oil. The test case used by Glovnea and Spikes [1] is moderately loaded in numerical terms having Moes and Bosma non-dimensional parameter values of M = 211 and L = 5.02 at their stated conditions. The computing domain used is $-2.5a \le x \le 1.5a$, $-2.0a \le y \le 2.0a$, with 50 nodes in each Hertzian dimension which is

sufficient to resolve smooth surface cases under these moderately loaded conditions. In the numerical solution the speed of the moving surface was increased linearly over time using a simple ramp with a series of acceleration rates corresponding to those used in the experimental work [1]. The acceleration rate of the entrainment velocity is then half this value in the simple sliding mode of the experiment. A timestep value of $\Delta t = 0.01$ ms was adopted for all the cases considered. This corresponds to splitting the acceleration ramp into 400 divisions for the highest acceleration of 50 m/s^2, and corresponds to a ball surface movement of 4 μm per timestep at the final steady state velocity. Tests with a finer timestep confirm that the results were not influenced by timestep size.

The oil parameters stated by Glovnea and Spikes [1] for the experiments modelled in the current work were as given in Table 1.

Table 1. Experiment parameters as given in [1]

R	9.5 mm
W	20 N
E_{ball}	210 GPa
E_{disk}	75 GPa
ν_{ball}	0.3
ν_{disk}	0.25
η_0 (40 °C)	0.445 Pa s
η_0 (80 °C)	0.052 Pa s
α (40 °C)	26.8 GPa^{-1}
α (80 °C)	19.6 GPa^{-1}
$\overline{U}$	0.2 m/s
ξ	2

As with the shut-down analysis reported by Holmes et al [4], some adjustment of the quoted oil parameters were necessary to obtain numerical results that correspond to the steady state film thickness profile given for the experimental contact. For the current paper this steady state is established at the end of the accelerating phase where the contact entrainment velocity is maintained at the steady state value of 0.2 m/s. The case considered for the numerical comparisons is the simple sliding configuration where the transparent disk is held stationary whilst the contacting ball is rotated according to the kinematic configuration examined. With this configuration it is stated to be easier to maintain a constant acceleration as it avoids the dynamic effects associated with changing the

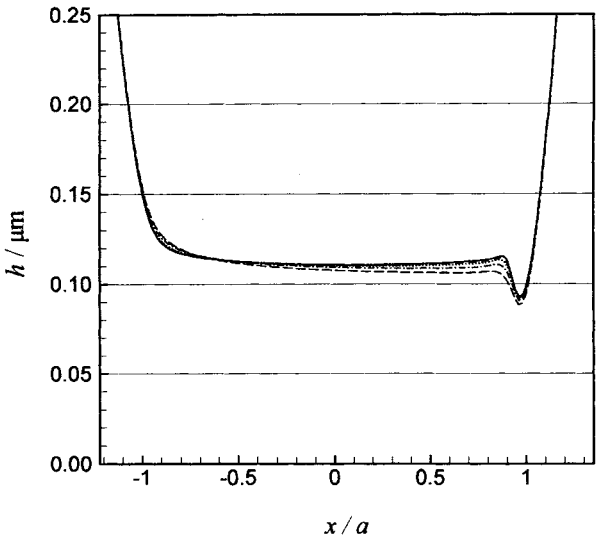


Figure 1. Centre line film thickness profiles for the steady state condition with $\tau_0 = 2$ MPa (lower h_c), 3, 5, 7, and 10 MPa (higher h_c).

rotational speed of the disk component which has a higher inertia than that of the contacting ball. Figure 1 shows steady state results obtained for simple sliding at the stated entrainment velocity for a range of values of parameter τ_0 with viscosity $\eta_0 = 0.061$ Pas. For the lubricant used in the start-up experiments there is less variation in the film thickness profile slope than for the lubricant used for the shut-down experiments analysed in [4]. A value of $\tau_0 = 5$ MPa was selected as giving a good match to the experimental profile and used for the transient numerical analyses.

4. RESULTS

The transient analysis was carried out for each of the four start-up accelerations used for the experimental work [1]. Figure 2 compares the central film thickness obtained during the numerical simulation of the start-up process with that observed experimentally [1]. During the numerical calculation the value of h_0 is held constant in equation (2) and this leads to an increase of about 13 % in the load supported during the start-up process. This does not cause any significant difference in film thickness response to that obtained by maintaining a constant load as is shown later in the paper. The actual load demand made of the contact during dynamic start-up is not known as dynamic effects may well change the load during acceleration.

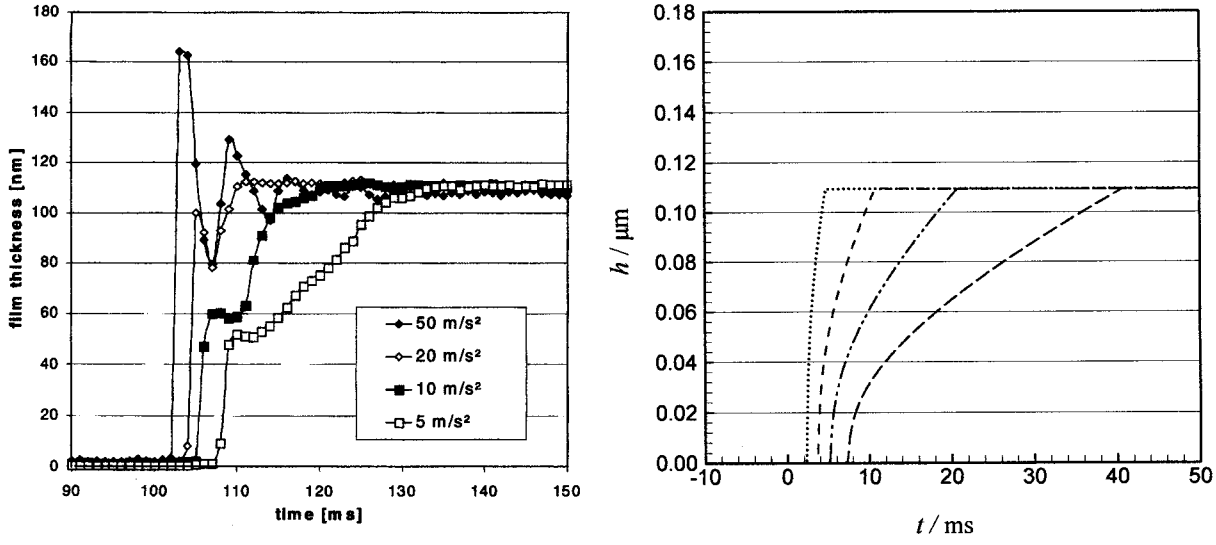


Figure 2. Variation of central film thickness during start-up. Left hand figure shows experimental results [1]. Right hand figure shows calculated numerical results for the four acceleration rates, ····· 50 m/s^2, - - - - - - - 20 m/s^2, — · — · — 10 m/s^2, — — — — 5 m/s^2

In figure 2 it is seen that the numerical results show a systematic increase in central film thickness once the wetted surface of the ball penetrates to that location. The rate of increase is greatest for the highest acceleration as would be expected, and the final steady state central film thickness obtained is identical for each case, as might also be expected. The experimental results show similarity at the lowest acceleration, but exhibit a different behaviour at the highest acceleration where the central film thickness measured has an oscillatory response before settling down to the steady state value after about 25 ms. Possible reasons for this oscillatory behaviour are discussed by Glovnea and Spikes [1] and in section 5 of the current paper.

Comparison of the calculated film thickness profile obtained from the numerical analysis with the experimental measurements at the highest acceleration is made in Figure 3. The profiles obtained 2, 3 and 35 ms into the start-up process are

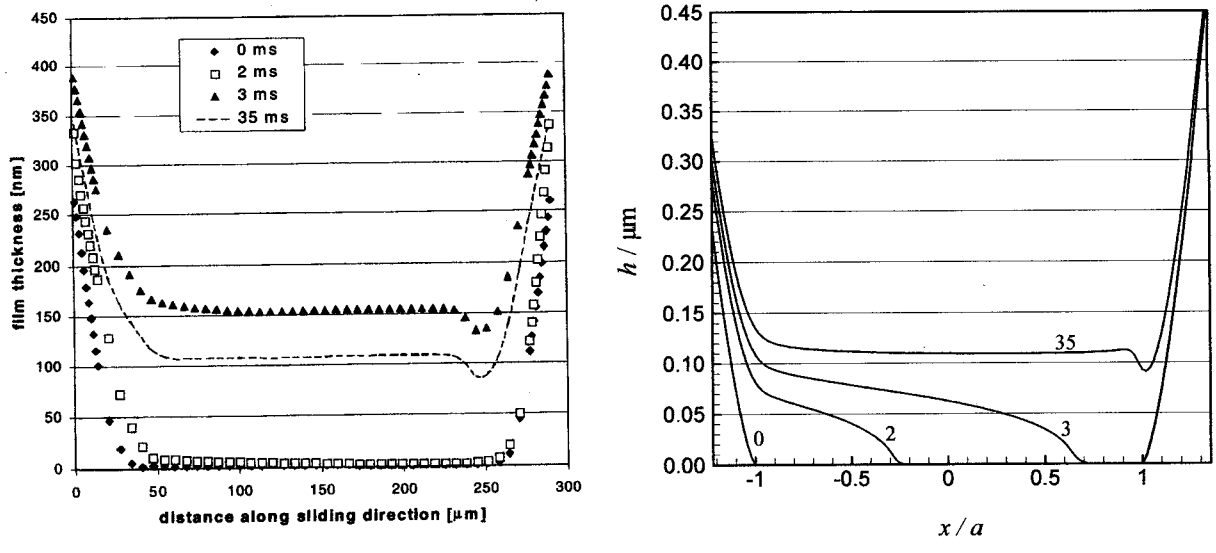


Figure 3. Entrainment centre line film profiles at start up for 50 m/s^2 acceleration. Left hand figure shows experimental results [1]. Right hand figure gives results of numerical analysis at times of 0, 2 ms, 3 ms and 35 ms after beginning of ball motion.

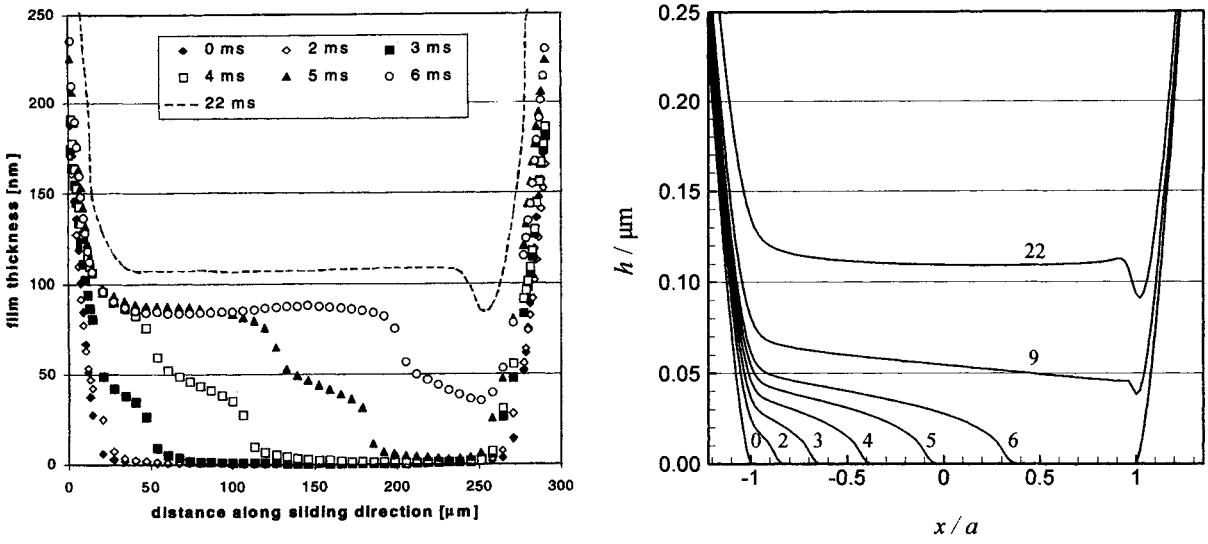


Figure 4. Entrainment centre line film profiles at start up for 10 m/s^2 acceleration. Left hand figure shows experimental results [1]. Right hand figure gives results of numerical analysis at times of 0, 2 ms, 3 ms, 4 ms, 5 ms, 6 ms, 9ms and 22 ms after beginning of ball motion.

shown and may be directly compared with the experimental results that are also given in the figure. For this case the experimental results exhibit an oscillatory central film thickness, as is shown in Figure 2. This behaviour can be seen to take place across the whole Hertzian contact. The 0.15 μm film thickness level seen at 3 ms is seen to extend over the whole contact area and exceed the steady state value by 50%. The numerical results are quite different with a progressive increase in film thickness at the times shown. The point on the entrainment centre line separating the dry and lubricated contact areas is seen to move forward in the numerical analysis at the entrainment velocity. Adopting the quoted acceleration of 50 m/s^2, lubricant initially at the edge of the Hertzian region and moving at the entrainment velocity will have moved forward by 0.1 mm after 2 ms and by a total of 0.225 mm after 3 ms. These are exactly the distances that the separation point has moved downstream in the numerical results at these times. The establishment of a much thicker film during this time in the experiment suggests that the entrainment velocity is possibly higher than supposed.

Figure 4 makes a similar comparison for the results obtained with a start-up acceleration of 10 m/s^2. The film thickness profiles obtained numerically form a family of curves which have a progressive increase in slope from the Hertzian inlet to the separation point between the wet and dry contact areas. The progress of the calculated separation point from inlet to exit during start-up is again entirely consistent with the distance moved at the entrainment velocity since motion began. For the elapsed times of 2, 3, 4, 5 and 6 ms shown the corresponding distances moved by lubricant travelling at the entrainment velocity are 0.02, 0.045, 0.08, 0.125 and 0.18 mm respectively. Making allowance for the Hertzian dimension of 0.134 mm, these distances correspond closely to the positions of the separation points at these five elapsed times. The position of the separation points for the experimental results are less clear cut and there is some uncertainty as to the position of the edge of the Hertzian region. For elapsed times of 2 and 3 ms there is a consistency between the distance moved by the ball and the position of the separation. For the later times, however, the separation has clearly moved in advance of the entrainment velocity. The 6 ms experimental profile shows oil entrained into the whole of the contact whereas the Couette entrainment of lubricant will not have progressed into the last third of the Hertzian area. Again, this suggests that the entrainment is faster than supposed in the experiment, but to a lesser extent than for the 50 m/s^2 case.

The film profile obtained numerically differs considerably from that given by the experimental measurement. In the experiment the profiles tend to exhibit two levels of film thickness with rapid transitions between the levels and between the lower level and contact. The 'front' associated with the

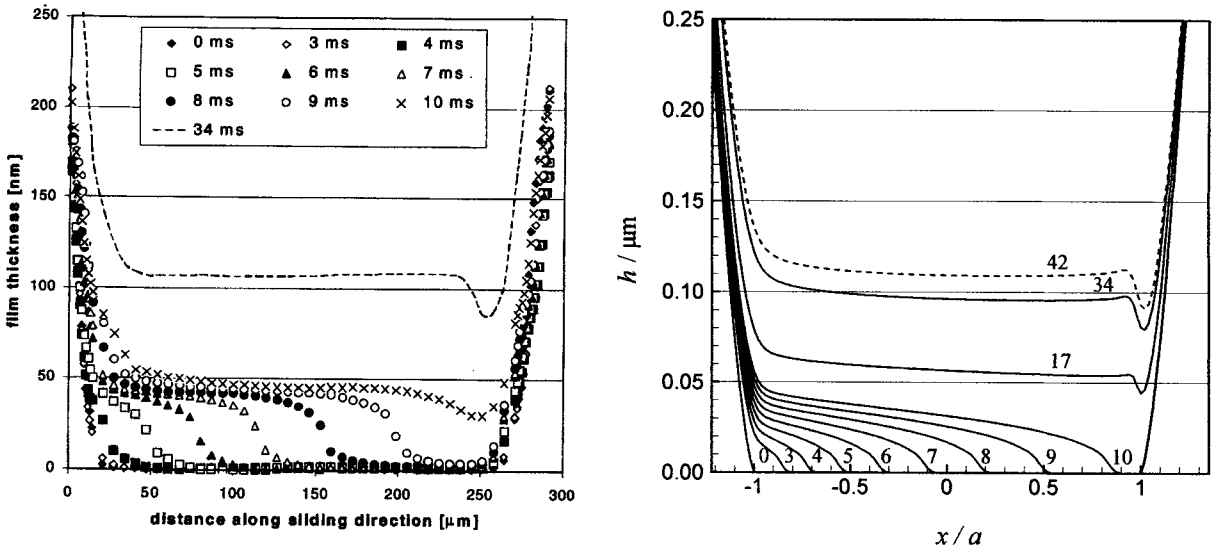


Figure 5. Entrainment centre line film profiles at start up for 5 m/s^2 acceleration. Left hand figure shows experimental results [1]. Right hand figure gives results of numerical analysis at times of 0, 3 ms, 4 ms, 5 ms, 6 ms, 7 ms, 8 ms, 9 ms, 10 ms, 17 ms and 42 ms after beginning of ball motion.

transition between the film levels can be clearly identified for the 4, 5 and 6 ms elapsed time profiles. The numerical film thickness is considerably smaller than that at the upper of these two levels for each of the elapsed times and broadly corresponds to the intermediate film thickness level seen experimentally. Steady state entrainment velocity is established at time 20 ms, and a further 2 ms ensures that all the oil within the contact area was entrained at the steady state velocity. Both numerical and experimental film thickness profiles are seen to be the same at elapsed time 22 m.

Figure 5 makes similar comparisons for the case having an acceleration of 5 m/s^2. The numerical results again are in the form of a family of curves which have a progressive increase in slope from the Hertzian inlet to the separation point between the wet and dry contact areas. The progress of the calculated separation point from inlet to exit during start-up is again entirely consistent with the distance moved by the ball since motion began. For the elapsed times of 3, 6, 8 and 10 ms, for example, the distances moved by the ball are 0.023, 0.09, 0.16, and 0.25 mm, respectively, which correspond to the separation locations. The numerical film thickness is again smaller than the experimental values, but the differences are less dramatic in this case. The experimental profiles tend to be almost constant from the inlet to the Hertz region to the position where they fall rapidly to an almost contacting condition. For this experiment the position of the dry/wet separation point is much more consistent with the motion of the ball than was the case at the higher accelerations.

The steady state profile is established in the experiment after 34 ms although the ball is still 15% below its steady state speed at that time. For the numerical result the 34 ms profile remains below the steady state value which is seen in Figure 5 at the 42 ms elapsed time.

For the 5 m/s^2 case Glovnea and Spikes [1] also present transverse film thickness profiles for specific elapsed times which are compared with the corresponding numerical results in Figure 6. The technique for measuring film thickness is carried out on one particular section so that the experimental results shown in Figure 6 have, of necessity, been obtained in a separate repeat experiment and have not been acquired simultaneously with the profiles of Figure 5. The film thicknesses obtained at the times for which the experimental results are given show a high degree of similarity in shape with the differences in central film thickness value remaining consistent with the differences seen in Figure 5.

The boundary between the dry and lubricated parts of the contact is illustrated by fringe interferograms for the 5 m/s^2 case [1], and comparison is made with corresponding numerical results in Figure 7. For the numerical result the area for which the contact remains dry is coloured black,

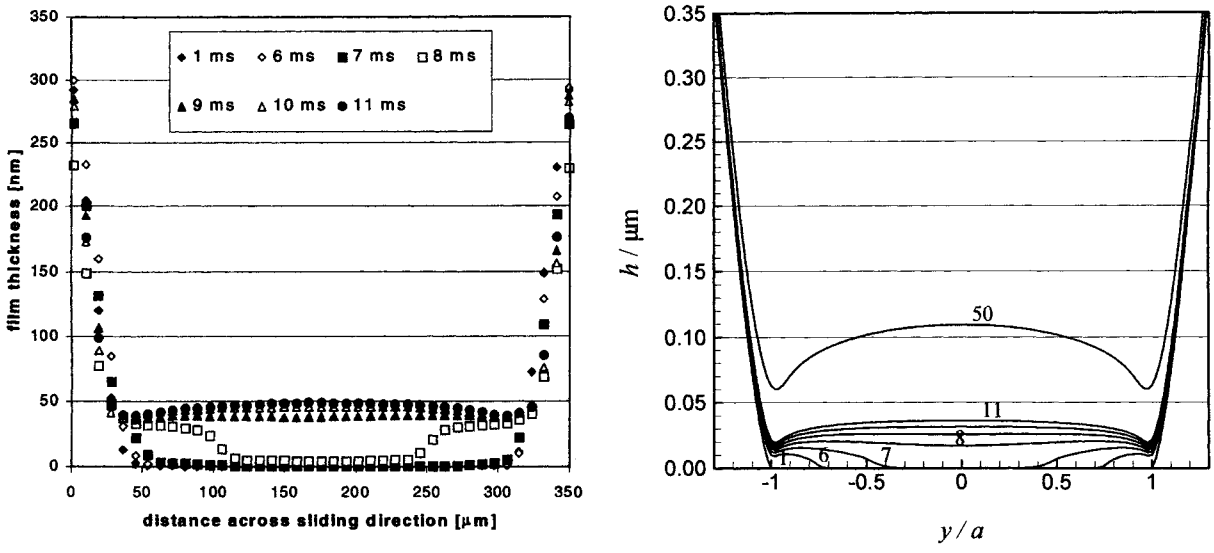


Figure 6. Transverse centre line film profiles at start up for 5 m/s^2 acceleration. Left hand figure shows experimental results [1]. Right hand figure gives results of numerical analysis at times of 1, 6 ms, 7 ms, 8 ms, 9 ms, 10 ms, 11 ms and 50 ms after beginning of ball motion.

whereas in the experimental interferograms contact will correspond to the fringe colour produced by the spacer layer. The position of the boundary separating the two parts of the contact may be seen as circular arcs that pass progressively through the Hertzian contact area, and the agreement between the boundary positions can be seen to be good.

5. DISCUSSION

In attempting to model the experiments numerically a decision regarding the elastic deflection is required. The value of the constant h_0 for the dry contact case is –1.885 μm, and for the steady state solution at the same load is -1.77 μm. The results presented in Section 4 have been obtained with h_0

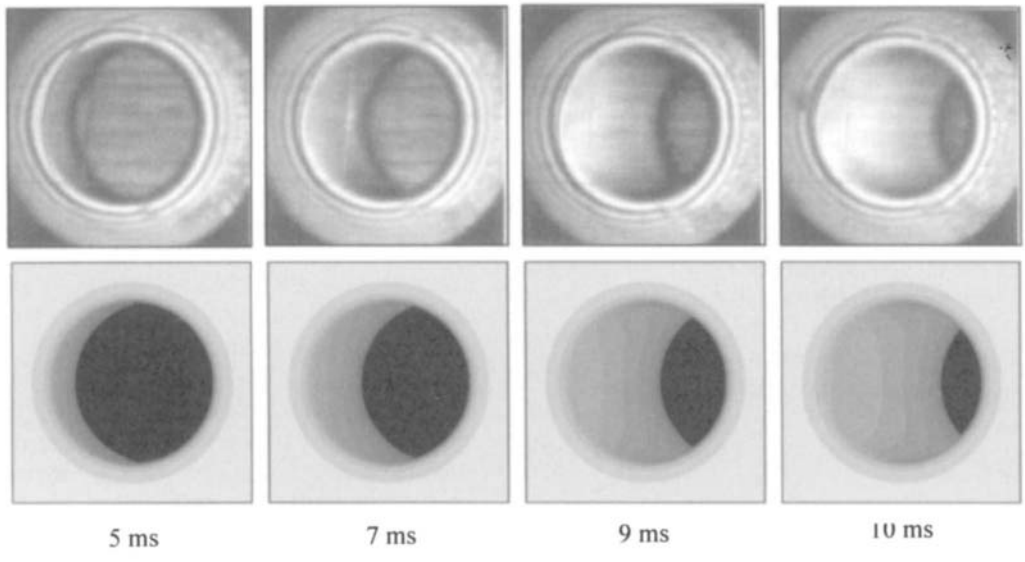


Figure 7. Comparison of interferograms obtained from the experiments at 5 m/s^2 acceleration (upper figures) [1] with calculated film contour plots at the corresponding times (lower figures) of 5 ms, 7 ms, 9 ms and 10 ms.

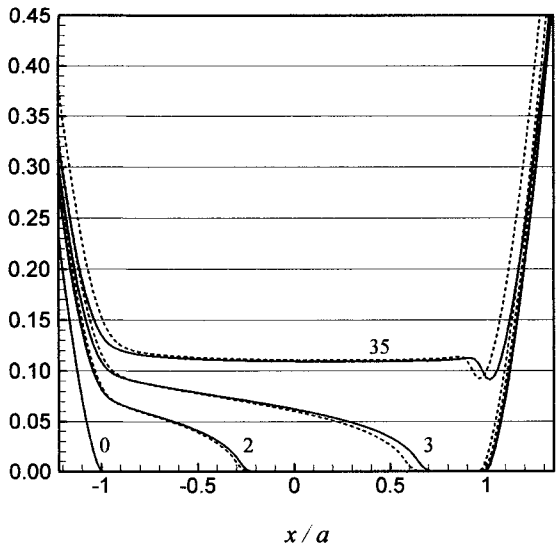


Figure 8. Centre line film thickness profiles obtained for the 50 m/s^2 case with fixed h_0 (solid curves) and fixed load (dotted curves).

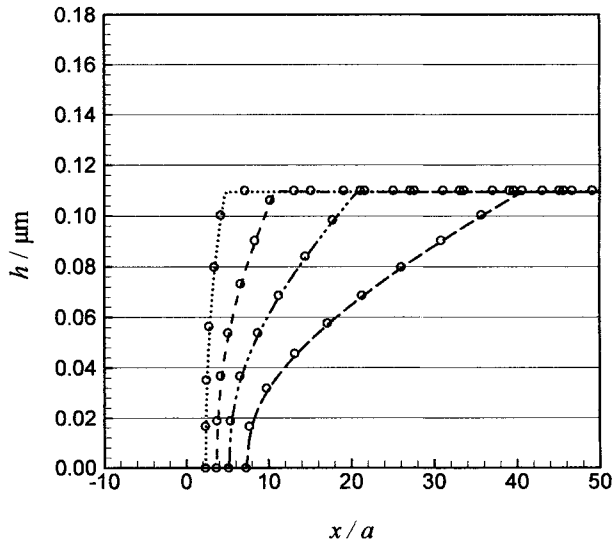


Figure 9. Comparison of h_c values obtained with the simple model (symbols) with those from the full numerical analysis for accelerations of 50 m/s^2 (left hand curve), 20 m/s^2, 10 m/s^2 and 5 m/s^2 (right hand curve).

fixed at -1.885 μm, which leads to a systematic increase in the load supported during the start-up process to a steady state value of 22.7 N. To test whether this factor can explain the differences observed between the numerical and measured films comparisons were made with results obtained where a fixed load is maintained. Although this may seem to be the preferred condition to adopt it is by no means clear that the load will be constant during the start-up process. Figure 8 compares the film profiles obtained with these two specifications for the 50 m/s^2 case. It is clear that whether or not the load is maintained makes little difference to the calculated films, and the fixed h_0 case simply leads to the increase in Hertzian area associated with the load change with no significant difference in the film profiles calculated.

The film thickness response at the centre of the contact for the numerical analysis shown in Figure 2 may be obtained from first principles utilising the following simple argument. If the quantity of oil entrained at the inlet to the Hertzian area is related to the steady state central film thickness by the expression

$$h_{ent} = h_{st}\left(\overline{U}/\overline{U}_{st}\right)^{0.68} \qquad (5)$$

and this oil is subsequently assumed to be transported at the entrainment velocity, then the central film thickness will vary with time during the start-up process. The speed dependence in equation (5) is taken from the formula of Chittenden et al [12]. Figure 9 shows the result of calculations based on this simple model for each of the acceleration rates. The figure also shows the corresponding results of the full numerical analysis taken from Figure 2. The strong correlation between the two sets of calculations indicates conclusively that this simple mechanism describes the essence of the way in which the Reynolds equation models the start-up situation for these four cases.

The simple model described above can also be used to calculate the film thickness profile at different times. A calculation for the 5 m/s^2 case is presented in Figure 10 together with the corresponding result from the full numerical solution. The profiles obtained from the simple analysis show good correlation with those from the full numerical analysis but with a tendency to overestimate the oil quantity entrained at the initial motion stages. This is seen in the simple model profile being in advance of the numerical model profile in the vicinity of the transition from lubricated to dry contact. This difference depends on the power (0.68) adopted in equation (5) and a higher power of 0.75, will reduce this difference considerably. There is excellent agreement on the position of the transition between lubricated and dry

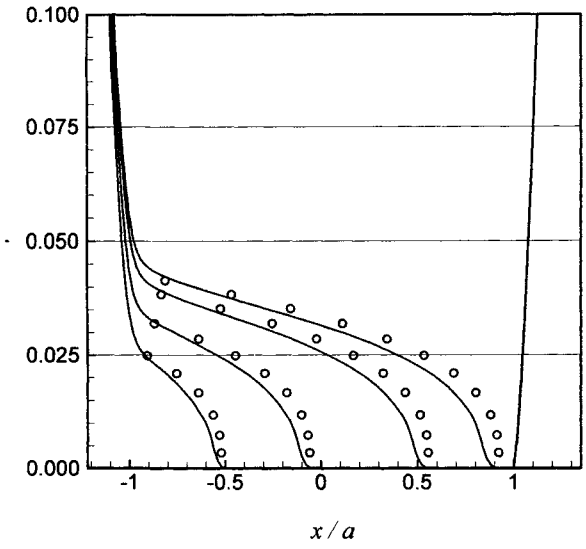


Figure 10. Film profiles obtained during start-up for the 5 m/s^2 case for the simple model (symbols) and full numerical analysis.

contact, which is not influenced by the power adopted in equation (5).

A factor to take into account in considering the differences shown in Figure 10 is that the shape in the inlet region corresponds to the dry contact shape initially as there is no deflection of the surfaces due to pressure outside the Hertzian region. Consequently the shape in the inlet may be less favourable for film generation than that implied in using equation (5) to associate entrainment capability with velocity. This effect may be quantified by an increase in the power adopted in equation (5), which improves the agreement as discussed above.

The differences between numerical and experimental profiles are significant, particularly in the case of the highest start-up acceleration. The mechanism described by the simple model transports lubricant into the conjunction at the current entrainment speed. This can be interpreted as the maximum rate at which mass flow into the contact area can occur through the Couette flow mechanism described by Reynolds' equation. Although the wetted ball surface will move at twice the entrainment speed in this simple sliding case, the quantity of oil entrained forward of the entrainment velocity front will be minuscule as the film thickness is tending to zero there. For the acceleration rate of 50 m/s^2 the distance moved by the entrainment front in 3 ms is 0.225 mm. The experimental profile shown in Figure 3 shows a film at that time that is fully developed and, at 0.16 μm, is 50% higher than that obtained at the final steady state velocity. A comparison of the 3 ms profiles shown in Figure 2 suggests that of the order twice as much oil has been entrained into the experiment contact by this time than is seen in the corresponding calculation.

What therefore has allowed this quantity of oil to enter the contact? It might be speculated that the rig dynamics causes a rapid reduction in load at start of motion. This would reduce the contact area and cause lubricant to be drawn into the gap between the surfaces. Rapid re-imposition of the load would entrap this lubricant and could lead to films thicker than the steady state value. If the load fell to zero, the quantity of oil enclosed by the contacting bodies within the radius a would be equivalent to a uniform film of 0.47 μm over the Hertzian area. This mechanism (of transient load reduction) could therefore quite easily provide means for a film of 0.16 μm to develop. However the simplest explanation is that the kinematic configuration is somehow different to that measured. The disk used in the experiments is itself clamped to ensure it remains at rest [13]. Friction in the dry contact may be sufficient to cause stick between the surfaces initially, which gives way suddenly when sufficient torque is developed. This could result in an initial overspeed of the ball when motion commences and subsequent damped torsional vibration. This mechanism has, however, been considered and discounted by Glovnea and Spikes [1]. For such a speed variation mechanism to explain the variation in h_c shown in Figure 2 for the 50 m/s^2 case, the ball velocity corresponding to the first peak would need to be 0.7 m/s compared to its final steady state value of 0.4 m/s.

Figure 11 shows the calculated variation in contact entrainment velocity with time based on the experimental data for h_c as a function of time [13] and the simple model of equation (5) This equation associates a value of U_e with each of the h_c measurements. A calculation of the times at which U_e had these specific values if the entrainment mechanism takes place according to the simple model of equation (5) has been carried out using a spreadsheet 'solver' facility. Comparison of Figure 11 with Figure 2 indicates the calculated velocity pattern is as seen in the measured film thicknesses values with the main difference being the period of the oscillations observed for the smaller acceleration rates. At the highest acceleration the key changes in

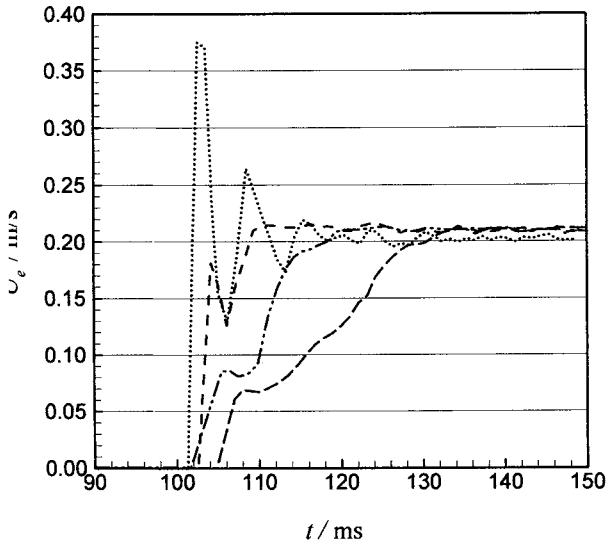


Figure 11. Variation of entrainment velocity, U_e, with time calculated from equation (5) and the experimental values of h_c given in Figure 2 for accelerations of 50 m/s^2, 20 m/s^2, 10 m/s^2 and 5 m/s^2.

the experimental film thickness take place more rapidly than the sampling rate of the film thickness measurements. An experiment that might throw further light on these effects is one where the acceleration takes place, not from rest, but from a very small steady entrainment velocity. This would ensure a full lubricant film existed prior to start up.

6. CONCLUSIONS

A transient numerical analysis has been carried out of the EHL start-up conditions examined experimentally by Glovnea and Spikes [1]. The results show considerable differences when detailed film profile comparisons are made. The transient behaviour of the front between dry and lubricated contact is reproduced, and the numerical results are closely aligned with a simple entrainment transport model. The differences between experiment and theory show an over entrainment of oil into the experimental contact compared to that calculated using the Reynolds equation. A possible explanation for this effect based on an initial sticking of the presumed slipping surfaces is suggested.

ACKNOWLEDGEMENTS

The work reported in this paper was carried out with the support of EPSRC grant GR/N33522.

NOTATION

a	contact semi-dimension	m
E'	reduced elastic modulus	Pa
E_b, E_d	elastic modulii of ball, disk	Pa
$f_{i,j}$	differential deflection influence coefficient	m/N
h	film thickness	m
h_{ent}	film thickness entrained at inlet	m
h_c	film thickness at centre of contact	m
h_0	constant in equation (2)	m
h_{st}	steady state value of h_c	m
i, j, k, l	mesh indices	
u	elastic deflection	m
$\overline{U}$	entrainment velocity	m/s
$\overline{U}_{st}$	steady state entrainment velocity	m/s
p	fluid pressure	Pa
R_x, R_y	radii of relative curvature	m
t	time	s
x, y	Cartesian co-ordinates	m
α	pressure-viscosity coefficient	Pa^{-1}
η	absolute viscosity Pa s	
ρ	density	kg/m^3
τ	shear stress	Pa
σ_x, σ_y	flow factors in x and y directions	ms
τ_0	Eyring shear stress	Pa
ξ	slide roll ratio $\xi = (U_b - U_d)/\overline{U}$	

Moes and Bosma non-dimensional groups

$$L = \alpha E' \left(\frac{2\eta_0 \overline{U}}{E' R} \right)^{1/4} \qquad M = \frac{w}{E' R^2} \left(\frac{E' R}{2\eta_0 \overline{U}} \right)^{3/4}$$

References

1. Glovnea, R. P. and Spikes, H. A. *Elastohydrodynamic film formation at the start-up of the motion.*, Proc. Instn. Mech. Eng., Part J, Journal of Engineering Tribology, Vol 215, pp 125-138, 2001.

2. Holmes, M. J. A. *Transient analysis of the point contact elastohydrodynamic lubrication problem using coupled solution methods.* PhD Thesis, University of Wales, 2002.

3. Holmes, M. J. A., Hughes, T. G., Evans, H. P. and Snidle, R. W. *Transient elastohydrodynamic point contact analysis using a new coupled differential deflection method: Part 1 Theory and validation.* Submitted to Proc. Instn. Mech. Eng., Part J, Journal of Engineering Tribology.

4. Evans, H. P. and Hughes, T. G. *Evaluation of deflection in semi-infinite bodies by a differential*

method, Proc. Instn. Mech. Eng., Part C, Journal of Mechanical Engineering Science, Vol 214, pp 563-584, 2000.

5. Venner, C. H. and Lubrecht, A. A. *Numerical simulation of transverse ridge in a circular EHL contact under rolling/sliding*, ASME Journal of Tribology, Vol 116, pp 751-761, 1994.

6. Holmes, M. J. A., Evans, H. P. and Snidle, R. W. *Comparison of Transient EHL Calculations with shut-down experiments.* Proc 29th Leeds-Lyon Tribology symposium, 2003.

7. Glovnea, R. P. and Spikes, H. A. *Elastohydro-dynamic film collapse during rapid deceleration Part 1: experimental results*, ASME Journal of Tribology, Vol 123, p 254-261, 2001.

8. Johnson, K.L. and Tevaarwerk. L.L. *Shear behaviour of elastohydrodynamic oil films*, Proc. R. Soc. Lond. A 356, p 215-236, 1977.

9. Sharif, K. J., Kong, S., Evans, H. P. and Snidle, R. W. *Contact and Elastohydro-dynamic Analysis of Worm Gears: Part 1 Theoretical Formulation.* Proc. Instn. Mech. Engrs Part C, Jn of Mechanical Engineering Science. Vol 215, pp 817-830; 2001.

10. Tao, J., Hughes, T. G., Evans, H. P. and Snidle, R. W. *Elastohydrodynamic response of transverse ground gear teeth.* To appear Proc. 28th Leeds-Lyon Symp. on Tribology, Elsevier, Amsterdam, 2002.

11. Tao, J., Hughes, T. G., Evans, H. P. and Snidle, R. W. *Elastohydrodynamic Lubrication of Gear Teeth in micropitting tests."* Submitted to Trans ASME Jn. of Tribology, 2002.

12. Chittenden, R.J.,Dowson, D., Dunn, J.F., and Taylor, C.M, *A theoretical analysis of the isotherma;l elastohydrodynamic lubrication of concentrated contacts Parts 1 & 2,* Proc Roy Soc Lond. A397, pp245-294, 1985.

13. Glovnea, R.P. Private communication, 2002.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Comparison of Transient EHL Calculations with Shut-down Experiments

M. J. A. Holmes., H. P. Evans, and R. W. Snidle,
Cardiff School of Engineering, Queen's Buildings, The Parade, Cardiff CF24 3TA, UK, e-mail: evanshp@cf.ac.uk

Keywords: elastohydrodynamic, transient, non-Newtonian, shut-down.

Abstract

The paper presents numerical results for the elastohydrodynamic lubrication of point contacts where the entrainment velocity is reduced with time. The cases considered correspond to experimental work presented by Glovnea and Spikes [1] for the case where entrainment is reduced to zero with constant deceleration.

The experiments modelled involve simple sliding of a steel ball on a glass disk and, due to the poor conductivity of the stationary glass component and the sliding configuration, the temperature is not known with accuracy. The steady state characteristics of the test case, detailed in the experimental result for the beginning of the transient behaviour, are therefore used to select appropriate values of the lubricant parameters for incorporating in a transient numerical analysis of the test conditions.

The paper compares the numerical results obtained using a differential deflection point contact transient analysis with the detailed published experimental film thickness measurements presented by Glovnea et al, and good correlation is shown.

1. INTRODUCTION

Transient analyses of the elastohydrodynamic lubrication (EHL) problem are necessary in attempting to develop an understanding of the lubrication of rough surfaces. The authors have developed a new point contact solution method [2,3] based on a novel differential deflection method [4]. The need to validate a new approach against existing published numerical solutions has been achieved [2,3] using the well known solutions of Venner and Lubrecht [5]. It is also advantageous to validate the solutions by comparison with experimental results, and that is the objective of the current paper and its companion [6].

Glovnea and Spikes [1,7] have recently published detailed results of transient point contact EHL experiments using their ball on transparent plane apparatus. The transient conditions considered consist of the shut-down of EHL action caused by the reduction of entrainment velocity under constant deceleration [1] and the start-up of EHL under a constant acceleration of the contacting components [7]. The shut-down experiments and their comparison with numerical results are the focus of the present paper.

The experimental work [1] consisted of two configurations. In one, the velocity of both contacting components was maintained so as to give pure rolling conditions as the entrainment velocity was reduced linearly to zero. In the second, the disk component remained fixed throughout and the contact operated in simple sliding, i.e. with $\xi = 2$, and the ball velocity was reduced linearly to zero. The kinematics of the second configuration were stated to be controlled more precisely than the first. This was due to difficulties caused by the high inertia of the disk component in comparison with that of the ball. The numerical analyses made in order to aid in validating the differential deflection method in transient mode therefore concentrated on the simple sliding results so as to minimise ambiguities concerning the time history of the component velocities. This introduces the complication of measuring the contact's operating temperature which will be determined by the heat input to the stationary glass disk and is affected by the disk's relatively poor conduction characteristics. This difficulty was overcome by adjusting the assumed operating temperature such that the calculated steady state result corresponded to the film profile information given in the experiment [1]

prior to the beginning of the deceleration. A non-Newtonian rheological model was used and its parameter values chosen so as to obtain the experimental steady state film profile.

2. THEORY

The Reynolds equation for the fluid is

$$\frac{\partial}{\partial x}\left(\sigma_x \frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\sigma_y \frac{\partial p}{\partial y}\right)-\frac{\partial(\rho \bar{U} h)}{\partial x}-\frac{\partial(\rho h)}{\partial t}=0 \qquad (1)$$

where the flow factors σ_x and σ_y ($\neq\sigma_x$) are obtained using the hyperbolic sine model proposed by Johnson and Tevaarwerk [8] as described in [9].

The novel feature of the solution method adopted is that the elastic deflection, u, which gives rise to the elastic film thickness according to

$$h(x,y) = \frac{x^2}{R_x} + \frac{y^2}{R_y} + u(x,y) + h_0 \qquad (2)$$

is specified in differential form [4]. The equation thus specifies the Laplacian of the deflection in terms of the influence of the pressure distribution according to

$$\frac{\partial^2 u(x_i,y_j)}{\partial x^2}+\frac{\partial^2 u(x_i,y_j)}{\partial y^2} = \sum_{all\ k,l} f_{k-i,l-j} p_{k,l} \qquad (3)$$

Equation (2) is then cast in the form

$$\frac{\partial^2 h}{\partial x^2}+\frac{\partial^2 h}{\partial y^2} = \frac{1}{R_x}+\frac{1}{R_y}+\sum_{all\ k,l} f_{k-i,l-j} p_{k,l} \qquad (4)$$

The advantage of the formulation is the extremely rapid decay of the influence coefficients $f_{i,j}$ as the indices i and j increase from zero. This enables equations (1) and (4) to be solved as a coupled pair as described by Holmes et al [2,3]. The corresponding approach for the line contact case has been found to be extremely useful in solving moving roughness problems where the composite roughness of the surfaces is an order of magnitude or more greater than the smooth surface film thickness [10,11].

3. LUBRICANT PARAMETERS

A finite difference Reynolds equation with a central difference formulation of the first order spatial derivatives and a Crank-Nicolson scheme for incorporating time dependence was used [3]. The results were indistinguishable from those obtained with an alternative finite element formulation also described in [3]. The non-Newtonian rheological model discussed above was adopted with the assumption of isothermal behaviour of the oil. The test case used by Glovnea and Spikes [1] is moderately loaded in numerical terms having Moes and Bosma non-dimensional parameter values of $M = 233$ and $L = 5.42$ at their stated conditions. The computing domain used is $-2.5a \le x \le 1.5a$, $-2.0a \le y \le 2.0a$, with 50 nodes in each Hertzian dimension which is sufficient to resolve smooth surface cases under these moderately loaded conditions. In the numerical solution, the smooth steady state solution was first obtained and then used as the starting condition for the transient analysis. The speed of the moving surface was reduced linearly over time using a simple ramp with a series of deceleration rates corresponding to those used in the experimental work [1]. The deceleration rate of the entrainment velocity is then half this value in the simple sliding mode of the experiment. A timestep value of $\Delta t = 0.04$ ms was chosen for all the cases considered. This corresponds to splitting the deceleration ramp into a hundred divisions for the highest deceleration of 75 m/s^2.

The oil parameters stated by Glovnea and Spikes [1] for the experiment were as given in Table 1. Comparison of a steady state EHL solution obtained with these parameters gave a thicker central oil film thickness than that quoted by Glovnea and Spikes [1] for the experiment conducted with the ball partly immersed in lubricant maintained at a temperature of 50 °C. This was partly redressed by adopting the value of $\alpha = 20.5$ GPa^{-1} suggested by Dr Glovnea [12] as being appropriate for the film forming capability of the oil at the experiment's temperature. The remaining discrepancy in central film thickness was removed by adjusting the viscosity to $\eta_0 = 0.42$ Pas, which corresponds to a

Table 1. Experiment parameters as given in [1]

R	9.5 mm
W	20 N
E_{ball}	210 GPa
E_{disk}	75 GPa
ν_{ball}	0.3
ν_{disk}	0.22
η_0 (50 °C)	0.609 Pa s
η_0 (70 °C)	0.245 Pa s
α	24 GPa^{-1}
$\bar{U}$	0.3 m/s
ξ	2

contact temperature of 57.5 °C. The exact contact temperature is not known for the experiment because, although the ball is semi immersed in temperature maintained lubricant, the stationary disk is subject to shear heating at the contact due to the simple sliding configuration. The poor conduction of the stationary glass component leads to an inevitable, unknown, temperature rise. The analysis used is isothermal and the adjustment made to viscosity to match the central film thickness corresponds to an inlet temperature rise of 7.5 °C To be confident that this temperature rise was reasonable, an independent thermal solution was obtained. This smooth, steady state solution was obtained using the thermal model described by Sharif et al [9] and showed an inlet temperature rise of at least 6.5 °C, which is consistent with the adjusted inlet viscosity. This process gave a much closer agreement with the starting condition, but there were still differences in the actual detailed film shape.

Figure 1 shows the steady state film profiles obtained on the entrainment centre line with a series of models where the non-Newtonian model parameter, τ_0, is given a range of different values. This variation in film profile slope with τ_0 is an interesting property of the pure sliding contact and may, incidentally, enable this feature to be used to calibrate non-Newtonian model parameters on a more extensive basis than in this simple study. Comparison of the profiles illustrated in Figure 1 with that obtained experimentally by Glovnea and Spikes led to the value τ_0 = 5 MPa being selected for the numerical modelling work. It may be noted that the value of τ_0 = 30 MPa gives a film profile that is essentially identical to that obtained with a Newtonian model, and this does not correspond to the experimental profile, indicating that including non-Newtonian behaviour is necessary to model the experimental results accurately in the simple sliding mode. The adjustment to viscosity described above was carried out with this value of τ_0 so that the steady state film thickness calculated numerically agreed both in value and slope with the experimental result.

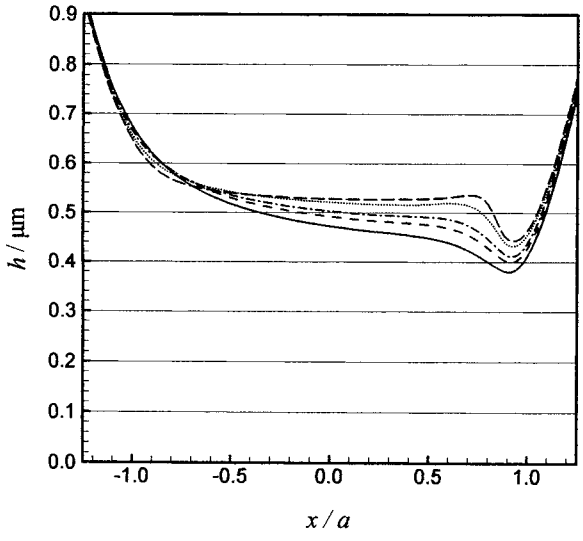


Figure 1. Centre line film thickness profiles for the steady state condition with τ_0 = 3 MPa (lower h_c), 4, 5, 10, and 30 MPa (higher h_c).

4. RESULTS

Detailed film thickness profiles for the entrainment and transverse directions are presented by Glovnea and Spikes [1] for an entrainment deceleration rate of 75 m/s^2. This corresponds to a stopping time of 4 ms for the initial entrainment speed of 0.3 m/s. The measurement technique used can determine the film thickness profile in one particular direction during the experiment, so the two profiles presented correspond to measurements taken from different repeat experiments. The position of the profiles is not stated exactly, but they are taken to be close to the centre of the contact, i.e. on $x = 0$ and $y = 0$, as is implied in the paper [1]. There are small inconsistencies between the experimental results in the two perpendicular profiles. The central film thickness for the transverse profile is initially some 5 % lower (0.465 μm) than that for the entrainment direction profile (0.49 μm), and is subsequently seen to remain at a lower level than the entrainment direction profile value as the film collapses over time. This small discrepancy is probably caused by minor temperature differences between the two repeat experiments and is not of any significance.

The solutions obtained using the coupled transient method are presented in the same format and scale as the experimental results [1] with film profiles plotted at the same time intervals as presented in the paper. Figure 2 shows the experimental and calculated film thickness profiles in the entrainment direction. It can be seen that the main features of the experimental profiles are

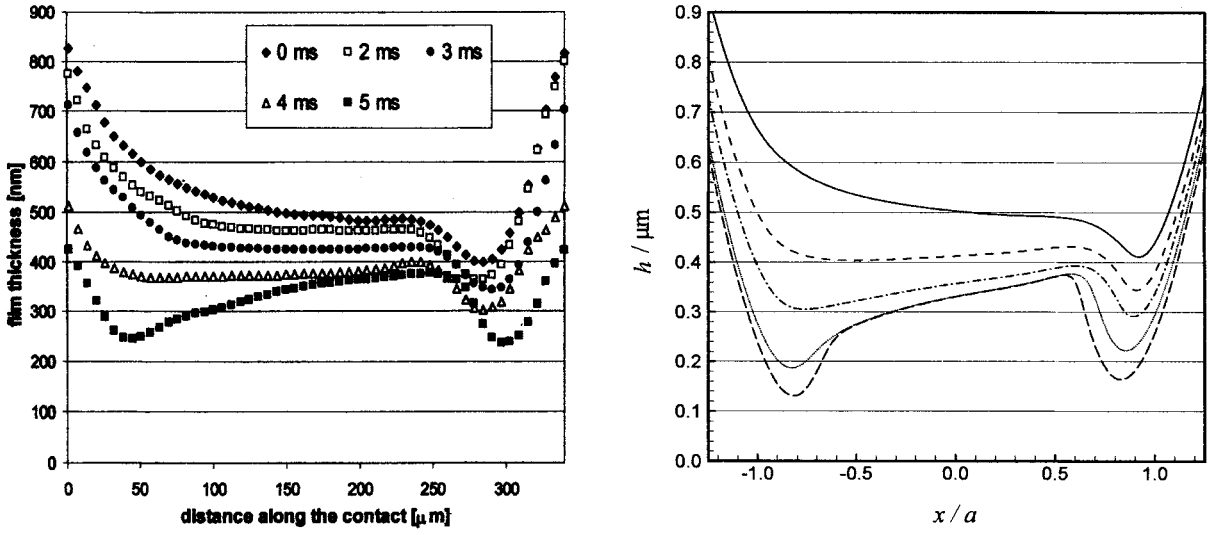


Figure 2. Entrainment centre line film thickness profiles at different times during the deceleration. Left hand figure shows experimental results [1]. Right hand shows calculated numerical results at times of 0 ms (upper curve), 2 ms, 3 ms, 4 ms, and 5 ms (lower curve)

reproduced in the numerical comparisons. The slope of the film thickness over the central part of the Hertz region is progressively diminished and then reversed as a result of entrainment being systematically reduced during the first four milliseconds analysed. After that time the lubricant leaks from the now stationary contact and a central entrapment is formed surrounded by film constrictions.

Figure 3 makes the corresponding comparisons with the transverse film profiles. This shows the formation of a symmetrical entrapment over the first four milliseconds with a relatively slow subsequent reduction in the entrapment film thickness. The main changes in the calculated film after entrainment has ceased at time 4 ms is in the vicinity of the film

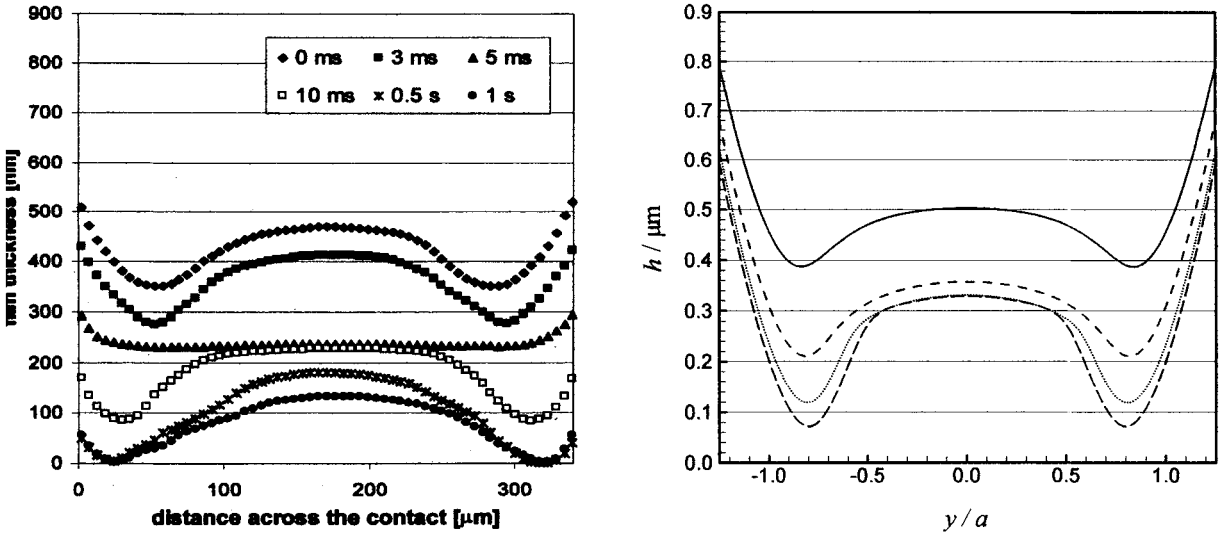


Figure 3. Transverse centre line film thickness profiles at different times during the deceleration. Left hand figure shows experimental results [1]. Right hand shows calculated numerical results

constrictions around the entrapment. The general behaviour of the experimental contact is seen to be reproduced in the numerical results but there are noticeable differences. For example, the feature of the almost constant transverse experimental film thickness of 0.24 μm at time 5 ms is not reproduced in the numerical results. This feature of the experimental measurements is rather unexpected and is indeed inconsistent with the corresponding entrainment centre line profile for the experiment shown in Figure 2, where at time 5 ms the central film thickness is about 0.36 μm. This latter value is more in keeping with the calculated numerical central film thickness value of 0.33 μm. It should be noted, however, that for the experiments reported in [1] for which the film profiles given in Figures 2 and 3 were obtained, "the deceleration rate was not explicitly controlled and the ball was simply halted as rapidly as possible." The constant deceleration rate of 75 m/s^2 is thus an estimate of the kinematic changes taking place, and this may account in some measure for the differences observed between the experiment and the numerical analysis.

Experiments were carried out [1] using a series of controlled deceleration rates from an initial entrainment velocity of 0.3 m/s, and the central film thickness value presented over analysis times of 70 ms. These experiments were simulated numerically in the current work. The numerical solutions were all obtained using the same time-step for each deceleration rate, as previously mentioned. The coupled solutions are presented in a similar format to those given for the experiments [1] with the central film thickness plotted over the duration of the analysis, and are shown in Figure 4. For the case with the least rapid reduction in entrainment velocity the contact does not come to rest within the 70 ms analysed. In each of the other cases the central film thickness falls steadily whilst the entrainment speed is reduced. Once the contact becomes stationary, after 4, 14, 23, 30, 60 ms, respectively for these five deceleration rates, the central film thickness changes very little over the remainder of the 70 ms analysed. At this stage its change is caused by oil leakage from the entrapment, and by adjustments to the shape of the entrapment as this shape becomes increasingly symmetrical.

The rate of central film thickness reduction during shut-down, and its subsequent behaviour during the entrapment stage is seen to be generally well modelled by the numerical method when the results are compared with experiment. The slowest deceleration case of 0.5 m/s^2 requires 600 ms to reach the entrapment stage, and consequently only part of the shut-down process is observed during the 70 ms analysed and recorded in the experiment [1].

The agreement with the higher deceleration rate is the least satisfactory as the experimental central film thickness reduction during shutdown can be seen to occur at a steepening rate in the experimental

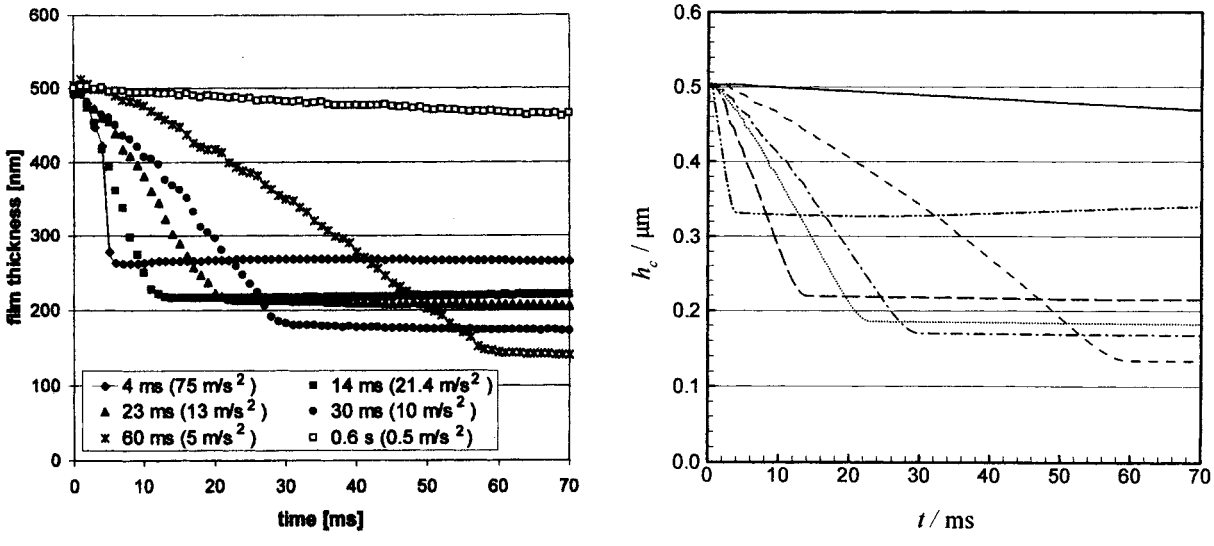


Figure 4. Central film thickness variation for six different deceleration rates. Left hand figure shows experimental results [1]. Right hand figure shows calculated numerical results for deceleration rates of

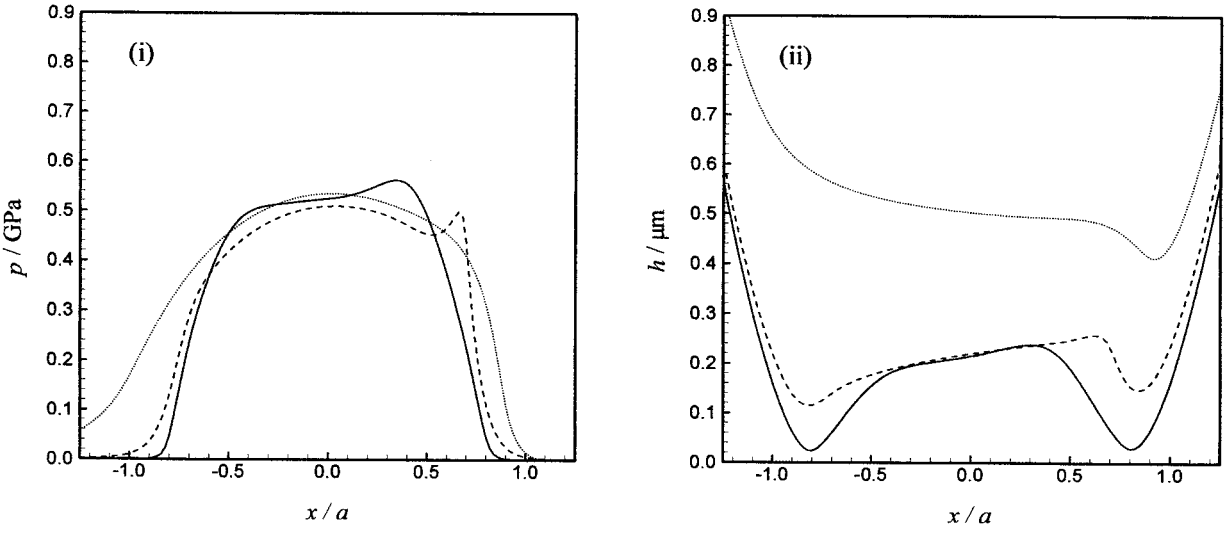


Figure 5. Centre line (i) pressure and (ii) film thickness profiles at start of slow down (dotted), end of rotation (dashed), and after 70 ms (solid line) for the 21.4 m/s^2 deceleration case.

result. The numerical counterpart takes place at a much more uniform rate, and does not reach as low a value as measured in the experiment. The differences observed may well be due to inadequacies in the rheological model adopted. Whether or not these discrepancies can be resolved by utilising a different non-Newtonian model remains to be seen. A further possible factor influencing the experimental results is inertia of the ball leading to wind up in the ball shaft as its drive motor speed is changed in a controlled way. This effect may be present to some extent and could explain the observation that the experimental value of h_c for the 75 m/s^2 deceleration case behaves in a similar way to that for the 21.4 m/s^2 case for the first three measurements taken after the ball velocity has begun to change from its previous steady value. The measurements are taken 1 ms apart and at the entrainment velocity of 0.3 m/s it takes 0.4 ms for lubricant to progress from the inlet of the Hertz contact to its centre. The change in entrainment due to the threefold difference in deceleration rate should therefore be apparent during the time that the measurements remain similar.

Although there is a discrepancy in the entrapment film thickness calculated for the 75 m/s^2 deceleration case when compared with experiment, both numerical and experimental results show a gradual increase of h_c during entrapment. This behaviour is not seen with the lower deceleration rates. Figures 5, 6 and 7 show entrainment centreline pressure and film thickness profiles prior to deceleration, at the instant the ball comes to rest, and 70 ms after deceleration commenced. Figure 5 shows the result for 21.4 m/s^2 where the shut-down time is 14 ms, Figure 6 shows similar results for 75 m/s^2 where the shut-down time is 4 ms, and Figure 7 corresponds to a higher deceleration of 150 m/s^2 where the shut-down time is 2 ms. The gradual increase of h_c during entrapment is seen to also take place in Figure 7. The reason for this difference in entrapment behaviour may be seen by comparing these figures. For the case of Figure 5, entrainment reduction during the slow down phase has reduced h_c to about 0.22 μm during the time taken for the ball to cease to rotate. The constrictions that have formed at the entry and exit sides are at the 0.11 μm and 0.15 μm levels, respectively. Leakage under these constrictions in the subsequent entrapment stage has led to symmetrical constriction clearances by 70 ms, but the central film thickness shape remains inclined and is hardly changed at all during this time. For the case of Figure 6, h_c is about 0.33 μm by the time the ball has ceased to rotate. The entry and exit constriction levels are about 0.18 μm and 0.22 μm so that leakage from the entrapment can take place more quickly. By the 70 ms stage the film thickness has become much more symmetrical than that shown in Figure 5, and a small increase in h_c has occurred. The differences in behaviour between Figures 5 and 6 are further demonstrated in the case of Figure 7. In this example the ball takes only 2 ms to stop and the entrapment h_c is about 0.42 μm with constriction levels of 0.24 and 0.28 μm. By

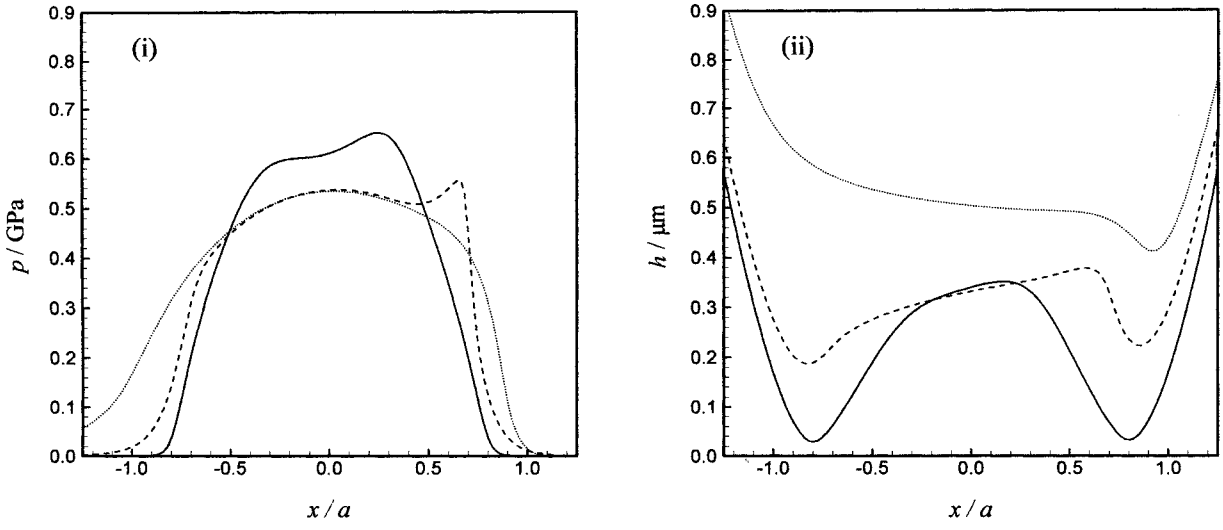


Figure 6. Centre line (i) pressure and (ii) film thickness profiles at start of slow down (dotted), end of rotation (dashed), and after 70 ms (solid line) for the 75 m/s^2 deceleration case.

the 70 ms stage the film is almost completely symmetrical, and a greater increase in h_c has occurred than in the case of Figure 6.

The differences in behaviour seen in Figures 5, 6 and 7 can be explained in terms of the different rates at which oil leaks from the entrapments formed during the time in which the entrainment velocity is reducing to zero. The film thickness shape at the instant that motion ceases is shown by the dashed curves in the figures. The radial oil flow takes place principally at the location of the constriction. At this location the pressure gradient is high but the pressure itself is sufficiently low for the pressure viscosity effect not to cause a large increase in viscosity. This radial flow can be expected to be proportional to the product of the radial pressure gradient and the cube of the film thickness at the constriction. Comparing the film thicknesses at the constrictions formed, and the calculated pressure gradients at these locations, it is clear that leakage from the entrapment will take place most rapidly in the case of Figure 7, and least rapidly in the case of Figure 5. Indeed for these two examples the film thicknesses at the constrictions at the instant that

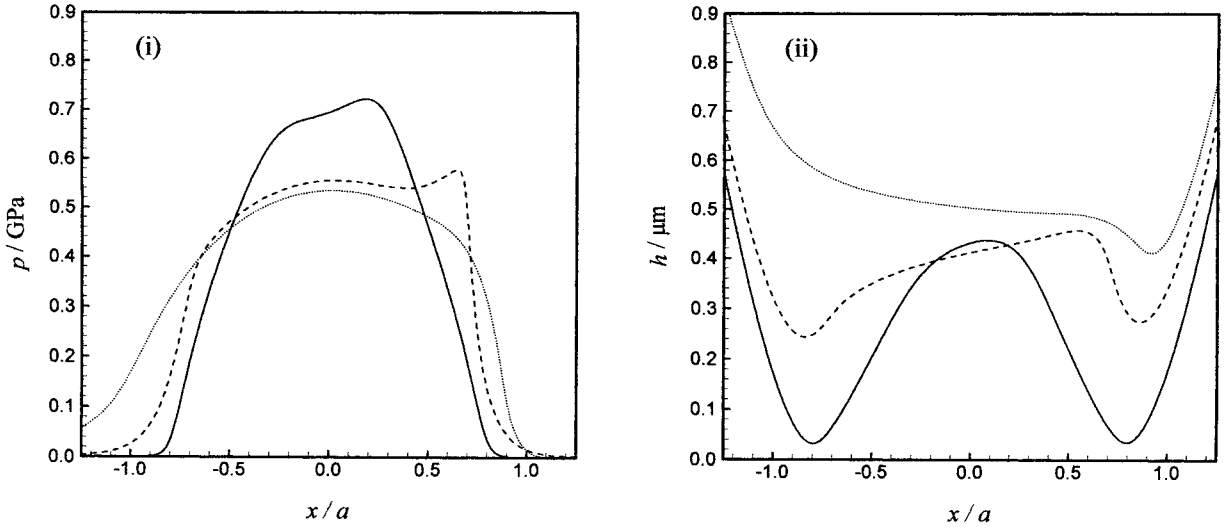


Figure 7. Centre line (i) pressure and (ii) film thickness profiles at start of slow down (dotted), end

rotation ceases are in the ratio 2:1. The pressure behaviour in the vicinity of the constrictions is very similar for the two cases and so the radial flowrates at these locations can be expected to be in the ratio of approximately 8:1. The flow of oil necessary to establish symmetry in the entrapment can thus occur much more quickly for the case of Figure 7, and this effect is apparent in the calculated film shapes after 70 ms.

5. CONCLUSIONS

A transient numerical analysis has been carried out of the EHL shut-down conditions examined experimentally by Glovnea and Spikes [1]. The main features of the experimental measurements are reproduced in the numerical results obtained. There are differences observed in the detailed comparison of film thickness profiles. These may be the result of the particular form of non-Newtonian rheological model adopted for the numerical analysis. If so, such comparisons may be a means of distinguishing between such models within an EHL contact.

6. ACKNOWLEDGEMENTS

The work reported in this paper was carried out with the support of EPSRC grant GR/N33522.

7. NOTATION.

a	contact semi-dimension	m
E'	reduced elastic modulus	Pa
E_b, E_d	elastic of ball, disk	Pa
$f_{i,j}$	differential deflection influence coefficient	m/N
h	film thickness	m
h_0	constant in equation (2)	m
i, j, k, l	mesh indices	
u	elastic deflection	m
$\overline{U}$	Entrainment velocity	m/s
p	fluid pressure	Pa
R_x, R_y	relative radii of curvature	m
t	time	s
x, y	Cartesian co-ordinates	m
α	pressure-viscosity coefficient	Pa^{-1}
η	absolute viscosity Pa s	
ρ	density	kg/m^3
τ	shear stress	Pa
σ_x, σ_y	Flow factors in x and y directions	ms
τ_0	Eyring shear stress	Pa
ξ	slide roll ratio $\xi = (U_b - U_d)/\overline{U}$	

Moes and Bosma non-dimensional groups

$$L = \alpha E' \left(\frac{2\eta_0 \overline{U}}{E' R} \right)^{1/4} \qquad M = \frac{w}{E' R^2} \left(\frac{E' R}{2\eta_0 \overline{U}} \right)^{3/4}$$

REFERENCES

1. Glovnea, R. P. and Spikes, H. A. *Elastohydrodynamic film collapse during rapid deceleration Part 1: experimental results*, ASME Journal of Tribology, Vol 123, p 254-261, 2001.

2. Holmes, M. J. A. *Transient analysis of the point contact elastohydrodynamic lubrication problem using coupled solution methods.* PhD Thesis, University of Wales, 2002.

3. Holmes, M. J. A., Hughes, T. G., Evans, H. P. and Snidle, R. W. *Transient elastohydrodynamic point contact analysis using a new coupled differential deflection method: Part 1 Theory and validation.* Submitted to Proc. Instn. Mech. Eng., Part J, Journal of Engineering Tribology.

4. Evans, H. P. and Hughes, T. G. *Evaluation of deflection in semi-infinite bodies by a differential method*, Proc. Instn. Mech. Eng., Part C, Journal of Mechanical Engineering Science, Vol 214, pp 563-584, 2000.

5. Venner, C. H. and Lubrecht, A. A. *Numerical simulation of transverse ridge in a circular EHL contact under rolling/sliding*, ASME Journal of Tribology, Vol 116, pp 751-761, 1994.

6. Holmes, M. J. A., Evans, H. P. and Snidle, R. W. *Comparison of Transient EHL Calculations with start-up experiments.* Proc 29th Leeds-Lyon Tribology symposium, 2003.

7. Glovnea, R. P. and Spikes, H. A. *Elastohydrodynamic film formation at the start-up of the motion,* Proc. Instn. Mech. Eng., Part J, Journal of Engineering Tribology, Vol 215, pp 125-138, 2001.

8. Johnson, K.L. and Tevaarwerk. L.L. *Shear behaviour of elastohydrodynamic oil films*, Proc. R. Soc. Lond. A 356, p 215-236, 1977.

9. Sharif, K. J., Kong, S., Evans, H. P. and Snidle, R. W. *Contact and Elastohydro-dynamic Analysis of Worm Gears: Part 1 Theoretical Formulation.* Proc. Instn. Mech. Engrs Part C, Jn of Mechanical Engineering Science. Vol 215, pp 817-830; 2001.

10. Tao, J., Hughes, T. G., Evans, H. P. and Snidle, R. W. *Elastohydrodynamic response of transverse ground gear teeth.* To appear Proc. 28th

Leeds-Lyon Symp. on Tribology, Elsevier, Amsterdam, 2002.
11. Tao, J., Hughes, T. G., Evans, H. P. and Snidle, R. W. *Elastohydrodynamic Lubrication of Gear Teeth in micropitting tests.*" To appear Trans ASME Jn. of Tribology, 2003.
12. Glovnea, R. P. *Private communication.* 2001.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Optical Interferometric Observations of the Effect of Moving Transverse Asperities on Point Contact EHL Films

M. Kaneta, N. Tani and H. Nishikawa

Department of Mechanical and Control Engineering, Kyushu Institute of Technology,
Kitakyushu 804-8550, Japan

The transient effects of transversely oriented moving asperities on point contact EHL films are examined by direct observations using the optical interferometry technique. One of the asperities, i.e., the highest asperity, has about twice height compared with other asperities, which are composed of the same height and wavelength. The highest asperity passing through the EHL conjunction remarkably affects the film profile or deformations of asperities downstream of the asperity when the moving speed of the rough surface is smaller than that of the smooth surface. When the moving speed of the rough surface is larger than that of the smooth surface, however, the roughness is hardly influenced by asperities including the highest asperity, although each asperity is deformed by the micro-wedge and squeeze actions. It is also found that the thermal effect plays an important role on such an experiment using a glass and steel contact.

1. INTRODUCTION

The film thickness in the elastohydrodynamic lubrication (EHL) regime is of the same order of magnitude as the surface roughness. Therefore, extensive investigations on rough surface EHL problems have been conducted by numerous researchers from different points of view. Experimental analyses on the effects of surface roughness on the film thickness variations have been carried out using optical interferometry by Jackson and Cameron [1], Wedeven [2], Wedeven and Cusano [3], Cusano and Wedeven [4, 5], Kaneta and Cameron [6], Silva et al. [7], Kaneta and Kanzaki [8], Kaneta, Sakai and Nishikawa [9, 10], Kaneta and Nishikawa [11, 12 13], and Kaneta, Kanada and Nishikawa [14] with single and multiple irregularities produced artificially formed on highly polished steel balls. Venner and Lubrecht [15] and Ai and Cheng [16] have simulated numerically the effects of slide-roll ratio on the pressure and film thickness profile for a transverse asperity and a dent passing through circular EHL contacts, respectively. It has been found through such experimental and numerical conjunction approximately at the average speed of the contacting surfaces while preserving its thickness and shape. This is because the flow in the EHL conjunction is mainly dominated by shear flow and pressure-induced flow is almost absent due to the very high viscosity [17]. Consequently, when the velocity of a smooth surface is larger than that of a rough surface, the surface roughness produces a strong downstream effect, while in the reverse situation the irregularity produces a strong upstream effect.

The authors [10] have proved using artificially-produced multiple transverse asperities with a constant height and wavelength that the above interesting phenomena occur when the speed of the smooth surface is larger than that of the rough surface. However, the authors could not observe the upstream effect in the reverse situation. Ai and Cheng [18] have shown through their transient EHL analyses for line contacts with measured surface roughness that when the moving speed of surface roughness equals to or exceeds the entraining speed, the roughness of the deformed surface profile in the EHL conjunction is close to the roughness of the undeformed roughness profile, although the

rough surface is slower than that of the smooth surface. However, the authors [10] have found in a point contact EHL that the vertical deformation of asperities always occurs when the film factor defined by the ratio of central film thickness calculated for smooth surfaces to root-mean-square roughness is less than 3, although the degree is rather small under pure rolling conditions. Therefore, more detail observations of transient phenomena induced by surface roughness would be necessary. In this paper, direct observations of the transient effect of oil films induced by the moving asperities having different heights will be described using the optical interferometry technique.

1.1. Nomenclature

a = Hertzian contact radius
E'= reduced elastic modulus, $2/E'=(1-\nu_B^2)/E_B+(1-\nu_D^2)/E_D$
G = material parameter, $\alpha E'$
h = film thickness
R = radius of steel ball
u_B, u_D = surface velocities of steel ball and glass disk
U = speed parameter, $\eta_0(u_B+u_D)/2E'R$
w = applied load
W = load parameter, $w/E'R^2$
x = coordinate along the entrainment
X = dimensionless coordinate along the entrainment, x/a
α = viscosity-pressure coefficient
ν_B, ν_D = Poisson's ratios of steel ball and glass disk
η_o = viscosity at atmospheric pressure
Σ = slide-roll ratio, $2(u_D-u_B)/(u_D+u_B)$

2. EXPERIMENTAL PROCEDURE

2.1. Apparatus and Specimens

The experimental apparatus is almost the same as that used previously [8-14]. The contacting bodies are a precision steel ball of 25.4 mm diameter and an optical glass disk of 165 mm diameter and 12 mm thickness, whose lower surface is sputtered with a semi-reflecting chromium layer. The reduced elastic modulus, E', is equal to 117 GPa. The steel ball was mounted between four taper rollers. The ball and the disk were driven separately so that the slide-roll ratio, Σ, could be adjusted to any desired value.

The duochromatic optical interferometry technique [19] was used to measure the shape and thickness of the lubricating film. A Xenon flash lamp with a flash duration of 20 μs was used as a light source and the fringe pattern was recorded with a high speed VCR (500 frames per second) and a 35 mm camera attached to a microscope.

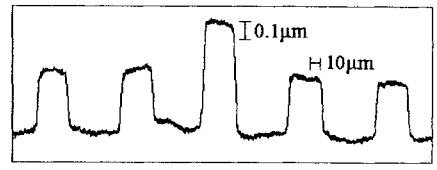


Fig.1 Stylus trace of asperities

Regular multiple long asperities, one of those has about twice as large in the height as the others, were produced onto one part of the ball surface by two stages of sputter deposition of chromium through a foil mask pasted on the surface of the steel ball: As the first step, regular multiple long asperities with a certain height and a constant wavelength were produced using a mask having nine slits, which have a length of 1 mm, a width of 30 μm, and a wavelength of 60 μm. After that, i.e., as the second step, the height of one of the regular multiple long asperities was increased using the same mask, where eight slits but for the central slit of the mask were covered with an adhesive tape. Figure 1 shows the stylus trace of part of the multiple asperities used in this experiment. The maximum height of asperities was about 0.58 μm and the heights of other asperities were about 0.3μm. The generatrix of each long asperity was arranged so as to be perpendicular to the oil entraining vector.

In photographing with the 35 mm camera, the flush of the Xenon lamp was triggered by a signal from the rotary encoder attached to the end of the shaft to which the steel ball was fastened. The exposure time was fixed by a single discharge from the Xenon lamp, which was short enough to freeze movement. Consequently, the transient behaviour of EHL films influenced by the transverse asperities passing through the contact region was recorded at a space of 55.4 μm; the pictures recorded were not continuous with time. The comparison between a series of interferograms recorded with the high speed VCR and that with the 35 mm camera proved that the transient effect induced by the moving asperities can be evaluated using interferograms recorded with the 35 mm camera. In this paper, therefore, the results obtained with the 35 mm camera were used, because of the good resolution.

2.2. Experimental Conditions

The oil used was a mineral bright stock (BS) having

a kinematic viscosity of 1388 mm^2/s at 21°C and 376 mm^2/s at 40°C, a specific gravity of 875 kg/m^3 at 21°C and 863 kg/m^3 at 40°C and a pressure-viscosity coefficient, α, of 23.7 GPa^{-1} at 21°C and 19 GPa^{-1} at 40°C.

The experiments were performed by changing slide-to-roll ratio, Σ, under a constant dimensionless entrainment or rolling velocity U. A load of w=39.2 N, which gives a maximum Hertzian pressure of 0.54 GPa and a Hertzian diameter of 0.37 mm, was used. The oil temperature in the inlet to the contact was 20.5± 1.0°C. These conditions correspond to the dimensionless load parameter $W=2.1\times10^{-6}$ and the materials parameter G=2770.

3. RESULTS AND DISCUSSION

Representative interferograms observed in cases where the highest asperity is located near the center of EHL conjunction are shown for three slide-roll ratios Σ=1, 0 and -1 under several values of dimensionless speed parameters in Figs. 2, 3 and 4, respectively, together with the corresponding midplane film profiles along the entraining direction. The red curves show the film profiles obtained for smooth surfaces. The direction of oil entraining velocity is from the left hand side to the right. An arrow shown in each film profile shows the position of the highest asperity. Figures 5, 6 and 7 show interferograms and the corresponding midplane film profiles along the entraining direction observed at various positions when the highest asperity passes through the EHL conjunction under conditions of Σ=1 (u_B=11.9 mm/s, u_D=35.4 mm/s), Σ=0 (u_B=23.8 mm/s, u_D=23.6 mm/s) and Σ=-1 (u_B=38.6 mm/s, u_D=12.8 mm/s), respectively, with $U=2\times10^{-11}$.

As has already been pointed out [6, 9, 10, 15, 20, 21], when the steel ball having the asperities rotates, each of the asperities entering the contact area receives the squeeze effect and consequently, the leading side (right hand side) of each asperity forms a divergent gap (reverse horseshoe shaped constriction). This effect will be defined as the micro-squeeze action. Furthermore, the profile of each asperity is markedly influenced by the slide-roll ratio Σ. In $\Sigma\neq0$, there is relative motion between each asperity and the mating surface, resulting in the change of the flow rate at each asperity. Consequently, when Σ>0 where the velocity of the glass disk (smooth surface) is larger than that of the steel ball (rough surface), the trailing side (left hand side) of each asperity forms a convergent gap by the hydrodynamic wedge action. When Σ<0, however, the leading side of each asperity forms a divergent gap. Such an effect will be defined as the micro-wedge action. In the present system of the contacting surfaces, the micro-squeeze action seems to become significant as the value of Σ shifts from positive to negative, and the micro-wedge action with an increase in $|\Sigma|$.

Within the limits of present experiments, direct contact of both surfaces could not be observed, and the minimum film thickness took place in the places where the highest asperity overlapped with side-lobes of the macroscopic horseshoe-shaped constriction. When the entrainment velocity or film thickness is small enough, within the range of resolution of interference fringe, the minimum film thickness at the tip of the highest asperity was almost the same as those at the tips of other asperities, regardless of the slide-roll ratio. However, when the overall film thickness becomes large, the minimum film thickness always occurs at the tip of the highest asperity. Furthermore, the film profile roughly approaches an envelope curve of the troughs of the asperities.

3.1. Effects of Slide-Roll Ratio

In cases where Σ=0, when the entraining velocity is low and the overall film thickness is rather thin, each asperity is depressed vertically, and forms a divergent shape due to the micro-squeeze action, but the original shape is roughly kept as a whole; the configuration is frozen during the passage through the EHL conjunction.

As seen from the comparisons among Figs. 2, 3 and 4, and also pointed out in the previous paper [10], there is a big difference in film profiles or deformation of each asperity between Σ=0 and $\Sigma\neq0$. At rather low entraining velocities, the vertical deformation of each asperity is remarkable and the asperity hardly keeps the original shape for both cases of Σ=-1 and Σ=1. As the entraining velocity increases, the deformation of each asperity decreases, though the divergent and convergent film shapes corresponding to Σ=-1 and Σ=1, respectively, are maintained especially at the highest asperity.

It should be noted from Fig. 3 corresponding to the case of Σ=1 at rather small values of U that the film thickness at the bottom of a trough just before the exit constriction almost equals to that at the tip of the adjacent asperity. In particular, the film thickness at that part for the case shown in Fig. 3(a) is thinner than

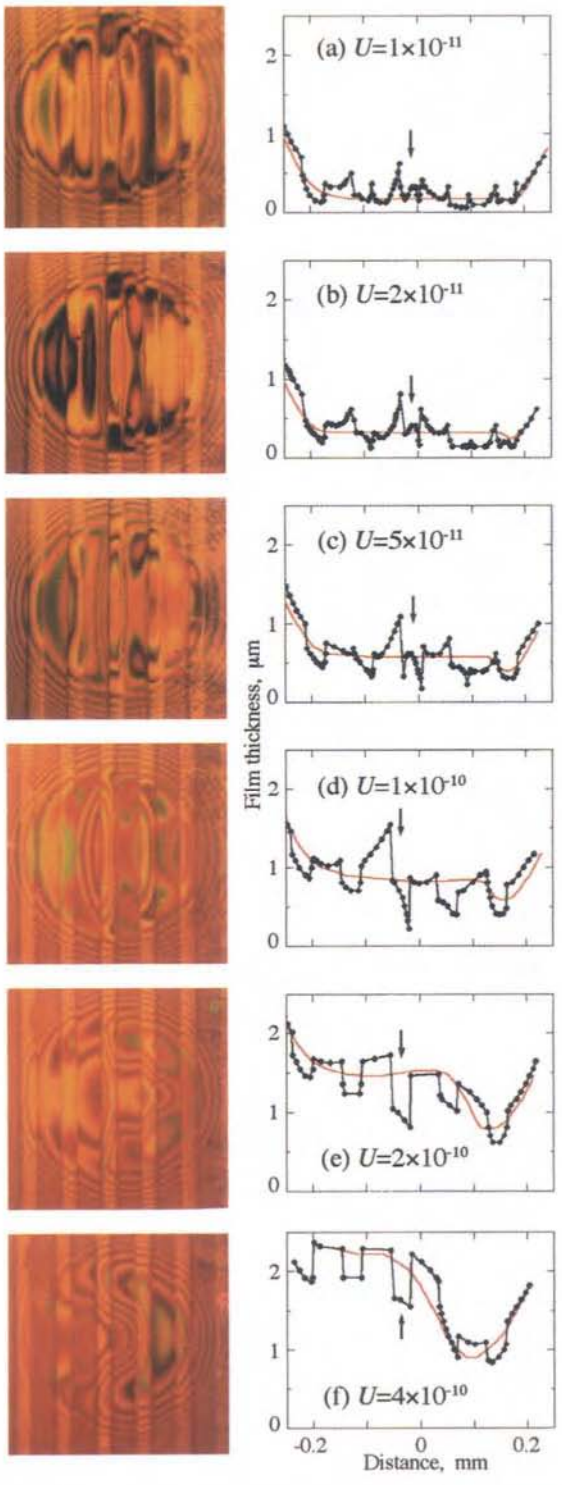


Fig.2 Interferograms and midplane film profiles where the highest asperity locates near the center of EHL conjunction; Σ=1

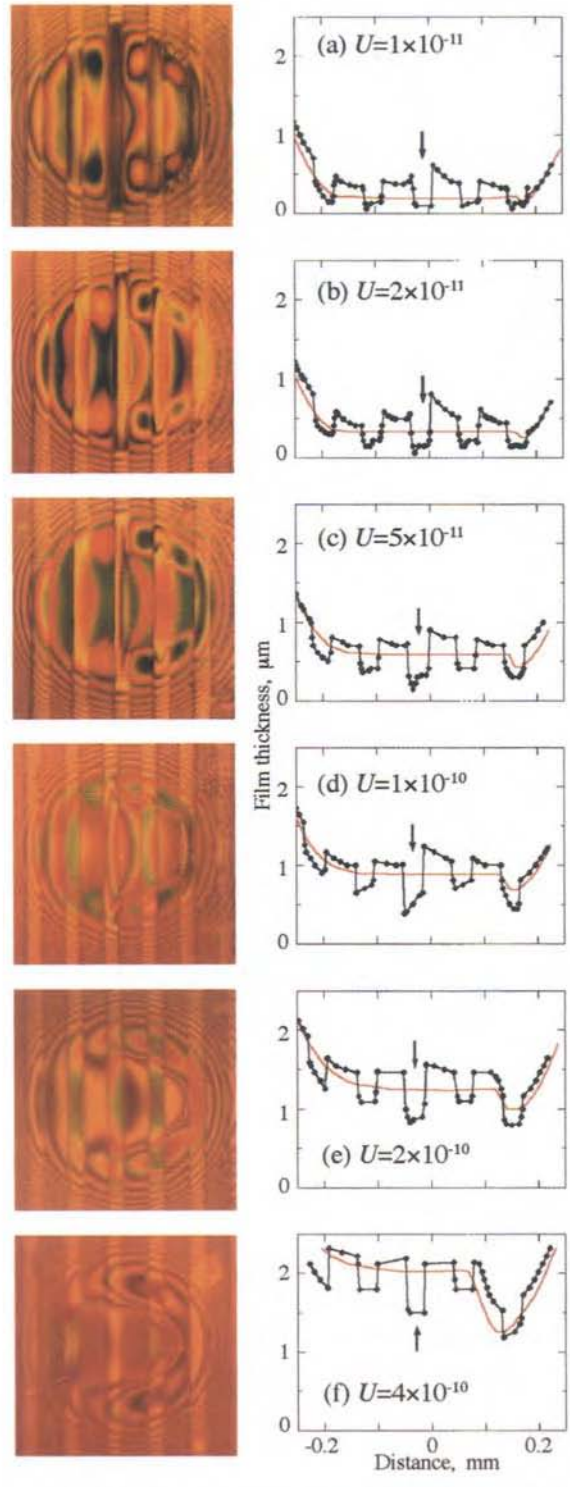


Fig.3 Interferograms and midplane film profiles where the highest asperity locates near the center of EHL conjunction; Σ=0

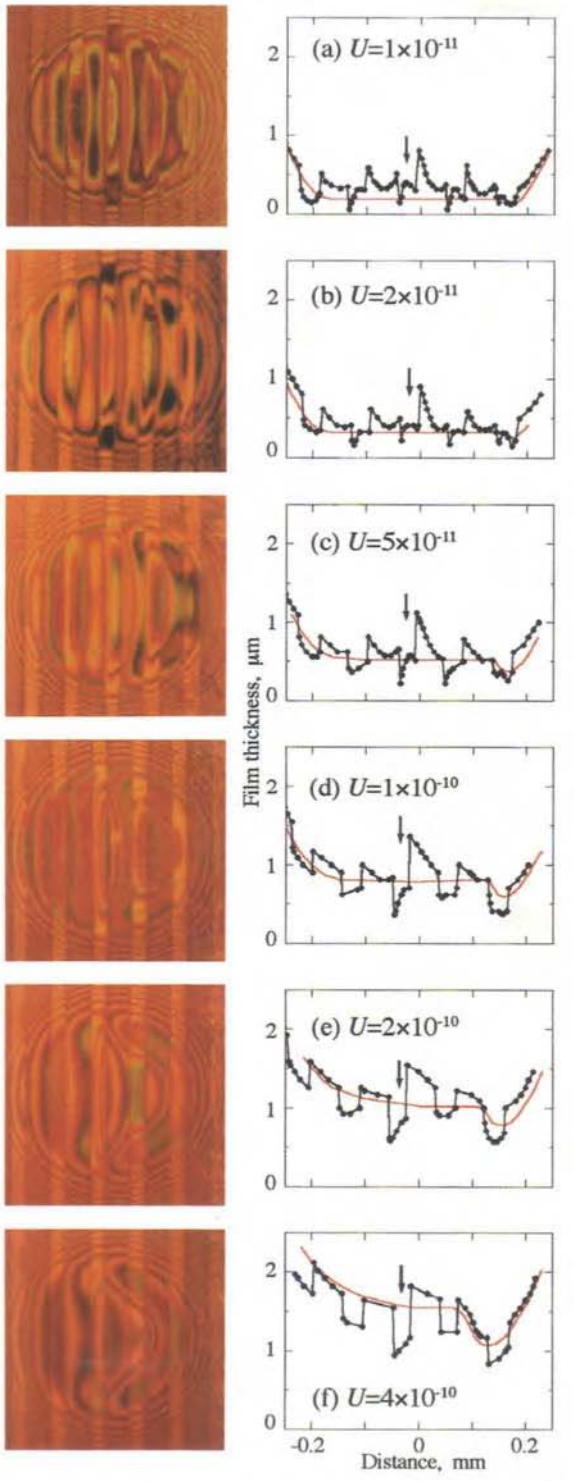


Fig.4 Interferograms and midplane film profiles where the highest asperity locates near the center of EHL conjunction; Σ= - 1

that at the highest asperity. Such a phenomenon seems to be greatly dependent on the fact that the Couette flow dominates over the Poiseuille flow in the EHL conjunction and consequently, a local film thickness moves at the average surface velocity. The detail will be explained using Fig. 5.

As seen from Fig. 5, the film thickness profile fluctuates complicatedly as each asperity passes through the EHL conjunction compared with the case of Σ=0. However, the detail observations prove that there is a remarkable regularity in the film thickness fluctuation. As the first example, let us consider the film thickness at the leading edge of the highest asperity. The film thickness at the position, which is denoted by O, becomes thicker in the order of Figs. 5(b), 5(c) and 5(d). Since the flow in the EHL conjunction is mainly dominated by the Couette flow, the oil moves with twice the velocity of each asperity when Σ=1. As a result, when the position shown by O in Fig. 5(b) moves to the position shown by O in Fig. 5(c), the film thickness at the position denoted by O in Fig. 5(c) becomes thick compared with that at the position denoted by O in Fig. 5(b), since its film thickness is influenced by rather thick film thickness at the position denoted by Δ in Fig. 5(b). In the same way, the film thickness at the position denoted by O in Fig. 5(d) becomes a little thicker than that in Fig. 5(c) by the effect of rather thick film thickness at the position denoted by Δ in Fig. 5(c).

As the second example, let us consider the change of the film thickness over one wave length, which corresponds to the region "A" shown in Fig. 5, composing of an adjacent asperity and trough. The film thickness at the position ● in Fig. 5(b) becomes thin as time passes or the steel ball rotates. That is, the position ● in Fig. 5(b) moves to the position ● in Fig. 5(c) and then to the position ● in Fig. 5(d). According to the effect of rather thin film thickness at the position ✖ in Fig. 5(b) and at the position ✖ in Fig. 5(c), the film thickness at the position denoted by the symbol ● becomes thinner in the order of Figs. 5(b), 5(c) and 5(d). The film thickness at the position ✖ in Fig. 5(c) is influenced by that at the position O in Fig. 5(b). That is, rather thin film thickness over one wavelength "A" shown in Fig. 5(d) is caused by the effect of the highest asperity. Therefore, it can be concluded that a high asperity passing through the EHL conjunction affects markedly local film profiles downstream of the highest asperity when Σ>0.

On the other hand, in the case where Σ=-1, as seen

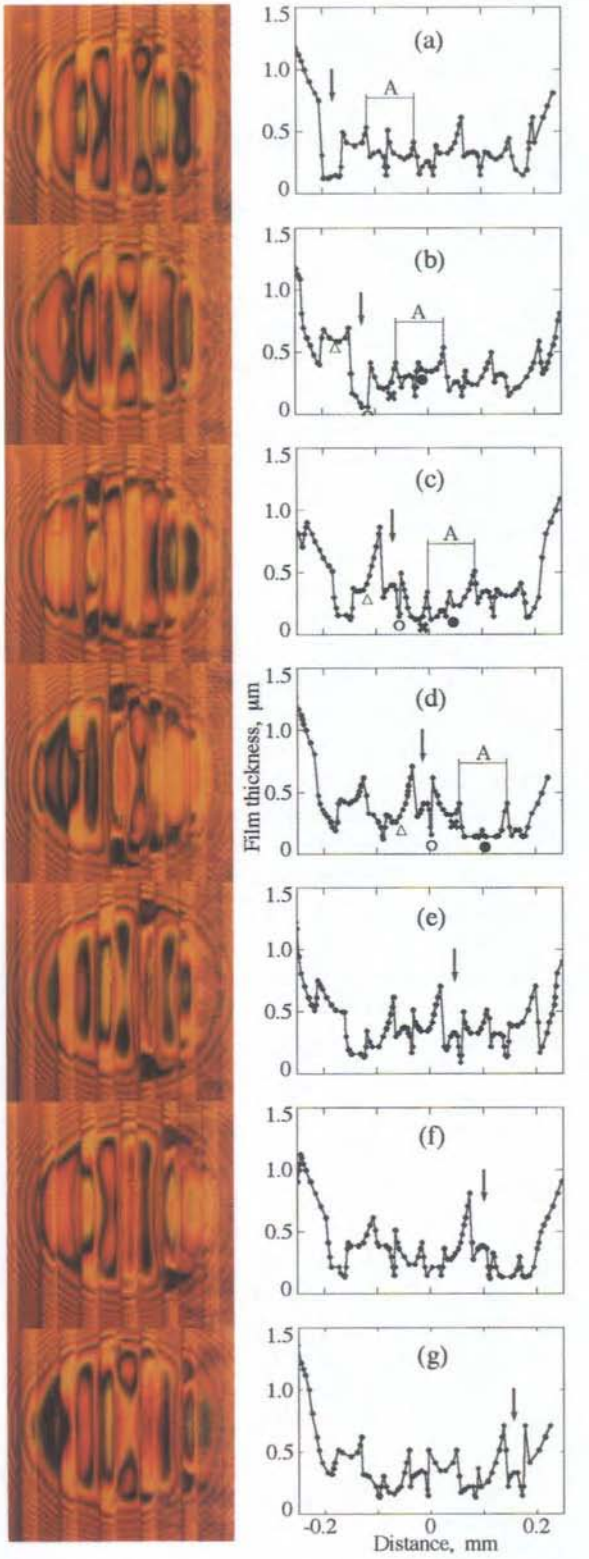


Fig.5 Interferograms and midplane film profiles of the asperity passing through the contact; Σ=1, U=2x10^{-11}

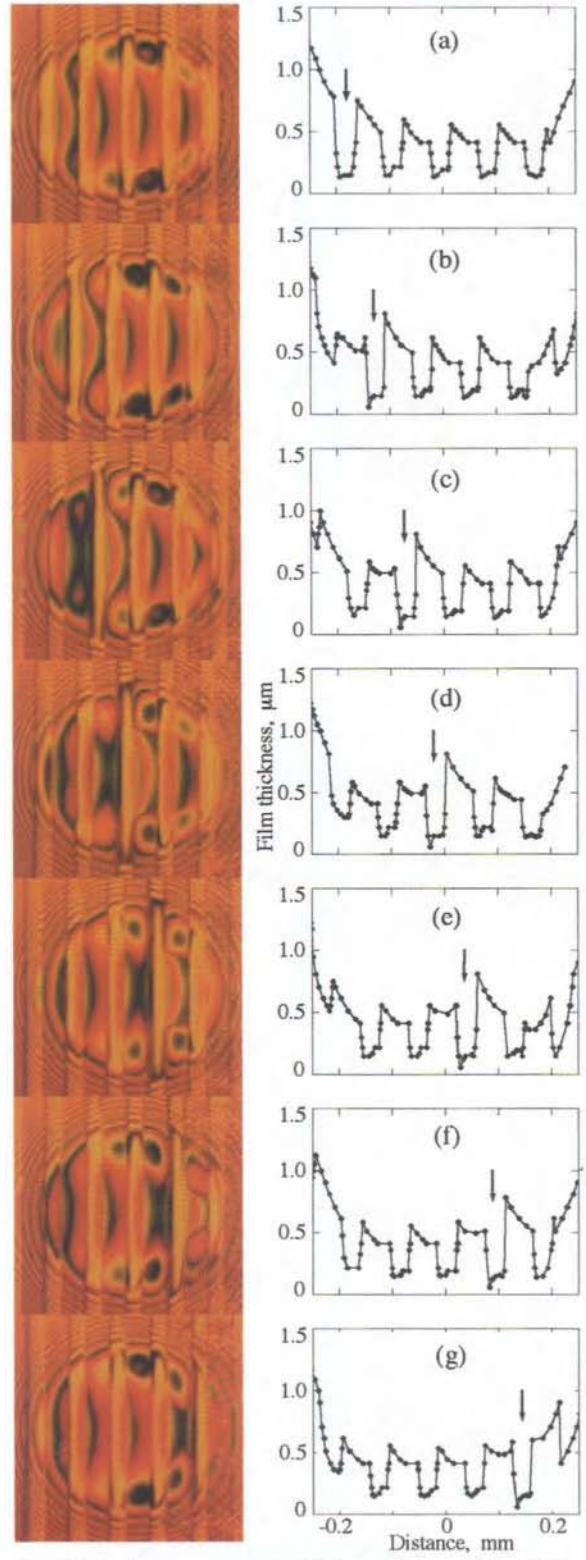


Fig.6 Interferograms and midplane film profiles of the asperity passing through the contact; Σ=0, U=2x10^{-11}

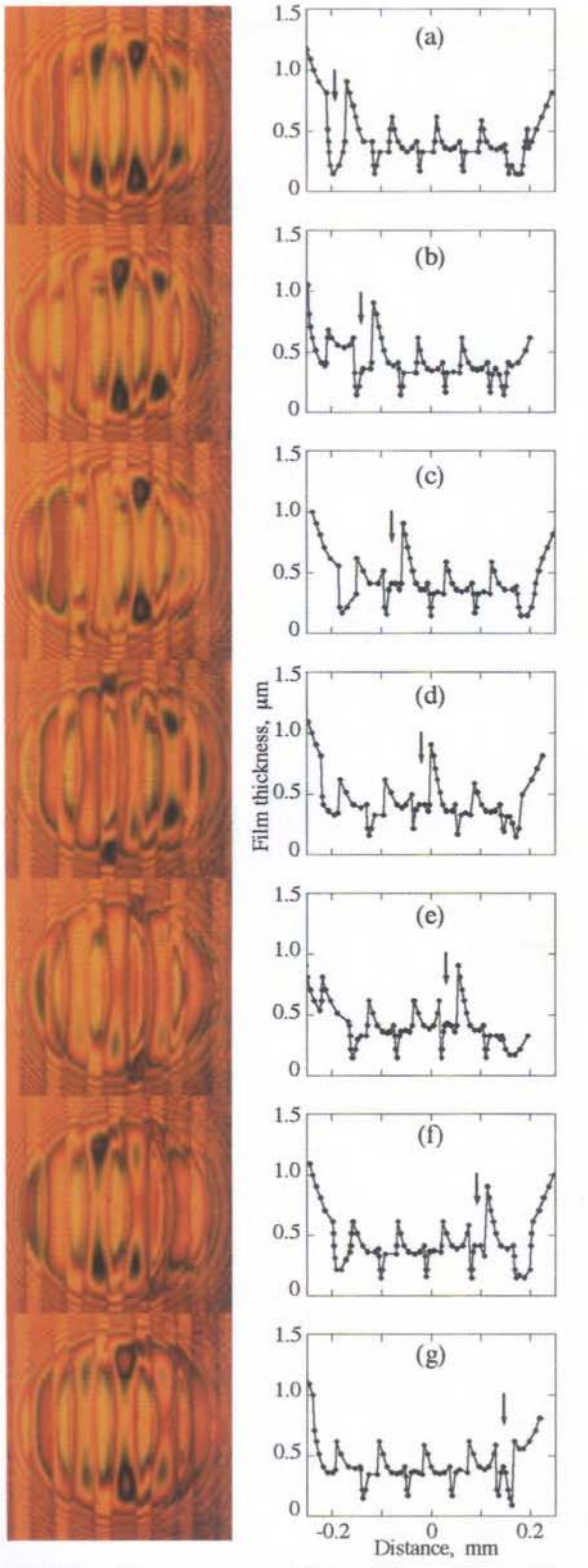


Fig.7 Interferograms and midplane film profiles of the asperity passing through the contact; Σ= -1, U=2x10^{-11}

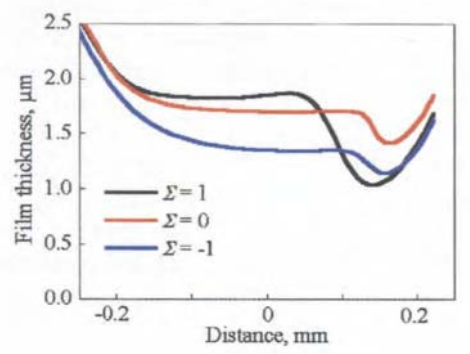


Fig.8 Midplane film profiles obtained numerically under conditons corresponding to Figs. 2(f), 3(f) and 4(f).

from Fig. 7, the highest asperity does not affect the film profile or deformations of asperities upstream of its asperity. Furthermore, the film thickness at the tip of each asperity in the EHL conjunction is almost the same. Similar phenomena were obtained even for the case of Σ=-1.5. These facts differ very much from the results obtained with single dent or ridge [2-5, 9-16]; with single dent or ridge, the strong upstream effect was produced under conditions of Σ<0.

It is an interesting fact that as seen from Figs. 4(a) and 4(b) the film thickness at the trailing edge of the highest asperity is slightly thicker than that at asperities existing in the both sides, when the entrainment speed is rather small. This seems to be due to the strong squeeze and wedge effects that the highest asperity receives.

3.2. Effect of Contacting Materials

Comparison of Figs. 2, 3 and 4 shows that as the velocity parameter increases the difference occurs in overall film thickness and shape, i.e., film profile for smooth surfaces, among Σ=-1, 0 and 1, although it is commonly accepted that film thickness in EHL contacts depends mainly on the conditions at the entrance of the contact and is independent of the slide/roll ratio if the velocity parameter is the same. As clearly seen from Figs. 2(e), 3(e) and 4(e), and Figs. 2(f), 3(f) and 4(f) there is a big difference in the film shape, and the minimum film thickness becomes thinner in the order Σ=0, -1, and 1. It can be considered that such a difference is caused by the temperature-viscosity wedge action [21] due to temperature distributions within the oil film, which is produced by a large difference in thermal conductivity of the

contacting materials. In order to make clear this fact, numerical simulations were carried out for smooth surfaces based on the Newtonian thermal EHL analysis in consideration of the variation of oil properties in all directions within the film. Figure 8 shows the mid-plane film profiles along the oil entraining direction obtained under conditions corresponding to Figs. 2(f), 3(f), and 4(f). The experimental and numerical results show almost the same tendency. The detail explanation will be reported in near future.

4. CONCLUSIONS

The effects of transversely oriented asperities passing through the EHL conjunction on micro-EHL films have been examined by direct observations using the optical interferometry technique. The main results obtained are as follows:

1. When the rough surface moves at a slower velocity than the smooth mating surface, each asperity affects local film profiles or deformations of asperities downstream of the asperity; in particular the effect of the highest asperity is remarkable.
2. However, when the moving speed of the rough surface is larger than that of the smooth surface, local films are hardly influenced by asperities including the highest asperity, although each asperity deforms by the micro-wedge and squeeze actions.
3. In order to examine the effect of slide-roll ratio on EHL films in detail under rather high sliding velocities using such a glass and steel contact system where the contacting surfaces are composed of materials having a large difference in the thermal conductivity, it is very important to consider the thermal effect.

ACKNOWLEDGMENTS

The authors wish to express their sincere thanks to Miss M. Sakamoto and Mr. T. Shigeta of Kyushu Institute of Technology for their help in this work. The program for numerical computation was offered by Professor P. Yang of Qingdao Institute of Architecture and Engineering. The properties of oil used were provided by Professor N. Ohno of Saga University, and the steel balls were supplied by Nippon Seiko K.K. This work was supported by Grant-in-Aid for Scientific Research (No. 13450067) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and the Satellite Venture Business Laboratory of Kyushu Institute of Technology.

REFERENCES

1. Jackson, A. and Cameron, A., 1976, "An Interferometric Study of the EHL of Rough Surfaces," ASLE Trans., Vol.19, pp.50-60.
2. Wedeven, L.D., 1978, "Influence of Debris Dent on EHD Lubrication," ASLE Trans., Vol.21, pp.41-52.
3. Wedeven, L.D. and Cusano, C., 1979, "Elastohydrodynamic Film Thickness Measurements of Artificially Produced Surface Dents and Grooves," ASLE Trans., Vol.22, pp.369-381.
4. Cusano, C. and Wedeven, L.D., 1981, "Elastohydrodynamic Film Thickness Measurements of Artificially Produced Nonsmooth Surfaces," ASLE Trans., Vol.24, pp.1-14.
5. Cusano, C. and Wedeven, L.D., 1982, "The effects of Artificially-produced Defects on the Film Thickness Distribution in Sliding EHD Point Contacts," ASME Journal of Lubrication Technology, Vol.104, pp.365-375.
6. Kaneta, M. and Cameron, A., 1980, "Effects of Asperities in Elastohydrodynamic Lubrication," ASME Journal of Lubrication Technology, Vol. 102, pp.374-379.
7. Silva ,G.M.S., Leather, J. A. and Sayles, R.S., 1985, "The Influence of Surface Topography on the Lubricant Film Thickness in an Elastohydrodynamic (EHD) Point Contact," Proc. 12th Leeds-Lyon Symp., pp.258-272.
8. Kaneta, M. and Kanzaki, Y., 1991, "Effects of Surface Roughness on Squeeze Film Characteristics in EHL," Tribologist, Vol.36, pp.319-326 (in Japanese).
9. Kaneta, M., Sakai, T. and Nishikawa, H., 1992, "Optical Interferometric Observations of the Effects of a Bump on Point Contact EHL," ASME Journal of Tribology, Vol.114, pp.779-784.
10. Kaneta, M., Sakai, T. and Nishikawa, H., 1993, "Effects of Surface Roughness on Point Contact EHL," Tribology Trans., Vol.36, No.4, pp.605-612.
11. Kaneta, M. and Nishikawa, H., 1994, "Local Reduction in Thickness of Point Contact EHL Films Caused by a Transversely Oriented Moving Groove and Its Recovery," ASME Journal of

Tribology, Vol.116, pp.635-639
12. Kaneta, M. and Nishikawa, H., 1999, "Experimental Study on Microelastohydrodynamic Lubrication," Proc. I. Mech. E., Vol.213, Part J, pp.371-381.
13. Kaneta, M. and Nishikawa, H., 1999, "The Effects of a Transversely Oriented Bump on Point Contact EHL Films in Reciprocating Motion with a Short Length of Stroke," Lubrication at the Frontier, Eds. Dowson et al., Elsevier, pp. 185-192.
14. Kaneta, M., Kanada, T. and Nishikawa, H., 1997, "Optical Interferometric Observations of the Effects of a Moving Dent on Point Contact EHL," Elastohydrodynamics '96, Eds. Dowson et al., Elsevier, pp. 69-79.
15. Venner, C.H. and Lubrecht, A.A., "Numerical Simulation of a Transverse Asperity in a Circular EHL Contact Under Rolling/Sliding," ASME Journal of Tribology, Vol.116, pp.751-761.
16. Ai, X. and Cheng, H.S., 1994, "The influence of Moving Dent on Point EHL Contacts," Tribology Trans., Vol. 37, No. 2, pp.323-335.
17. Venner, C.H., 1991, "Multilevel Solution of the EHL Line and Contact Problems," Ph.D. thesis, Twente University, The Netherlands.
18. Ai, X. and Cheng, H.S., 1994, "A Transient EHL Analysis for Line Contacts With Measured Surface Roughness Using Multigrid," ASME Journal of Tribology, Vol.116, pp.549-558.
19. Foord, C.A., Wedeven, L.D., Westlake, F.J., and Cameron, A., 1969/1970, "Optical Elastohydrodynamics," Proc. Instrn. Mech. Engrs., Vol.184, Part 1, No.28, pp.487-505.
20. Cheng, H.S., 1977, "On Some Aspect of Microelastohydrodynamic Lubrication," Proc. 4th Leeds-Lyon Sympsium on Tribology, pp.77-79.
21. Chang, L., Cusano, C., and Conry, T.F., 1989, "Effects of Lubricant Rheology and Kinematic Conditions on Micro-Elastohydrodynamic Lubrication," ASME Journal of Tribology, Vol.111, pp.344-351.
22. Yang, P., Qu, S., Kaneta, M., and Nishikawa, H., 2001, "Formation of Steady Dimples in Point TEHL Contacts," ASME J. Tribol., 123, pp. 42-49.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Mechanism of Dimple Formation under Elastohydrodynamic Conditions

K. Yagi*, K. Kyogoku and T. Nakahara

Graduate School of Science and Engineering, Tokyo Institute of Technology
12-1, Ookayama, 2 chome, Meguro-ku, Tokyo, 152-8552, Japan

This paper describes an experimental study to investigate the mechanism of dimple formation under elastohydrodynamic conditions. The authors measured the temperatures of both the sliding surfaces and the oil film, furthermore, estimated the temperature profiles across the oil film with experimental results. The experimental results show that a temperature difference between the surfaces exists at the dimple zone and varies in the sliding direction. The temperatures of the oil film and the surfaces increase dramatically at the dimple zone. Furthermore, the couette flow predicted by the temperature difference between the surfaces decreases as is close to the centre of the dimple i.e. the pressure increases by the viscosity wedge action at the dimple zone.

1. INTRODUCTION

It has been believed that the film shape in the elastohydrodynamic (EHD) conditions would be characterized by the parallel film at the centre of the contact area and the minimum film at the exit zone and the film thickness would become zero as the entrainment speed goes zero until recently. Dowson et al [1] presented the first exact solution for line EHD contacts by combination of the Reynolds equation and the elastic deformation equation of the bounding surfaces. Their numerical solutions indicated that the film profile was formed by the parallel film at the centre zone and the minimum film constriction at zone just before the exit zone and later such a profile was indicated experimentally by Crook [2], who measured the EHD film profile with a two disk apparatus by the electrical capacitance method. Gohar et al [3] showed EHD film shape in point contacts by the optical interferometry method. The film shape in the point contacts was also characterized by the parallel film and the horse shoe shaped constriction at the exist zone. Therefore, the two parameters of the thicknesses in the parallel region and the minimum film thickness have been used for the design of machine elements and the formulas of the thickness in the line contacts [4] and the point contacts [5,6] have been provided. Since the hydrodynamic film thickness become zero as the entrainment speed goes zero according to the above theory, it has been believed commonly that the hydrodynamic action would not occur in the conditions of the zero entrainment velocity.

However, Cameron [7] had already indicated that the existence of lubricated film even under opposite sliding and rolling conditions i.e. in zero entrainment velocity, which has not been hardly known until recent years in spite of discovery about 50 years ago. He carried out an experiment with two disks of bronze and steel and measured traction and electrical resistance between the disk surfaces. The experimental results indicated the resistance increased and the traction decreased as well as scuffing did not occur with increasing the sliding speed. Cameron supposed the film formation would be caused by the temperature difference of the bounding surfaces at both the inlet zone and called

*Present address : Department of Mechanical Engineering, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan

the action the viscosity wedge action. Dyson et al [8] also carried out an experiment at the conditions in 1967. They measured the film thickness by the electrical capacitance method and showed an abnormal phenomenon that was that the film thickness increased as load was increased. Although their studies were significantly impact results and hence referred by several researchers [9-12], the study of the film formation in the conditions had not continuous until recently.

Additionally, Kaneta et al [13] also found such a conical depression in glass-steel contacts that was not able to predict by the conventional theory with a ball-on-disk type apparatus and called the depression "dimple". When the glass surface is stationary and the ball surface is moving, the dimple does not exist. However, at the pure sliding conditions with moving the surface of the disk only, the dimple appeared at the centre of the contact area. They thought that the mechanism would be based upon squeeze action toward the sliding direction because the ball surface sank in the disk surface. Furthermore, they supposed that the oil film in the dimple would behave as solid-lilke state due to be very high pressure [14].

After the Kaneta's discovery, some researchers indicated models to explain the dimple phenomena, and during the progress of the research, a dimple was also discovered in the opposite sliding and rolling conditions. Ehret et al [15] presented a dimple model based upon the oil solidification and indicated the shape of the dimple similar to the Kaneta's results. However, since their model treated the tendency of the oil solidification as one of input parameters, the model was not able to explain the difference of the film profiles between the pure sliding conditions with moving the ball surface only and that moving the disk surface only. Kudish [16] supposed the difference of the film profiles the stretch action of the surfaces and explained the dimple formation but the elastic properties which Kudish used were unrealistic ones. Qu et al [17] presented a full solution using the generalized Reynolds equation and energy equation and their numerical solution could explain the dimple formation by the viscosity wedge action comparatively. Additionally, the authors [18] measured temperature rises of the oil film and the surface of the ball by an improved infrared technique developed by Ausherman et al [19] as well as the film profile by the optical interferometry method in the opposite sliding and rolling conditions. The results showed that a large dimple was formed in the contact area and the temperature rise increased at the dimple region markedly. Yang et al [20] presented a theoretical solution using generalized Reynolds equation coupled with energy equation taking into account for the temperature profile across the oil film in the conditions. They also showed that the dimple existed in spite of the zero entrainment speed.

For better understanding of the mechanism of the dimple occurrence and the oil state at the dimple zone, experimental investigation of the temperature profile across the oil film as well as the temperature distribution in the contact area would be needed and a key factor. The viscosity wedge action suggested by Qu et al [17] and Yang et al [20] is caused because the temperature profile across the oil film varies in the sliding direction. Additionally, the extent of the oil solidification suggested by Ehret et al [15] is also influenced by the temperature rise of the oil film. Furthermore, if the state of the oil film is the solid one or close to it, the temperatures on the bounding surfaces and temperature profile across the oil film would be important information to discuss the slip of the solidified oil at the interface between the surface and oil film.

Therefore, in this study the temperature rises and the film thickness were measured to discuss the mechanism of the dimple formation in the EHD conditions.

2. EXPERIMENTAL PROCEDURE

2.1. Experimental apparatus

Figure 1 shows a ball-on-disk apparatus. A sapphire disk or CaF_2 disk of 96 mm diameter and 5 mm thickness and a steel ball of 25.4 mm diameter were driven by AC servo motors independently. Load was applied through a liner guide below an oil bath by an air cylinder. Traction and load were measured by a load cell which was able to detect the forces in three directions.

2.2. Temperature measurement method

Temperature rise was measured by CCD type of an infrared microscope, which was positioned above the ball-on-disk apparatus. The spatial resolution of the detector is about 20 μm and the detection range in wavelength is from 3.0 μm to 5.0 μm.

Temperature rises of both the sliding surfaces and

Table 1
Properties of ball and disks

Properties	Steel (SUJ2)	Sapphire	CaF_2
Young modulus, GPa	210	365	75
Poisson's ratio	0.3	0.2	0.28
Density, kg/m^3	7850	3980	3180
Thermal conductivity, W/(mK)	41	27	9.7
Specific heat, J/(kgK)	418	577	854
Thermal diffusivity, $kJ/(m^2s^{0.5}K)$	11.6	7.87	5.13

the oil film were measured by an improved technique and the temperature profile across the oil film was estimated with the experimental results [21]. The method of separation between the radiation from the ball surface and the radiation from the oil film used optical band bass filters developed Ausherman et al [19]. Figure 2 depicts the spectrum of the transmissivity of the filters and disks and the emmisivity of oil. Figure 3 shows the disks for measuring each the temperature. When temperature of the ball surface and the oil film were measured, a disk without coat was used. The emmisivity of oil film has a specific emmitance band from 3.0 μm to 4.0 μm whereas that of steel is constant approximately. Filter B detected the radiation from the ball surface only and Filter F detected the radiation from the oil film as well as from the ball surface. Calibrations in the relation between temperature and intensity of the infrared emission from the oil film were carried out at each oil film thickness calculated by geometrical shape and the intensity of the radiation from the oil film was regarded as difference between calibration data with oil film and that without the oil film. To measure the temperature of the disk surface, another disk was used, which was coated with about 300 nm Cr in thickness on the contact surface to cut off the radiations from the ball surface and the oil film as well as to radiate from itself. The repeatability of the temperature measurements was about 10 % of the value. The temperature profile across the oil film was estimated assuming it to be a parabolic curve with experimental results shown in Fig.4.

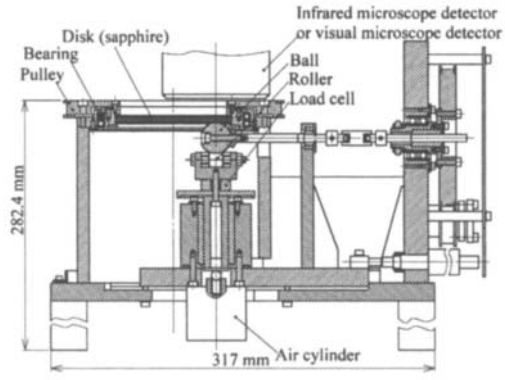


Figure 1. Experimental apparatus

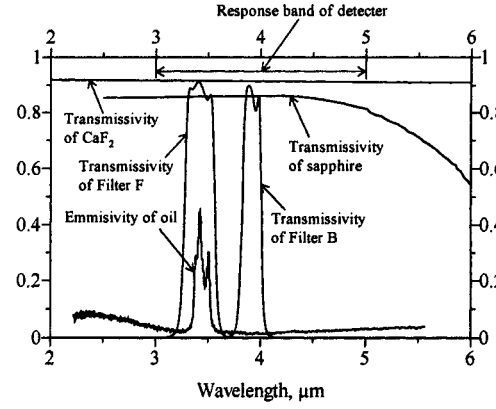


Figure 2. Spectrums of transmissivity of filters, disks and emmisivity of oil

2.3. Film thickness measurement method

It is necessary for exchanging from the intensity

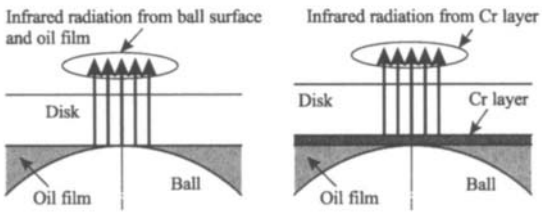


Figure 3. Disks for measuring temperatures

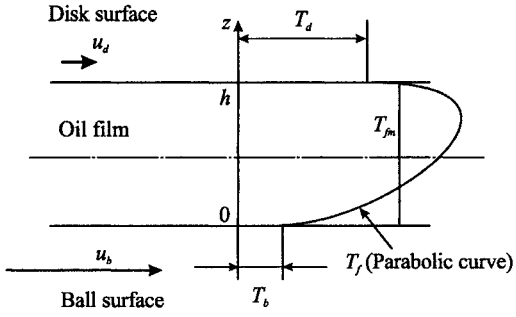


Figure 4. Temperature profile across oil film

of the oil emission to the temperature as well as for knowing the film thickness itself to measure the oil film thickness. The film profile was measured by the two colour optical interferometry technique developed by Wedeven [22]. The light was illuminated in the contact area by a metal halide lamp through a optical fiber, filters of green and red colour and a microscope whose a 5 × objective lens is NA number of 0.13. The interference fringe pattern of the film thickness was recorded by digital camera attached to the microscope.

2.4. Experimental conditions

Experiments were carried out in two basic conditions. One condition was sapphire-steel contacts with a paraffinic oil (P100) to investigate characteristics of the film formation in the opposite sliding and rolling conditions. The other condition was CaF_2-steel contacts with an aromatic traction oil (Santotrac100) to investigate the mechanism of the dimple occurrence in the Kaneta's conditions. The properties of the oils are shown in Table 2. The experiments were carried out at room temperature of 23 ± 1.0 °C.

The slip ratio used in this paper is defined as follows.

$$S = \frac{u_b - u_d}{u_m} \tag{1}$$

where, u_b is the speed of the ball surface, u_d is the speed of the disk surface and u_m is the entrainment speed defined as $u_m = (u_b + u_d)/2$. When the slip ratio $S < 2$, both the surfaces are moving in same direction. When the slip ratio $S > 2.0$, both the surfaces are moving against opposite directions. When the slip ratio equals to infinity, both the speeds of the surfaces are same and the entrainment speed equals to zero.

3. EXPERIMENTAL RESULTS

3.1. Opposite sliding and rolling conditions

Figure 5 shows the interference fringe patterns of the film thickness for various slip ratios at a constant sliding speed of 2.0 m/s and load of 123 N (mean Hertzian pressure p_m of 0.97 GPa and Hertzian radius a_r of 201 μm). As the slip ratio is close from -8.0 to infinity, the dimple appears and increases at the left side of the contact area in spite of decreasing the entrainment speed. When S is infinite, the

Table 2
Properties of oils

Oil	Specific density 15/4 °C	Kinematic viscosity mm²/s			Pressure-viscosity coefficient, (GPa)⁻¹		
		23 °C	40 °C	100 ° C	23 °C	40 °C	100 ° C
P100	0.868	231	88	11	24	21	15
Santotrac100	0.978	318	91	8	41	32	17

dimple exists at the centre of the contact area. Furthermore, as the slip ratio varies from infinity to 8.0, the dimple decreases and moves toward the right side of the contact area with changing the dimple shape. Moreover, as is focusing attention on the symmetry of the slip ratio conditions, which is, for example, difference of the results between S = -18.0 and S = 18.0, it is found that there is not large difference between the dimple shapes but there is some difference between the film thicknesss according to the colours of both the interference images. Next, it is found that the cavitation region varies out of the contact area for various slip ratios. When the slip ratio S equals to ±8.0, within the width of the contact diameter across the sliding direction, the cavitation zone disappears and the zone filled with the oil was formed. Out of the width, the striated cavitation was formed and disappears gradually as S is close to infinity.

Figure 6 illustrates temperature rises and film profile in the centre of contact area along the sliding direction for S = -8.0, ±∞ and 8.0. The left graph indicates the data in the centre of the contact area along the sliding direction and the right one indicates the temperature distributions in the oil film of the cross section in the centre line. When S = -8.0, the film thickness at the right side of the contact area is about 0.2 μm and increases gradually toward the left side of the area. Passing through the centre of the contact area, i.e. the film thickness increases markedly. The maximum surface temperature rises of both the surfaces are about 100 K and the profiles cross at the centre of the area. The temperature of the oil film, which is markedly higher than that of both the surfaces, increases markedly at the dimple zone and reaches 357 K. The temperatures of both the surfaces also increase at the dimple zone. When the slip ratio is infinite, the film thickness still keeps about 0.1 μm at both the side of the contact area and the maximum film thickness reaches up 2.0 μm at the centre of the contact area, where the oil temperature rise increases dramatically. For comparison between S = -8.0 and 8.0, it should be noted that the film thickness for S = -8.0 is somewhat thicker than that for S = 8.0 and the temperature rise in case of S = -8.0 is also higher than that of S = 8.0.

3.2. Glass-steel contacts with pure sliding

Figure 7, 8 and 9 show the interference fringe pattern of the film thickness, profiles of the temperature rises and the film thickness and temperature distribution across the oil film, respectively, at the CaF_2-steel contacts with pure sliding for a disk surface speed u_d of 0.8 m/s against the stationary ball surface, mean Hertzian pressure p_m of 0.36 GPa and Santotrac100 oil. It is found that the dimple appears at the centre of the contact area similar to Kaneta's results [1,2]. The temperature rise of the disk surface, which is lower thermal conductivity and moving, is higher than that of the ball surface. The oil temperature rise increases at the dimple zone and reaches 211 K. As is seen in Fig.9, it is found that the heat transfers into the ball surface and the peak of the temperature profile across the oil film is close to the side of the disk surface.

4. DISCUSSION

4.1. Temperature profile in contact area

The results of temperature measurements show such characteristic results as high temperature rise in the dimple zone, difference of temperature profiles between the bounding surfaces and an asymmetrical profile of temperature in the symmetrical conditions of the speed.

The extraordinary temperature rise in the dimple zone would be caused by high pressure. It can be predicted easily that the pressure would be higher than the pressure predicted by the Hertzian theory at the dimple zone according to the deformation. As shear force is added in the high pressure zone, heat is generated in the oil film. Furthermore, the heat generated in such thick film as the dimple is hard to diffuse into the surfaces. Therefore, the temperature would increase in the dimple zone markedly and sometimes reaches above 300 K.

The asymmetrical profile shown in Fig.6 and the higher temperature rise of the moving surface shown in Fig.8 would seem to be caused by difference of thermal properties between the bounding surfaces. According to the theory of temperature rise on a semi infinite volume [23], which has been used for the thermal EHD theory [24], it is well known that the temperature of surface where heat is moving would be influenced by the thermal diffusivity as well as the sliding speed. In the opposite sliding and rolling conditions, when the slip ratio is lower than zero, the speed of the lower thermal diffusivity surface (sapphire) is higher than that of higher thermal diffusivity surface (steel) and when the slip ratio is higher than zero, the speed of higher thermal diffusivity surface (steel) is higher than that of the

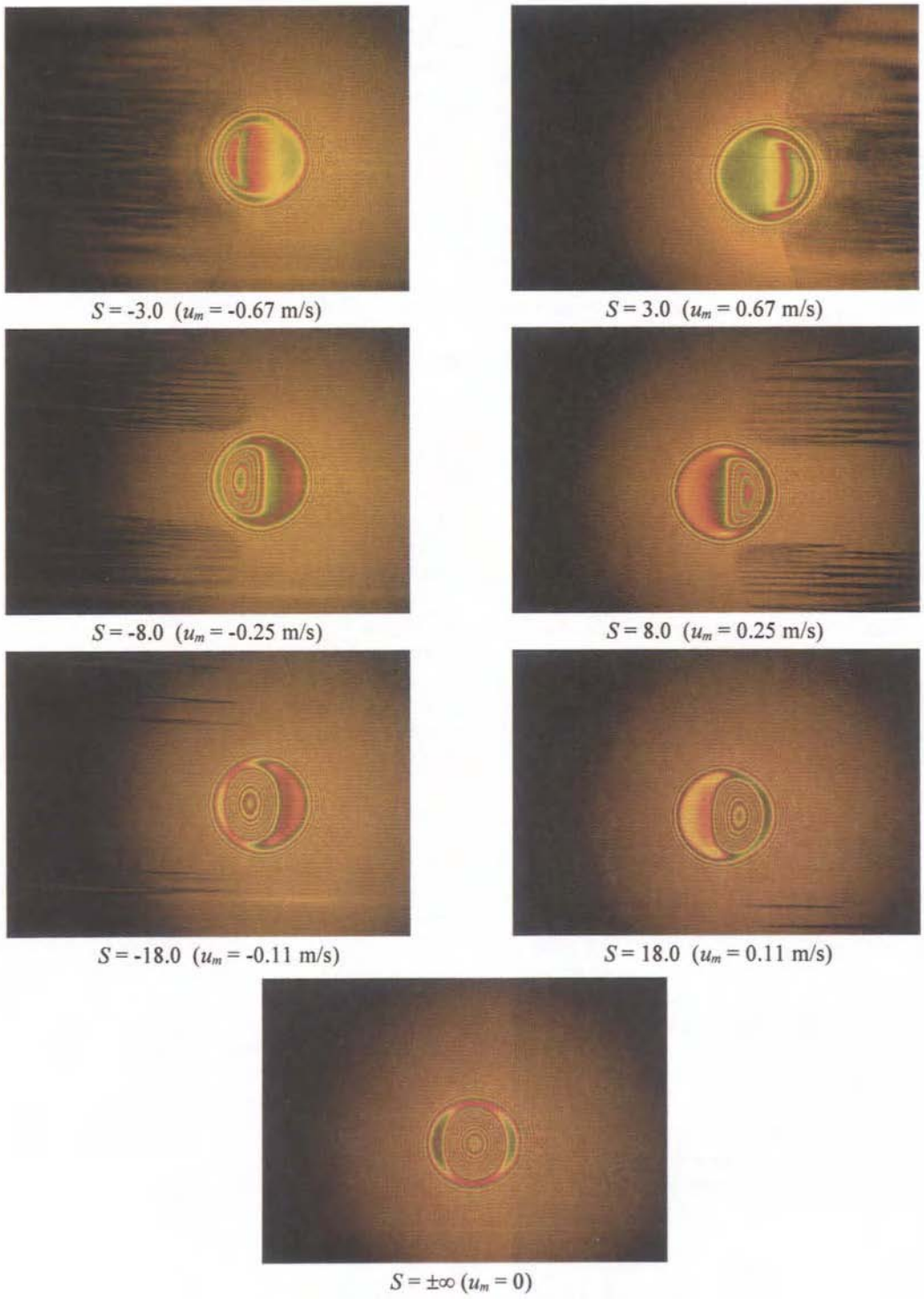
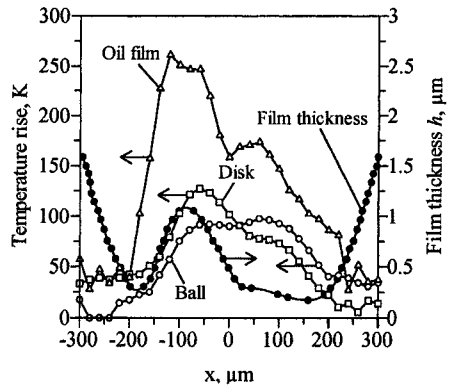
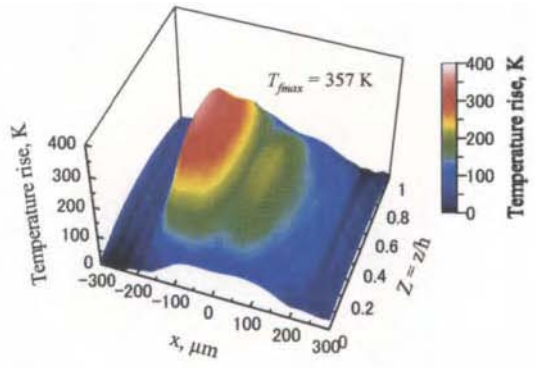


Figure 5. Interference fringe patterns of film thickness for various slip ratios
(p_m = 0.97 GPa, Sapphire-SUJ2 contacts, oil : P100)

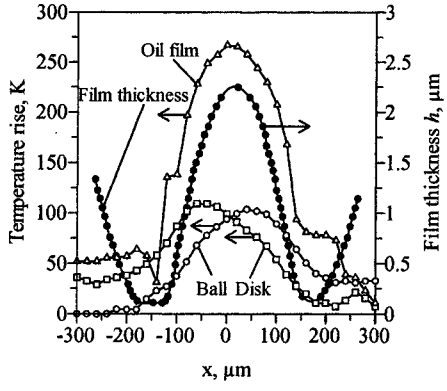


Profiles in sliding direction

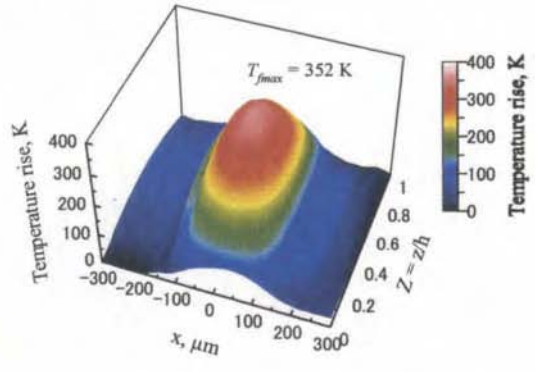


Temperature distribution in oil film

(a) S = -8.0

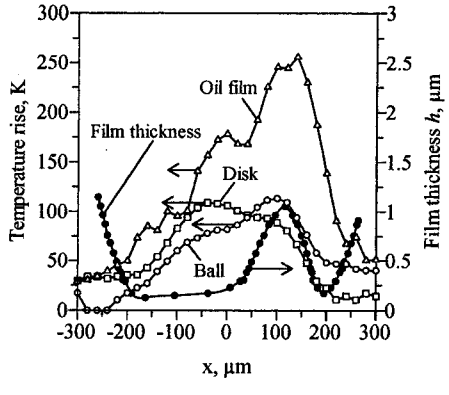


Profiles in sliding direction

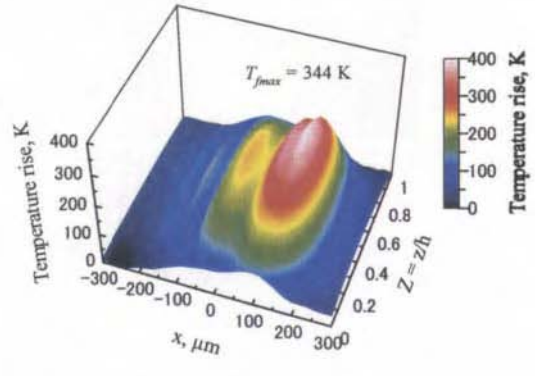


Temperature distribution in oil film

(b) $S = \pm\infty$



Profiles in sliding direction



Temperature distribution in oil film

(c) S = 8.0

Figure 6. Temperature rises and film thickness in centre line along sliding direction (p_m = 0.97 GPa, Sapphire-SUJ2 contacts, oil : P100)

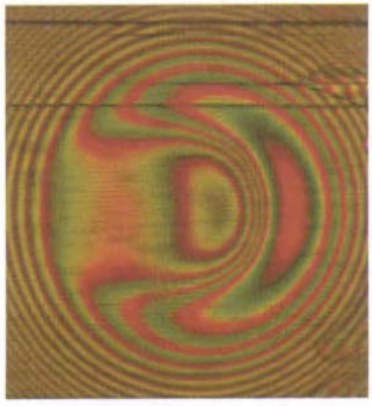
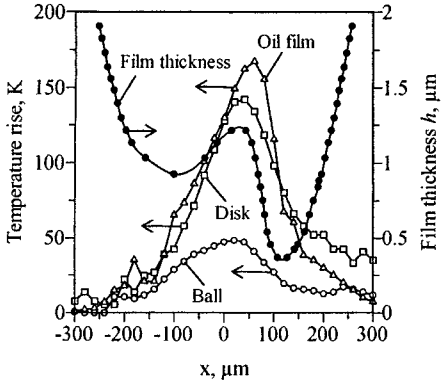


Figure 7. Interference fringe pattern of film thickness in CaF_2-steel contacts



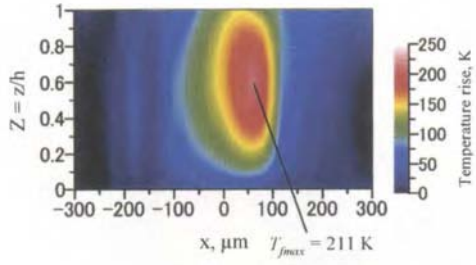


Figure 8. temperature rise of dimple zone in CaF_2-steel contacts

lower thermal diffusivity surface (sapphire). Thus, the conditions of the heat transfer are not symmetrical although the conditions of the speeds are so. Additionally, in the CaF_2-steel contacts, the thermal diffusivity of the CaF_2 disk is lower than the steel ball. When the thermal diffusivity of the

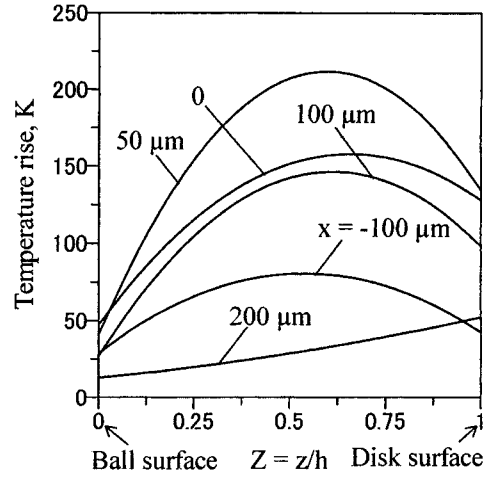


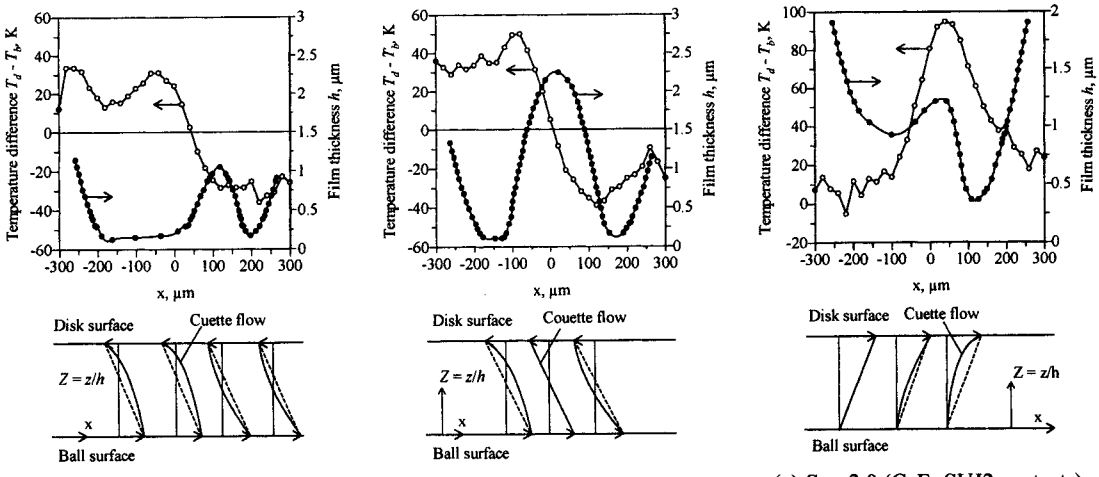
Figure 9. Temperature profiles across oil film in CaF_2-steel contacts

bounding surface are same or grade of difference between sapphire and steel, the surface temperature of the lower speed is higher than that of higher speed [21,24]. However, the thermal diffusivity of CaF_2 is still lower than that of sapphire. Hence, the temperature results would become the higher temperature rise of the moving surface in the CaF_2-steel contacts.

4.2. Mechanism of the dimple formation

From the experimental results, the authors consider the viscosity wedge action to be the cause of the dimple formation because the great temperature difference has been found at the dimple zone. For more quantitatively discussion of the relation between the temperature difference and the film profile, Figure 10 shows the temperature difference between both the sliding surfaces, the film profile and the couette flow suggested by the temperature difference. The figures (a) and (b) indicate the results at the opposite sliding and rolling conditions and the figure (c) indicates that in CaF_2-Steel contacts. It should be noted that great temperature difference exists in the contact area and, especially, the temperature difference varies markedly along the oil film at the dimple zone.

As expected, in the opposite sliding and rolling conditions, the temperature difference between both the surfaces varies dramatically at the dimple zone;



(a) $S = 8.0$ (Sapphire-SUJ2 contacts) (b) $S = \pm\infty$ (Sapphire-SUJ2 contacts) (c) $S = -2.0$ (CaF_2-SUJ2 contacts)

Figure 10. Explanation of dimple formation by viscosity wedge action

hence, the viscosity wedge action would build the hydrodynamic pressure. For the infinite slip ratio, since at both the side of the contact area the oil is more dragged by the lower temperature surface moving into the contact area, the oil flows into the contact area. Thus, the total couette flow into the contact area does not equal to zero. Additionally, since the temperature difference of both the surfaces decreases as is close to the centre of the contact area, the couette flow decreases and pressure increases markedly at the dimple zone.

The disappearance of the caviation zone as seen in Fig. 5 can also explain by the viscosity wedge action. For example, when $S = 8.0$, the direction of the average speed of the bounding surfaces is from left to right on this paper. Focusing attention on the temperature difference at the right side of the contact area, it can be seen that the oil tends to be dragged by the disk surface according to the temperature difference between the bounding surfaces. As a result, the oil would flow from right to left on this paper.

In the glass-steel contacts with pure sliding conditions, the temperature difference also varies at the dimple zone. The couette flow would equal to the entrainment speed at the inlet zone approximately because the temperature difference is very small. However, at the contact area, the flow decreases as is passing downstream because the temperature of the moving surface (disk surface) increases than that of the stationary surface (ball surface).

4.3. State of oil at dimple zone

The state of the oil film in the dimple zone would behave as liquid state according to the high temperature rise in the dimple zone. At the beginning when Kaneta's dimple was discovered, it was thought that the state of the dimple behaved as solid-like state. Ehret et al [15] developed a model based upon the oil solidification to explain the dimple phenomena. However, the solidification of oil is hard to be considered in the conditions of above 200 K. Therefore, we can conclude that the dimple would be formed by the viscosity wedge action due to the temperature variations across the oil film.

For better understanding of the dimple phenomena in the future, the accurate prediction of the viscosity and durability at high temperature and pressure conditions would be needed because the temperature at the dimple zone reaches high temperature which exceeds sometimes above 300 K.

5. CONCLUSIONS

The above results and discussion suggest the following conclusions:

1 Great temperature difference between both the sliding surfaces exists at the dimple zone and varies in the sliding direction.
2 The temperatures of the oil film and both the sliding surfaces increase dramatically at the dimple zone.
3 The viscosity wedge action based upon the results of temperature measurements can explain the dimple formation under the EHD conditions.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Nissan Motor Co., Ltd. for supplying steel balls and disks and Idemitsu Kosan Co., Ltd. for providing oils. The authors are also grateful to Mr. T. Sato who was an undergraduate student of Tokyo Institute of Technology for his effort in carrying out experiments.

REFERENCES

1. D. Dowson and G. R. Higginson, J. Mech. Eng. Sci., 1 (1959) 6.
2. A. W. Crook, Nature, 190 (1961) 1182.
3. R. Gohar and A. Cameron, Nature, 200 (1963) 458.
4. D. Dowson , Proc. Inst. Mech. Engrs., Pt. 3A, 182 (1968) 151.
5. R. J. Chittenden, D. Dowson, J. Dunn and C. M. Taylor, Proc. Roy. Soc. Lond., Ser. A, 387 (1985) 245.
6. R. J. Chittenden, D. Dowson, J. Dunn and C. M. Taylor, Proc. Roy. Soc. Lond., Ser. A, 387 (1985) 271.
7. A. Cameron, J. Inst. Petrol., 37 (1951) 471.
8. A. Dyson, H. Naylor and A. R. Wilson, Proc. Inst. Mech. Engrs., Pt. 3B, 180 (1965-66) 119.
9. F. Hirano, J. Jap. Soc. Trib., 16, 4 (1971) 535.
10. A. Dyson, Trib. Int., 8 (1975) 117.
11. S. Asanabe, H. Ayabe, S. Matsumoto and M. Kanamaru, J. Jap. Soc. Trib., 22, 2 (1977) 77.
12. P. M. Ku, Proc. the 6th Leeds-Lyon Symp. on Trib., (1979) 237.
13. M. Kaneta, H. Nishikawa, K. Kameishi, T. Sakai and N. Ohno, Trans. ASME. J. Trib., 114 (1992) 75.
14. M. Kaneta, H. Nishikawa, T. Kanada and K. Matsuda, Trans. ASME. J. Trib., 118 (1996) 886.
15. P. Ehret, D. Dowson and C. M. Taylor, Proc. Roy. Soc. Lond., Ser. A, 454 (1998) 763.
16. I. I. Kudish, Proc. 1999 STLE/ASME H. S. Cheng Trib. Surve., Orland, Florida (1999) 188.
17. S. Qu, P. Yang and F. Guo, Trib. Int., 33 (2000) 59.
18. K. Yagi, K. Kyogoku and T. Nakahara, Proc. Int. Trib. Conf. Nagasaki 2000, (2001) 567.
19. V. K. Ausherman, H. S. Nagaraj, D. M. Sanborn and W. O. Winer, Trans. ASME. J. Lub. Tech., 98 (1976) 236.
20. P. Yang, S. Qu, Q. Chang and F. Guo, Trans. ASME. J. Trib., 123 (2001) 36.
21. K. Yagi, K. Kyogoku and T. Nakahara, J. Jap. Soc. Trib., 47, 4 (2002) 321.
22. L. D. Wedeven, Ph.D Thesis, University of London (1970)
23. J. C. Jaeger, Proc. Roy. Soc. Lond. N. S. W, 76 (1942) 203.
24. H. S. Cheng and B. Sternlicht, Trans. ASME, J. Bas. Eng., 87 (1965) 695.

SESSION IV
TRIBOLOGICAL COATINGS - ADVANCES IN SURFACE ENGINEERING (1)

Chairman : Professor A Neville

Paper IV (i) **Deformability of the Near Surface Layer of the First Body**

Paper IV (ii) **Thickness Dependency of the Lubricity of Ag Thin Films Deposited on Si (111)7 x 7 Surface**

Paper IV (iii) **Sliding Wear Behaviour of Nanocrystalline Cr_3C_2-25(Ni20Cr) Coatings at Elevated Temperature**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Deformability of the near surface layer of the first body

A. Eleöd[a], L. Baillet[b], Y. Berthier[b], T. Törköly[a]

[a]Budapest University of Technology and Economics,
1111 Budapest, Bertalan L. u. 2., Z/709, Hungary

[b]INSA de Lyon, Laboratoire de Mécanique des Contacts, UMR CNRS/INSA 5514
Bâtiment Jean d'Alembert, 20, Av. Albert Einstein, 69621 Villeurbanne Cedex, France

The material law of the near surface layer of the friction surfaces was described using the equations of the classical theory of plasticity, taking also into consideration the influence of the state factors (temperature, strain rate, strain and stress state), which modify the plasticity and the deformability in a very wide range. Using unalloyed carbon steel and rapidly solidified, power metallurgy, hypereutectoidal aluminium alloy, the influence of the hydrostatic pressure on the plasticity and deformability was determined experimentally.

1. INTRODUCTION

Under the influence of the frictional load, transitional or residual changes can be produced in the near surface layer of the first body. These changes represent the response of the first body to the friction load.

All reversible processes, which occur in the near surface layer, are considered as transitional changes. Residual changes remain after the friction load has been removed, i.e. they are irreversible and create new conditions under repeated loading. The development of TTS (tribological transformation of surface) and cracking, including the particle detachment too, as residual changes of the surfaces, can be traced back to purely tribological origins. The transition between temporary and residual changes, especially in the case of ductile materials, is fundamentally related to yielding. But *the range between reversible and irreversible changes, as far as the onset of the irreversible processes is concerned, may vary over a very wide range under the influence of state factors* such as temperature, strain rate and stress state.

Some aspects of the accommodation of the near surface layer to the load during the friction can be described by the classical theory of elasticity and plasticity, but the development of TTS and the particle detachment cannot be treated by the classical theories.

To usefully extend the modelling of the TTS [1-4], it is essential to know its mechanical characteristics. Before developing special investigations to determine these characteristics, it is necessary to study these possibilities of characterisation, which are offered by the classical plasticity. Considering, that the activated deformation during the developing of the TTS are higher than the deformations, which can be explained by the classical theory of plasticity, it seems to be necessary, based on the pioneering work of Kármán [5], to take into consideration the effect of the hydrostatic pressure on the plasticity and deformability.

To determine the effect of the hydrostatic pressure on the plasticity and deformability, we need at least four independent investigations. The objective of the present work is also to choose those, which permits to extrapolate on the activated deformation of the TTS.

2. THE CONCEPT OF DEFORMABILITY AND INFLUENCING FACTORS

In the present discussion 'deformability' is construed as including the whole domain of deformation, both elastic and plastic, and means the maximal deformation, which can be supported by the given contact without inducing fracture.

The deformability of a purely elastic contact can be characterised by the strain to elastic shakedown limit, which depends on the actual yield stress. The elastic deformability is often related to a ratio of some power of the hardness and elastic modulus of the surface layer [6]. This approach is based on the Hertzian theory. Furthermore, it assumes that the average Hertzian pressure is equal to the hardness at the onset of plastic flow.

One of the characteristics of elastic Hertzian contact is that over the elastic shakedown limit, the place of the beginning of the plastic flow varies in function of the friction coefficient [7]. In the case of low friction coefficient it begins in a known depth below the surface due to the maximal shear stress. On the surface, at the contact point or on the contact line, the stress state is hydrostatic compression and the deformation remains elastic, even if a residual change happens below the surface. If the friction coefficient increases, the plastic flow can begin on the surface.

In the case of plastic deformation, many terms have been used to characterise the concept of deformability such as plastic deformability, capacity of deformation and plastic strain to failure. These all refer to the accumulated plastic strain which can be tolerated by the material prior to plastic instability or fracture.

Plasticity and plastic deformability, however, *are not identical, but are connected terms.* The plasticity of a given material can be characterised by its resistance against the plastic deformation. The plasticity is usually described by the flow curve, i.e., the plot of actual yield stress with respect to plastic strain. Studies of the effects of state factors on the plasticity of materials were undertaken quite early (Nádai 1925, Prandtl 1930, Alder-Phillips 1954, etc.). All experimentally verified empirical and semi-empirical equations show that increasing strain rate and decreasing temperature, influence the yield stress in the same sense. The increasing of temperature, below the temperature of recrystallisation, increases the kinetic energy of the dislocation, above this temperature it activates the dynamic recrystallisation and the polygonisation. Increasing the strain rate yields the decreasing of time available for the moving as far as for the interaction of the dislocations. The effect of the stress state on the plasticity characterised by the flow curve, is less significant.

As an example, to represent the influence of the strain rate and the temperature on the flow curves, we used the flow curves of the C45 unalloyed carbon steel (its chemical composition is shown in Table 1.). The curves were determined by upsetting tests and the results were corrected to frictionless value [8].

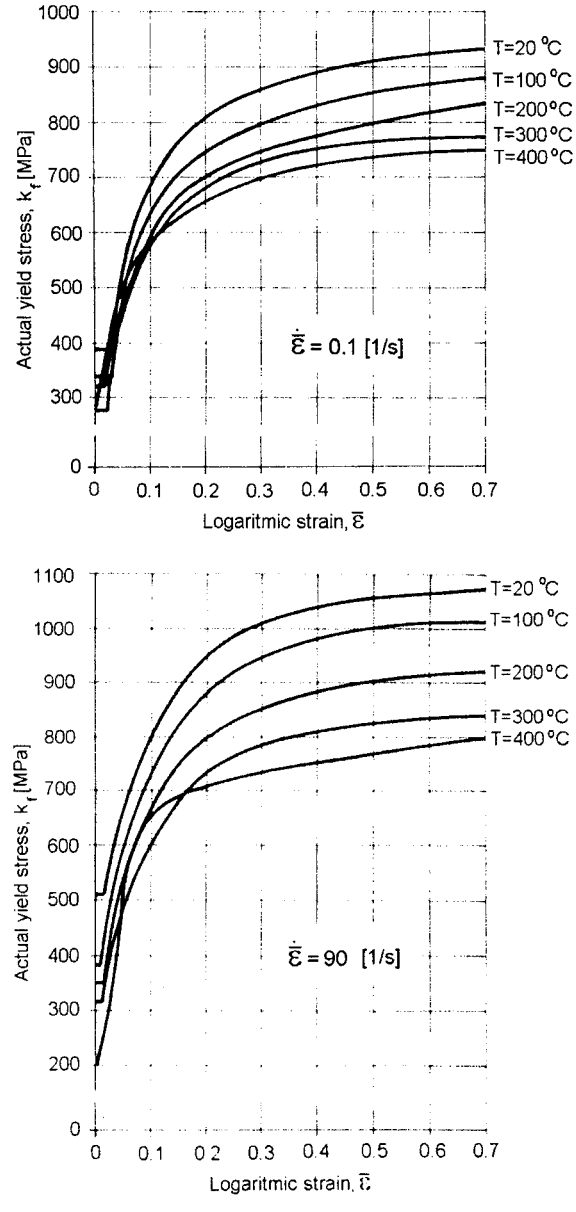


Figure 1. Dependence of the actual yield stress of the C45 steel on temperature (T) and on strain rate ($\dot{\bar{\varepsilon}}$), without superimposed hydrostatic pressure.

For the analytical description of the flow curves power law is generally used:

$$k_f = C \cdot \bar{\varepsilon}^n \tag{1}$$

where: k_f – is the actual yield stress, which depends on the structure of the material, on temperature and on the stress and strain state,

$\bar{\varepsilon}$ - is the comparative strain,

$n = f(T)$ – is the hardening exponent ("n value"), which depends on the temperature.

The C constant of the power law can be determined from the continuity of the transition between elastic and plastic deformation. At the border of the uniaxial and biaxial stress state the strain calculated by the Hooke-law and the strain from the flow curve have to be the same:

$$\frac{\bar{\sigma}}{E} = \left(\frac{k_f}{C}\right)^{\frac{1}{n}} \tag{2}$$

where: $\bar{\sigma}$ - is the comparative stress,

E – is the Young-modulus.

At this point both the comparative stress and the actual yield stress are equal to the standardised yield stress ($R_{p0.2}$), which was determined at room temperature and in uniaxial stress state. So the C constant of the power law yields from the equation (2):

$$C = R_{p0.2}\left(\frac{E}{R_{p0.2}}\right)^n \tag{3}$$

In Equation (3) n is, as appropriate, the exponent of the flow curve at room temperature.

Taking into consideration, that the yield stress of the C45 steel is $R_{p0.2} = 345\ MPa$, as well as the Young-modulus is $E = 2.1 \cdot 10^5\ MPa$, we can fit a power law into the quasi-static ($\dot{\bar{\varepsilon}} = 0.1$ [1/s]) flow curve belonging to room temperature (see Figure 1.). Using Equation (1) and (3) we can determine the constant and the exponent of the power law of the flow curve at room temperature:

$C = 987\ MPa,$
$n = 0.16$

The effect of the strain rate is described by the Alder-Phillips formula, which had been determined earlier empirically, but its faithfulness was verified later in thermodynamical way [9]:

$$k_f = k_{f0}\left(\frac{\dot{\bar{\varepsilon}}}{\dot{\bar{\varepsilon}}_0}\right)^m \tag{4}$$

where: k_{f0} – is the actual yield stress determined by referent strain rate,

$\dot{\bar{\varepsilon}}_0$ - is the referent deformation velocity, [1/s],

$\dot{\bar{\varepsilon}}$ - is the actual deformation velocity, [1/s],

$m = f(T)$ – is the velocity exponent ("m value"), which depends also on the temperature.

The temperature dependence of the calculated n and m values of the C45 steel is shown in the next figure:

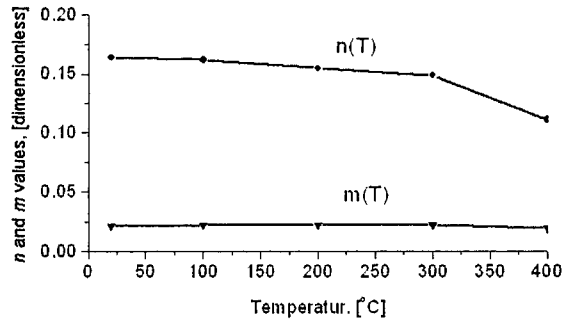


Figure 2. The temperature dependence of the n and m values in the case of C45 steel.

Figure 2 shows that the velocity exponent (m value) seems to be constant in the examined range of temperature. That means, in the examined range of temperature the dynamic processes of structure changes have not started yet.

There is a known relation between the hardening exponent and the velocity exponent at the limit of the stable (uniform) plastic deformation [9]:

$$\bar{\varepsilon}_m = \frac{n}{1-m} \tag{5}$$

In Equation (5) $\bar{\varepsilon}_m$ is the comparative strain, which belongs to the maximal force during the tensile test. Substituting the exponents belonging to

room temperature for Equations (5), we received $\bar{\varepsilon}_m$ = 0.17, which value is equal to the measured limit of the uniform plastic deformation during the tensile test.

According to experience, the actual yield stress of the metals, which are devoid of allotropic transitions, decreases exponentially, if the temperature stays below the temperature of recrystallisation [9]:

$$k_f = k_{f0} \cdot e^{-\beta \cdot \Delta T} \tag{6}$$

where: β – is a constant, which depends on the material,

ΔT - is the deviation from room temperature, [°C].

Based on the measured results, we calculated $\beta = 8 \cdot 10^{-4}$ for the C45 steel.

It is possible to express the effect of the state factors influencing the actual yield stress a single equation:

$$k_f = C \cdot \bar{\varepsilon}^n \cdot \left(\frac{\dot{\bar{\varepsilon}}}{\dot{\bar{\varepsilon}}_0} \right)^m \cdot e^{-\beta \cdot \Delta T} \tag{7}$$

where: both n and m depend on temperature.

Equation (7), completed with the Hooke-law, is suitable to define the material law of the near surface layer of the first body, but it is not enough to describe either the development of TTS or the beginning of the particle detachment.

Apart from the state factors depending material law, it is also necessary to know the critical strain to fracture, i.e. the deformability, if we want to model the friction processes reliably.

Concerning the deformability, especially because of its sensitive dependence on the state of stress, there is, typically, insufficient information available. *This indicates the importance of the stress state while only the deviator-tensor is responsible for plastic deformation, fracture is determined by the whole stress state.* The clearly ductile fracture is a consequence of the joining of concentrated dislocation masses to a continuos deficiency at the grain boundaries. This process is retarded by the component of hydrostatic compression of the stress state. Ductile fracture is determined by the crack propagation, which is a consequence of an interaction between the shape and size of the initial crack as well as its stress field.

Maier [10] represented first the plastic strain to fracture against the ratio of the average of the three principal stresses and of the Misesian yield stress (see Equation 8). The importance of the Maier's stress parameter is that it is the ratio of two principal components of the stress state, i.e. it is the quotient of the hydrostatic sphere-tensor and the flow stress, which is connected to the second invariant of the deviator-tensor (see Equation 15).

Adapting Maier's representation of deformability, Steiger [11] established that the plastic strain to fracture couldn't be represented by only one curve and in practice it must be determined by an array of curves.

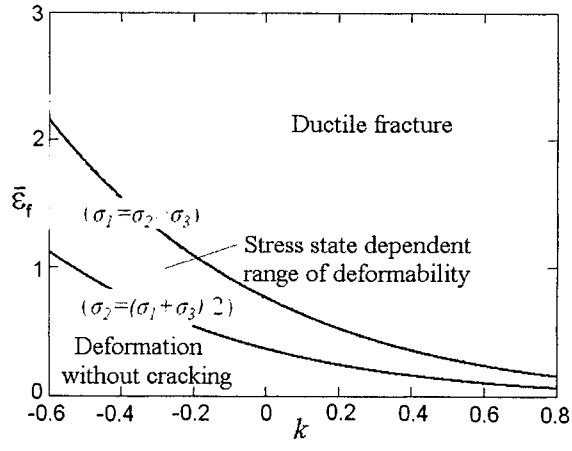


Figure 3. The limit curves of deformability of the C45 steel (at atmospheric pressure and room temperature).

In Figure 3, $\bar{\varepsilon}_f$ is the equivalent strain to fracture, k is the Mayer's stress state index:

$$k = \frac{\bar{\sigma}_m}{k_f} = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3k_f} \tag{8}$$

where: k_f is the actual yield stress, which is equal to the comparative stress at the moment of the plastic flow:

$$k_f = \frac{1}{\sqrt{2}} \sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2} \tag{9}$$

The diagram verifies the experience, that the deformability of a given material is maximal, if its stress state is near hydrostatic ($\sigma_1=\sigma_2>\sigma_3$), or

contrarily, it is less deformable under plain strain conditions ($\sigma_2=(\sigma_1+\sigma_3)/2$).

Bogatov [12], completing the Maier-Steiger representation of deformability by a third invariant of the stress state (by the Lode-parameter), represented the variation of the plastic strain to fracture as a function of the stress state indexes (k, μ_σ) by a four-parametrical function (see Equation 10), which describes a surface:

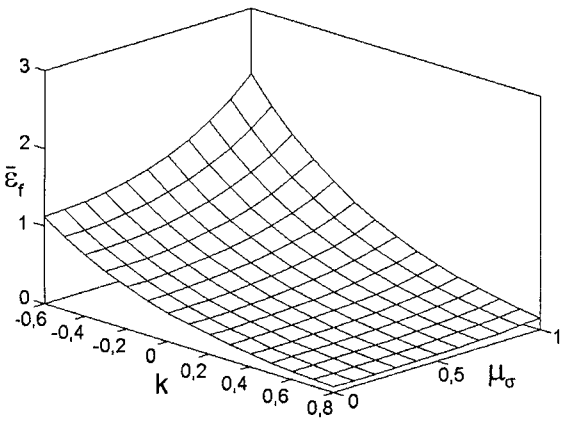


Figure 4. The boundary surface of deformability of C45 steel (at atmospheric pressure and room temperature)

The second independent index of the stress state is the Lode-parameter [13]:

$$\mu_\sigma = \frac{2\sigma_2 - \sigma_1 - \sigma_3}{\sigma_1 - \sigma_3} = \frac{\sigma_2 - \bar{\sigma}}{\sigma_1 - \sigma_3} \qquad (10)$$

The Bogatov's boundary surface of deformability (see Figure 4.) can be described by the following equation:

$$\bar{\varepsilon}_f = (a + bk + c\mu_\sigma^2)^d \qquad (11)$$

In Equation (11) a, b, c, d are material dependent constants which can be determined experimentally, using four independent investigations.

From the point of view of deformability, the importance of the Lode-parameter can be interpreted indirectly. The Lode-parameter can also be expressed by the invariants of the deviator-tensor and it also contains the hydrostatic sphere-tensor. In fact, the Lode-parameter grants the position of the centre of the circle of principal stresses (i.e. the Mohr-circle) compared to the hydrostatic component and to its diameter. The Lode-parameter demonstrates clearly that the greater the hydrostatic component of the stress state is, the smaller the importance of the difference is between the principal stress components.

The four-parametrical Bogatov's description of the deformability was extended by Darvas and Ziaja [14]. Based on their experimental work, they used an additional function to describe the plastic strain to fracture:

$$\bar{\varepsilon}_f = (C_1 + C_2|\mu_\varepsilon|)\exp[(C_3 + C_4|\mu_\varepsilon|)k] \qquad (12)$$

In Equation (12) $C_1 ... C_4$ are material depending constants, k is the Maier's stress parameter, μ_ε is the Lode-parameter. They suggested to determine the Lode-parameter using the strain components from the real history of deformation:

$$\mu_\varepsilon = \frac{2\varepsilon_2 - \varepsilon_1 - \varepsilon_3}{\varepsilon_1 - \varepsilon_3} \qquad (13)$$

3. THE INFLUENCE OF THE HYDROSTATIC PRESSURE ON THE PLASTICITY AND DEFORMABILITY

Based on the previous descriptions, it is conceivable, that the plasticity and deformability of an active contact region of the friction surface, depend on the actual stress and strain state of the given environment. The stress and strain state of a loaded contact can be described by a stress and a strain tensor.

As it is well known, the stress tensor can be decomposed into two components, a deviator-tensor which describes the deviation from the hydrostatic sphere tensor and a hydrostatic sphere tensor:

$$\sigma_{i,j} = s_{i,j} + \sigma_m \cdot \delta_{i,j} \qquad (14)$$

where: $\delta_{i,j}$ – is the Kronecker symbol,

$\delta_{i,j} = 1$ *if* $i = j$ and $\delta_{i,j} = 0$ *if* $i \neq j$

σ_m – is the hydrostatic sphere tensor which can be expressed as one third of the first scalar invariant of the stress tensor,

$s_{i,j}$ – is the stress deviator.

The stress deviator, which expresses the difference from the hydrostatic stress state (see Equation 16), causes plastic deformation, while the sphere tensor causes only elastic deformation.

The intensity of the stresses can be characterised by the relation between the second invariant of the stress tensor and the uniaxial comparative stress:

$$f = \frac{1}{2} s_{i,j} s_{i,j} - \frac{\bar{\sigma}^2}{3} \tag{15}$$

As long as the value of the function f is negative, the material sustains only elastic deformation. If $f=0$, plastic deformation happens. If $f>0$, then the material will break, or, in the case of ductile material, below the critical strain to fracture, strain hardening will happen and the state of $f=0$ sets in, which means that $\bar{\sigma} = k_f$.

Characterising the stress state by principal stresses, Equation (14) can be developed:

$$\sigma_{i,j} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} = \begin{bmatrix} (\sigma_1 - \sigma_m) & 0 & 0 \\ 0 & (\sigma_2 - \sigma_m) & 0 \\ 0 & 0 & (\sigma_3 - \sigma_m) \end{bmatrix} + \begin{bmatrix} \sigma_m & 0 & 0 \\ 0 & \sigma_m & 0 \\ 0 & 0 & \sigma_m \end{bmatrix} \tag{16}$$

In Equation (16) σ_m is the hydrostatic component of the stress state, which is equal to the average of the principal stress components:

$$\sigma_m = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3} \tag{17}$$

Equations (14) and (16) make it clear, that every compression stress state always has a hydrostatic component. Research on the influence of the hydrostatic pressure on different material behaviour goes back to over 90 years. First Kármán [5] documented with his upsetting tests under hydrostatic pressure, that the brittleness or plasticity are only state of the materials, not characteristics, as well as that the deformability is basically dependent on stress state. The effect of the hydrostatic pressure on the deformability was first studied entirely by Bridgman [15]. The monograph of Pugh [16], who resumed not only his own, but also the entire published results on this field, followed Bridgmann's scientific achievement. Today the name Khulman-Wilsdorf hallmarks the papers publishing new results about investigations under high pressure [17].

Concerning the effect of the hydrostatic pressure on the mechanical properties of metals, it was found that the hydrostatic pressure had no effect on the lower yield stress for mild steel nor on the stress-strain curve for silicon-aluminium alloy. However, a small reduction in the upper yield stress was observed on mild steel [16]. For more strain hardened specimens the decrease in yield stress was much higher. In unanimous option of all researchers the deformability increases under effect of superimposed hydrostatic pressure [16].

According to experience, the ductile fracture of a free surface ensues always under plane stress state condition [14]. On one hand, we can assume that the presence of the hydrostatic sphere tensor eliminates, or holds up the development of this plane stress state, on the other hand, it can be conducive to rejoin the discontinuity at the border of the grains.

4. INVESTIGATIONS

4.1. Tested materials

To determine the increasing of deformability under the effect of superimposed hydrostatic pressure, we used two different materials (see Table 1.).

Table 1.
Composition of the investigated materials

Material / Components	C	Mn	Si	Cr	Ni	Mg	Cu	Fe	Al
C 45 steel	0,45	0,5	0,2	<0.3	<0.3		<0.3	rest	
AlSi26Ni6 RS/PM			26,4		6,0	0,04	<0,01	0,41	rest

On one hand, a ductile material, C45 unalloyed carbon steel was investigated. On the other hand, a rapidly solidified, power metallurgy, hypereutectoid aluminium alloy (AlSi26Ni6), which is practically not deformable at room temperature and atmospheric pressure.

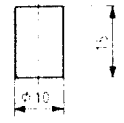
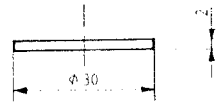
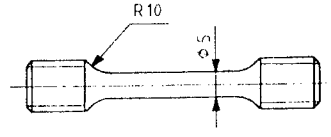
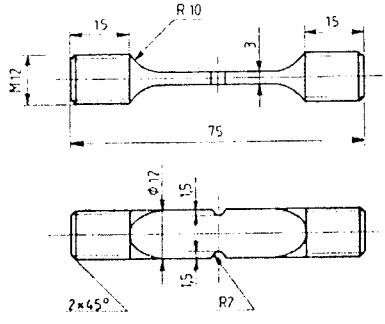
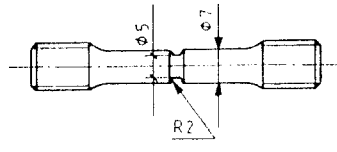
For the experiments under hydrostatic pressure, an isostatic pressure chamber (see Figure 5.) was installed on a four-column hydraulic press. The diameter of the punch is 40 mm, the maximal pressing force is 1000 kN, so the maximum of the available nominal hydrostatic pressure was 800 MPa. The range of the hydrostatic pressure for each experiment was 400, 600, 800 MPa. The velocity of the punch was constant, u = 1 mm/s.

To realise the tensile tests, the use of a special cartridge (see Figure 6.) was necessary. For the upsetting and the punching tests we used different specimen holders (see Figure 7.).

4.2. Experiments

To calculate the parameters of the boundary surface of deformability, upsetting tests, three different tensile tests and a punching test were used.

Table 2.
Investigations and specimens to determine the deformability under hydrostatic pressure

Investigation	Geometrical properties of the specimens [mm]	Stress state coefficients	
		k	μ_σ
Upsetting test:		-1/3	1
Cylindrical punching test:		0	0
Cylindrical tensile test:		0,41...0,45	1
Notched planar tensile test:		$1/\sqrt{3}$	0
Notched cylindrical tensile test:		0,72...0,79	1

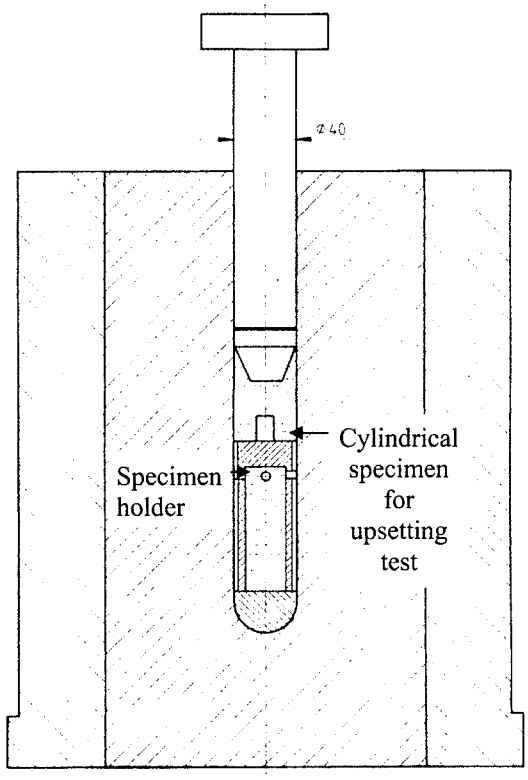


Figure 5. The isostatic pressure chamber with the cylindrical specimen and the specimen holder installed for the upsetting test.

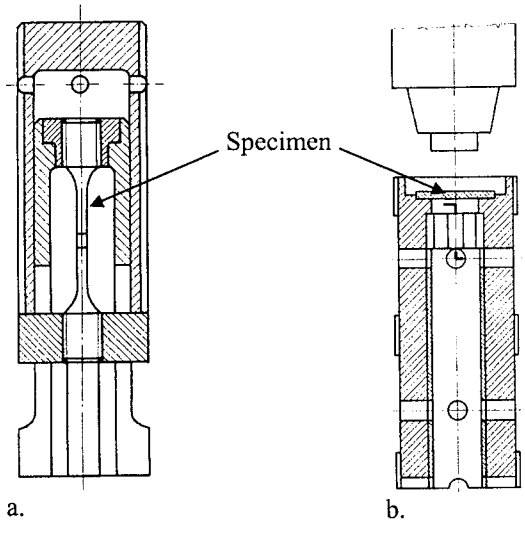
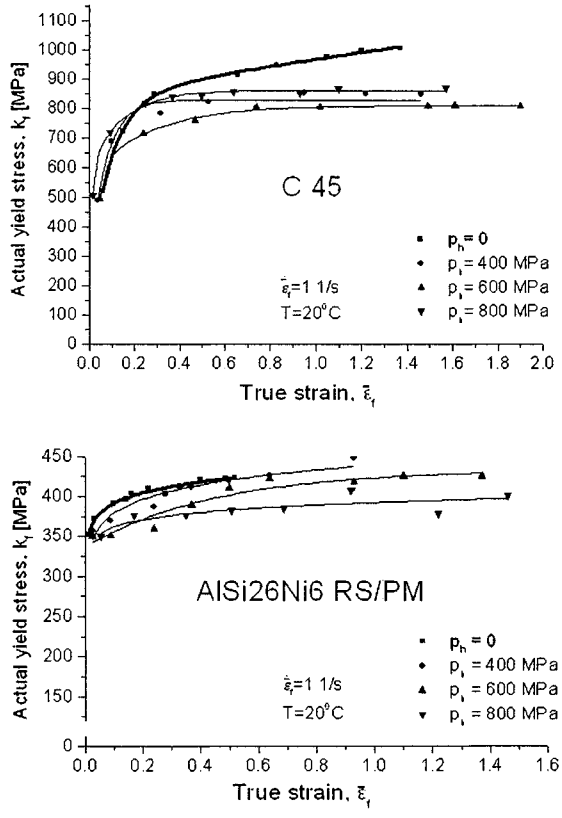


Figure 6. Cartridges for tensile test (a) and specimen holder for punching test (b).

4.3. Evaluation of the results

The flow curves received by upsetting tests are shown in Figures 7 and 8. The curves prove that the tested materials became more plastic under hydrostatic pressure, as well as their plastic state was similar to the perfect plasticity (there is no more hardening).



Figures 7 and 8. Flow curves of the tested materials.

The increase in plasticity is different for the tested materials. The aluminium alloy, which is more brittle at room temperature and atmospheric pressure, became more plastic under hydrostatic pressure than the ductile steel. It is also interesting, that the highest hydrostatic pressure does not cause the highest plasticity.

The results of the deformability tests are shown in Figures 9 and 10. The conclusions are similar to the conclusions given by the flow curves. Under the effect of the hydrostatic pressure, the strain to fracture increased, and in this case, the 600 MPa pressure delivered the best results as well. On the other hand, the difference between the upper and

lower limit curves was smaller than at atmospheric pressure.

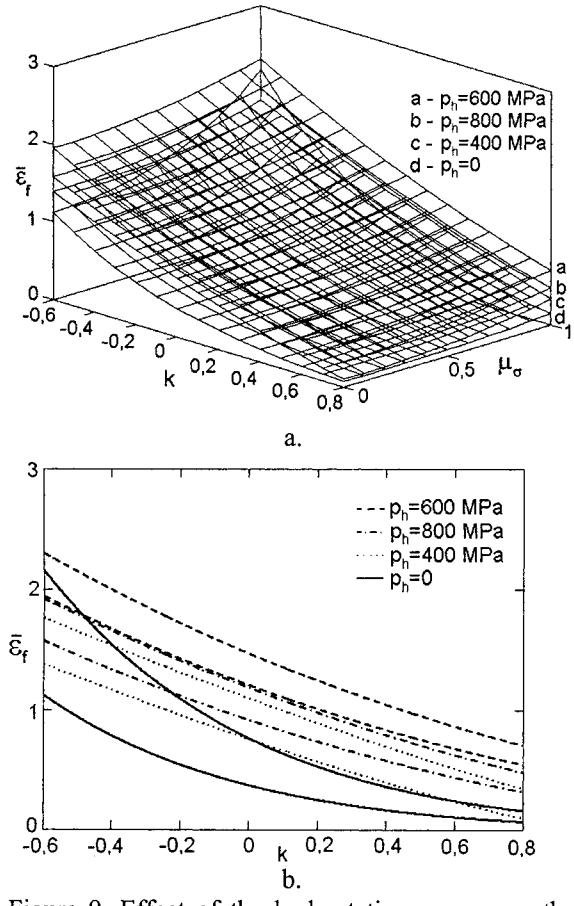
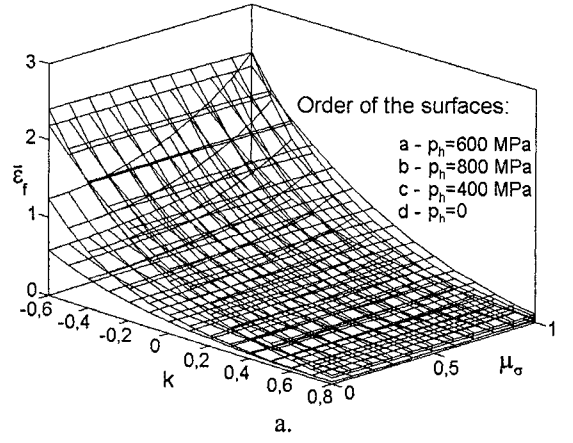


Figure 9. Effect of the hydrostatic pressure on the deformability of C45 steel; boundary surface of deformability (a), limit curves of deformability (b), (upper curves belong to μ_σ=1, lower curves to μ_σ=0).



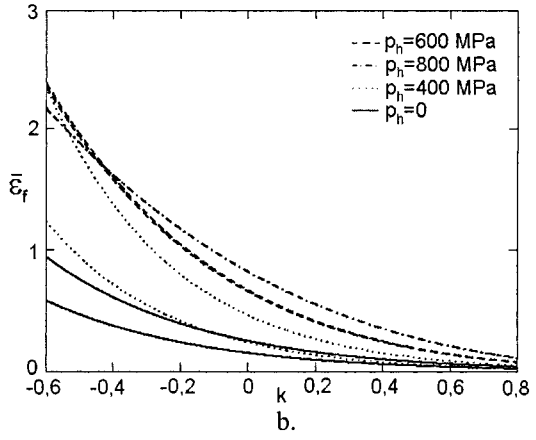


Figure 10. Effect of the hydrostatic pressure on deformability of RS/PM aluminium alloy; boundary surface of deformability (a), limit curves of deformability (b), (upper curves belong to μ_σ=1, lower curves to μ_σ=0)

In the case of the hypereutectoid aluminium alloy, it is especially visible, that the dependence of the deformability on the Lode-parameter became negligible, even in the region of the higher hydrostatic pressure, there is no more difference between μ_σ=0 et μ_σ=1 curves (see Figure 10).

5. DISCUSSION

The stress deviator is responsible for the plastic deformation, but the deformability is governed by the hydrostatic sphere tensor. In other words, the plastic deformation of the near surface layer of the first body is caused by the deviator of the local stress state, but the particle detachment, which is a consequence of the ductile fracture of the near surface layer, is determined by the whole stress state. For the analysis of the behaviour of the near surface layer, we have to take the dependence of the deformability on the actual stress state into consideration too, beside the dependence of the plasticity on the temperature and the strain rate.

Considering that the plasticity and the deformability of a friction surface changes in time and space, Equation (7), completed with the temperature dependent n and m values, as well as Equation (12) may be used successfully to define the material law of the near surface layer. The previously experimentally demonstrated accommodation of the near surface layer can be

better understood by its quasi-perfect plasticity under hydrostatic pressure, which can create new ground for the explanation of the friction processes, especially in the case of brittle coatings and materials.

To analyse the influence of the hydrostatic pressure on the deformability, it is enough to investigate only axi-symmetric specimens, which are characterised by a Lode-parameter $\mu = 1$, because the experiences show that increasing of the hydrostatic pressure, the stress state dependent limit curves close to each other.

6. CONCLUSIONS

Using the classical theory of plasticity, we defined a material law for the near surface layer of the first body. This material law depends on temperature, strain rate and on actual stress state. Furthermore, it takes the deformability into consideration too, which depends on the hydrostatic component of the actual stress state.

Using ductile unalloyed steel and a brittle, power metallurgy aluminium alloy, the dependence of the flow curves and the deformability were also investigated experimentally.

ACKNOWLEDGEMENTS

The authors would like to express their thanks to the Hungarian Research Found (OTKA T 029963, T 025208) for sponsoring the above-mentioned researches, as well as, to the CNRS (reference number HJBCU45995) and the Hungarian Academy of Sciences for the support of mobility.

REFERENCES

1. A. Eleőd, F. Oucherif, J. Devecz, Y. Berthier, Conception of numerical and experimental tools for study the TTS. "Lubrication at the frontier, Dowson, D. and al., Amsterdam, Elsevier, 1999, vol. 36, p.673-682.
2. D.A. Rigney, Microstructural Evolution during Sliding, Procs. ASM Sympos. on Wear of Engineerng Materials, ed. J. Hawk, ASM Intl., Indianapolis, IN, Sept., 1997, p.3-12.
3. D.A. Rigney, Transfer, Mixing and Associated Chemical and Mechanical Processes during the Sliding of Ductile Material, ed. D.A. Rigney, based on an Intl. Sympos. in Hyderabad, India, Wear 245(2000)1-9.
4. Y. Berthier, Background on friction and Wear, in Lemaitre Handbook of Materials, ISBN 0-12-443341-3 Academic press, p.676-699.
5. Th. Kármán, Festigkeitsversuche unter allseitig. Druck, Zeitschrift VDI, 1911, 55, 1749-57.
6. A. Leyland and A. Matthews, On the significance of the H/E ratio in wear control, Wear 246 (2000) p.1-11.
7. Johnson, K.L.: Contact mechanics and the wear of metals. Wear 190 (1995), p:162-170.
8. G. Fritzsch, R. Siegel, Kalt- und Warmfliess-kurven von Baustählen. Zentralinst. für Fertigungstechn. des Maschinenbaues, Karl-Marx-Stadt, Inform. für Betriebe des Maschinenbaues, 68.
9. K. Lange, (ed): Handbook of Metal Forming. McGraw-Hill Book Comp. New York etc. 1985.
10. A.F. Maier, Einfluß des Spannungszustandes auf das Formänderungsvermögen der metallischer Werkstoffe. VDI Verlag GmbH, Berlin, 1935.
11. H. Steiger, Über die Abhängigkeit des Formänderungsvermögens metall. Stoffe vom Spannungszust. Dissertation, TH Aachen, 1965.
12. A.A. Bogatov, O.I. Mizhirickij, V.F. Sisminsev, Ju.A. Aksenov, Fizika metallov I metallovedenie, 45(1978) 5. p.1089-1094.
13. W. Lode, Versuche über den Einfluß der mittleren Hauptspannung aus dem Fließen der Metalle. Zeitschrift für Physik 36(1926) 11/12 p.913-939.
14. Gy. Ziaja, Z. Darvas, Die Abhängigkeit der Umformvermögens Metallischer Werkstoffe von Formänderungszustand, Metal Forming Theory and Practice, Proc. 5th Int. Conf. on Metal Forming, ICMF'91. Győr, 19-21 June 1991. Ed. M. Tisza, GTE Miskolc, p. 40-47.
15. P.W. Bridgman, Studies in Large Plastic Flow and Fracture, McGraw-Hill, London, 1952.
16. H.Li.D. Pugh, Mechanical Behaviour of Materials Under Pressure, Elsevier, Amsterdam-London-New York, 1970.
17. D. Kuhlmann-Wilsdorf, Effects of Local High Pressures and Shears at Tribological Contact Spots, Tribology Research, G. Dalmaz et al. (eds), 2001 Elsevier Science B.V.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Thickness dependency of the lubricity of Ag thin films deposited on Si(111)7×7 surface

M. GOTO [a], T. YAMAMOTO [a]*, and F. HONDA [a]

[a] Toyota Technological Institute, Hisakata 2-1-12, Nagoya 468, Japan

Abstract

Tribological properties of Ag films on Si (111) 7×7 surface having a wide range of thicknesses have been studied using a pin-on-plate type tribometer under ultrahigh vacuum (UHV) conditions. The coefficient of friction of thicker Ag films and Si (111) Ag-$\sqrt{3}\times\sqrt{3}$ ordered surface show approximately 0.1, whereas the coefficient of friction of 0.4 - 8.4 nm thick Ag films is one order lower than that obtained from both the thicker Ag films and the Ag-$\sqrt{3}$ surface. Auger electron spectroscopy (AES) measurements show that the Ag MNN intensity ratio of the worn track surfaces to non-rubbed surfaces changes depending on the Ag thickness, and sub-nanometer thick films show the highest ratio. Reflection high-energy electron diffraction (RHEED) patterns show that the Ag film deposited on Si (111) 7×7 surface at room temperature forms a two-dimensional fiber crystal structure and gradually changes to a single crystal Ag film as the deposition progresses, whereas the films annealed at 500℃ form a $\sqrt{3}\times\sqrt{3}$-R30° super-structure.

The mechanism of the lubricity of the Ag film in the nanometer-thickness region is discussed in terms of the crystallographic structure of Ag depending on the film thickness.

1. INTRODUCTION

Thin soft metallic lubricant layers such as gold, lead and silver (Ag) are common materials for solid lubricants. Such materials are used in aerospace mechanical components to obtain lubricity in ultrahigh vacuum environments. The film thickness effect for friction force was reported by Bowden and Tabor with indium thin films on tool steel [1]. The optimum thicknesses to obtain the lowest friction force varied, depending on the surface roughness of the substrate for ion-plated lead, indium and Ag layer deposited on SAE 52100 steel substrates which had various surface roughness from 25 nm to 700 nm Rrms [2]. Talivaldis et al. [3] reported the optimum thickness for ion-plated gold film deposited on SUS440 substrate with a surface roughness of 20 nm Ra. They found that the coefficient of friction increased as film thickness increased above the optimum thickness at which the lowest coefficient of friction was obtained. The stated reason for such film thickness effect was an increase in real contact area as film thickness increases.

On the other hand, a thinner lubricant layer range from sub-nanometre to sub-micrometer will be needed if the machine components become even smaller, and micro machining more common. In thinner soft metallic lubricant layers, it is anticipated that the material properties of ultra-thin films differ from those of bulk properties, i.e., the nano size effect might affect the frictional performance. For example, the coefficient of friction, durability of the film and influence of the atmosphere might be different from those of a thicker lubricant layer. However, studies concerning such thin solid lubricant layers are very few.

Recently, we reported that sub nanometer-thick Ag films deposited on Si (111) surface showed good lubricity under ultrahigh vacuum (UHV) environment [4, 5]. In these cases, Si wafer with a very smooth surface was selected as the substrate, and the counter surface was finely polished diamond, which is not reactive with Ag. As a result, adhesive force and displacement taking place in the Ag layer could be minimized, and the coefficient of friction decreased to the 0.01 range. The film thickness effect might be apparent in such an ultra-thin region;

*Present address: KOYO SEIKO CO., LTD. 24 - 1, Kokubu Higanji - cho, Kashiwara - shi, OSAKA 582 - 8588, Japan

however, the effect in the nanometer thickness region is still unclear.

The objective in this paper is to clarify the film thickness effect of Ag films in the range from 0.4 nm to approximately 170 nm under limited conditions due to the asperities in the contact area and the atmospheres. The relationship between the tribological properties and nano structure of Ag is the special focus of this study.

2. EXPERIMENTAL

2.1. Apparatus

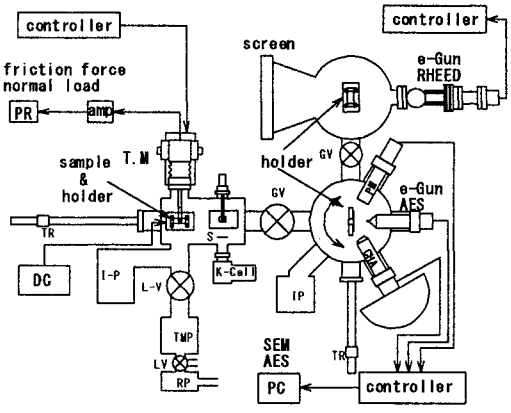


Figure 1. Schematic diagram of the experimental apparatus

The experimental apparatus used in this study is schematically shown in Figure 1. The apparatus consisted of three UHV chambers; a reflection high-energy electron diffraction (RHEED) chamber, Auger electron spectroscopy (AES) chamber and frictional experiment chamber connected in line. The basic pressure of the RHEED and AES chamber was less than 1×10^{-8} Pa, and that of the frictional experiment chamber was less than 4×10^{-8} Pa. The evacuating system was composed of turbomolecular, ion-sputter and titanium sublimation pumps to avoid oil fume contamination on surfaces. Ag films were deposited on Si (111) 7×7 surface using a water-cooled pyroritic boron nitride (PBN) crucible with temperature control, and supplied for the frictional experiments without exposure to atmosphere. A laboratory-built sliding test apparatus was mounted in the experimental chamber. A detailed description of the sliding test apparatus was presented elsewhere [5]. A Si (111) substrate was fixed rigidly on the sample stage of the experimental chamber, and the frictional experiments were carried out by sliding the pin against the fixed substrate.

2.2. Testing specimens

Si (111) substrates were cut from an n-type wafer with a resistivity of 1-3 Ωcm to $18\times4\times0.6$ mm^3. Hardness was 7.4 GPa by micro Vickers hardness measurement. The direction $<211>$ corresponds to the sliding direction, and was confirmed by RHEED diffraction pattern. The surface roughness of Si substrates was 0.68 nm root-mean-square (Rrms) by coherent laser roughness measurement (WYCO, TOPO-3D) for an area of 200×200 μm^2. The single-crystal diamond pin, which was polished hemispherically with a curvature radius of 3 mm, and fixed on top of the SUS304 holder. Roughness of the diamond surface was 9.7 nm for an area 200×200 μm^2. The crystal orientation of both the diamond and the Si sliding plane was (111), and that of the sliding direction was parallel to <211> axis, which was ascertained by a back-reflection Laue spot.

2.3. Experimental procedure

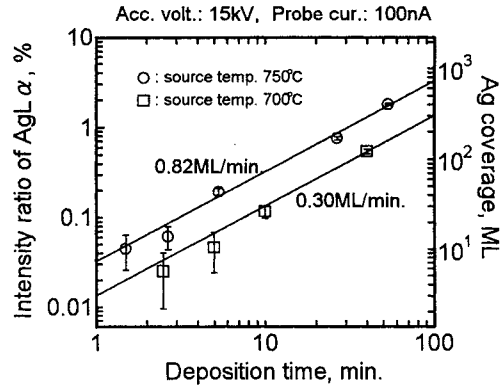
Specimens were ultrasonically cleaned in both acetone and methanol several times, before mounting the Si substrate and the pin in the UHV chamber. The diamond pin was cleaned with dilute nitric acid, to eliminate heavy metallic contaminations before ultrasonic cleaning. After the cleaning, the substrate was attached on the transfer rod in the frictional experiment chamber, and the diamond pin was also attached on the load sensor of the sliding test apparatus. Subsequently, the experimental chamber with specimens was baked out at a temperature of 150℃. The pressure reached 10^{-8} Pa at room temperature after the baking, and the Si substrate was transferred to the RHEED chamber.

The oxide layer on the substrate surface was removed by iterated flashing under 10^{-8} Pa vacuum, at around 1200 ℃ by a direct current through the specimen, just before the Ag deposition [6, 7]. After flashing, a clear 7×7 RHEED pattern indicating clean Si (111) surface was observed, then the Si substrate was cooled to room temperature.

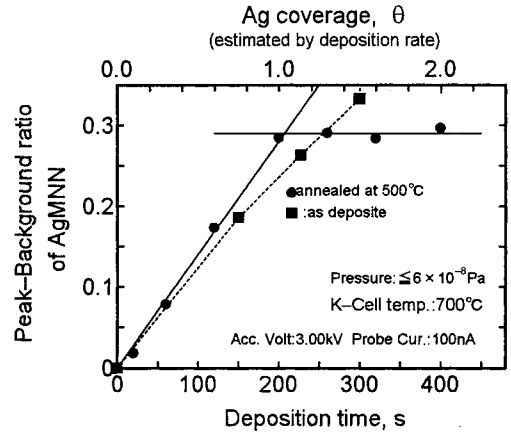
Ag films were deposited on Si (111) 7 × 7 surfaces using a water-cooled Knudsen Cell at source temperatures of 700 or 750℃. The Ag film thickness was defined here as the product of the deposition time and the deposition rate, as measured by electron probe microanalysis (EPMA). Figure 2 (a) shows the relation between deposition times and relative intensity of AgLα, which reflects the good linearity with the deposition rate. In the sub-nanometer region, the reliability of the film thickness was confirmed by a condensation curve of Ag, which was obtained by AES measurements. Figure 2 (b) shows the condensation curve of Ag on the Si (111) surface at the crucible temperature of 700 ℃. The Ag MNN intensities were measured from the annealed surface on which Ag was deposited. The Ag MNN intensity ratio of peak to background was saturated at a deposition time of 200 s, which corresponded to 1 monolayer (ML) thickness, estimated by the deposition rate. By controlling the evaporation time at a crucible temperature of 700 and 750℃, a given thickness of Ag was deposited at a pressure less than 4×10^{-8} Pa. Prior to frictional experiments, the structures of deposited Ag films were observed by RHEED again, with accelerating voltage of 15 kV and an incident angle of approximately 4° .

The frictional experiments in reciprocal sliding were carried out under normal load condition of 250 mN and sliding speed of 0.1 mm/s in the UHV at a pressure of 10^{-8} Pa. The number of reciprocal sliding cycles was limited to 100, and the stroke span of the sliding track was 1.0 mm for all experiments. During sliding, the friction force and normal load were continuously recorded. The friction force is defined here as the average maximum force as a reciprocal sliding cycle, and the coefficient of friction was defined by the friction force over the average normal load for each cycle.

After the frictional experiments, the Si substrate with Ag film less than 4 nm thick was transferred to the AES chamber to perform AES analysis on a worn surface. The peak-background ratio of Ag MNN Auger electron intensity on the worn surface compared with the unworn surface on the substrates was obtained by *in-situ* AES with an acceleration voltage of 5 kV and a probe current of 100 nA. The locations of worn surfaces on Ag films were identified by scanning electron microscopy (SEM) after frictional experiments.



(a). Deposition time vs. intensity ratio of AgLα



(b). Condensation curve of Ag on the Si (111) surface

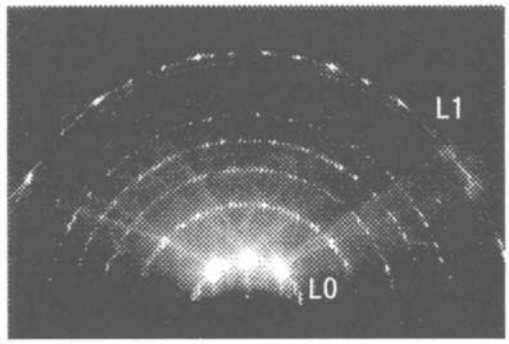
Figure 2. Ag deposition rate measurements b EPMA and AES

3. RESULTS AND DISCUSSION

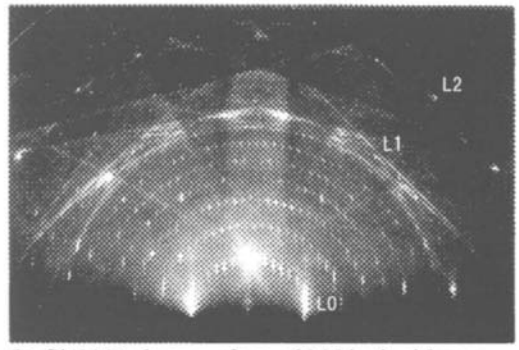
3.1. Ag film structure

Figure 3 (a) and (b) show the RHEED patterns of Si (111) clean surface, and (c) ~ (h) indicate the patterns of 0.4, 1.6 and 8.4 nm Ag films. The patterns ((a), (c), (e) and (g)) shown in the left column are obtained by <110>azimuthal incidence, and the patterns in the right column are obtained again by <211>azimuthal incidence.

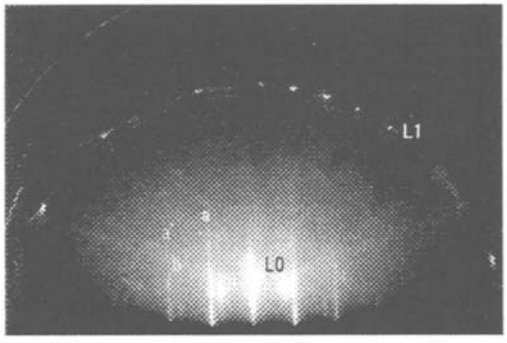
In Fig. 3 (c), in the pattern of 0.4 nm Ag film, streaks "a" and "b" were observed, which originated from the Ag fine crystals on the Si (111) 7×7



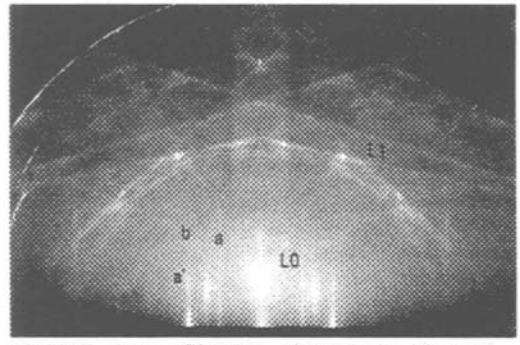
(a) Si(111) clean surface <110> incidence



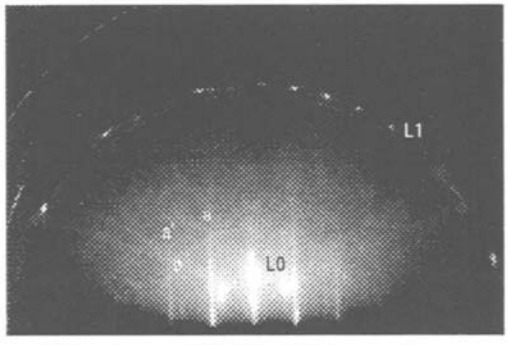
(b) Si(111) clean surface <211> incidence



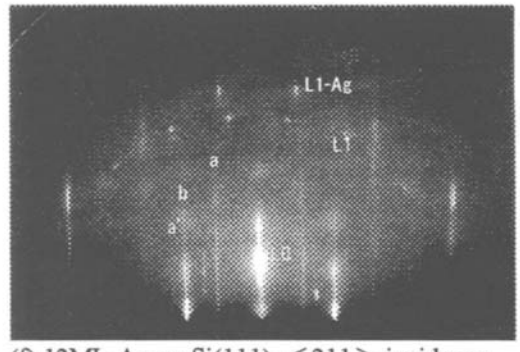
(c) 3ML-Ag on Si(111) <110> incidence



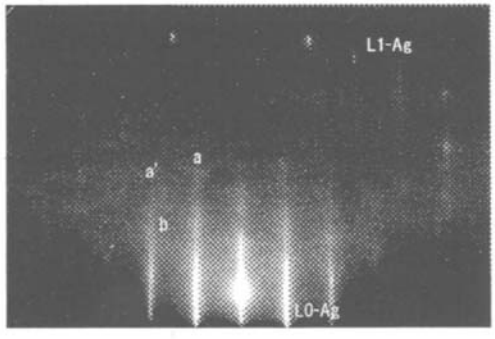
(d) 3ML-Ag on Si(111) <211> incidence



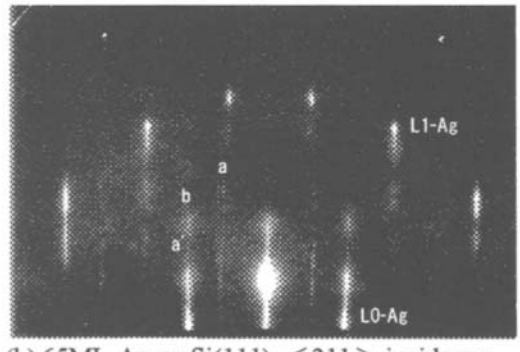
(e) 12ML-Ag on Si(111) <110> incidence



(f) 12ML-Ag on Si(111) <211> incidence



(g) 65ML-Ag on Si(111) <110> incidence



(h) 65ML-Ag on Si(111) <211> incidence

Fig. 3. RHEED patterns of Ag films on Si(111) surface

surface. The streak "a" in Fig. 3 (c) is a diffraction pattern from epitaxally grown Ag crystals, and streak "b" is that from Ag crystals of which the < 211 > axis is parallel to the < 110 > direction of the substrate. The intensity of streak "a" is grater than that of "b". If the substrate is rotated in the azimuthal orientation, streak "a" and "b" are always observed, but the relative intensity relation between "a" and "b" is changed. In contrast, if the incidence of electron beam is changed to the < 211 > direction of the substrate, the intensity of "b", is greater than that of "a", as shown in Figure 3 (d). This shows that the Ag film consists of fine crystals that have a (111) texture structure in which they show a preferred orientation of $<211>_{Ag} \parallel <211>_{Si}$ on $(111)_{Ag} \parallel (111)_{Si}$ [8, 9]. As the film thickness increases, the streak b gradually vanishes in Figs. 3 (c), (e) and (g), and becomes stronger in Fig. 3 (d), (f) and (h). This indicates that the orientations of the Ag crystals on the surface of the film gradually change to a single direction with film growth.

The streak "b" shown in Figs. 3 (d), (f) and (h) change to the transmission pattern of Ag crystals with increasing film thickness. This indicates that the morphology of the film becomes rougher, and the change suggests that crystal growth of Ag occurs. From the crystallographic point of view, crystal growth of the fine crystals is expected so that the average Ag grain size will increase as the growth process progresses. The average size of an Ag crystal is, however, difficult to estimate by RHEED patterns. Another approach, scanning tunnelling microscopy, should be used to estimate the average size.

3.2. Film thickness effect of Ag films in nanometre region

Figure 4 indicates the average coefficient of friction from start to 100 reciprocations, as a function of Ag film thickness in the range of 0 to 170 nm. The average coefficient of friction in the nano region was approximately 0.02, and the coefficient of friction increased as film thickness increased. The coefficient of friction of the thicker Ag films ranged 0.06 – 0.11.

This is similar to the result of Arnell et al. [2], but the thickness range, in which the lowest coefficient of friction was obtained, was two orders thinner than their results. In our case, film penetration and/or degradation caused by protrusions on the contact surface is limited in the sub-nano region of the thickness, since the roughness of the substrate was in the sub-nano meter range and that of the pin was in the nano meter range. The Ag MNN intensity ratio of the worn surface to the non-slid surface still remained at 0.9 - 0.6 after 100 reciprocal slidings with a film thickness of more than 0.4 nm, as shown in Figure 5. This indicates that the sliding plane was on the film/pin interface and/or on film, and that a smaller surface roughness does not affect it.

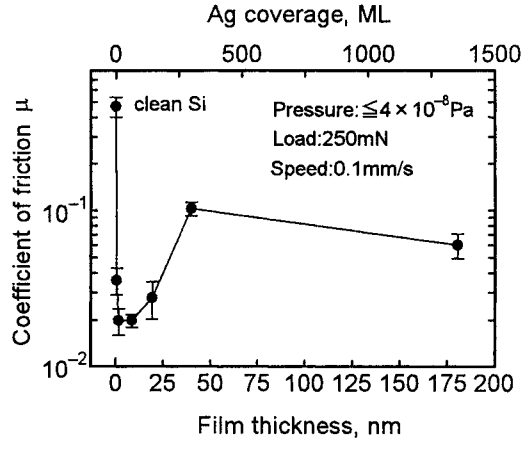


Figure 4. Average coefficient of friction as a function of Ag film thickness in the range of 0 to 170 nm

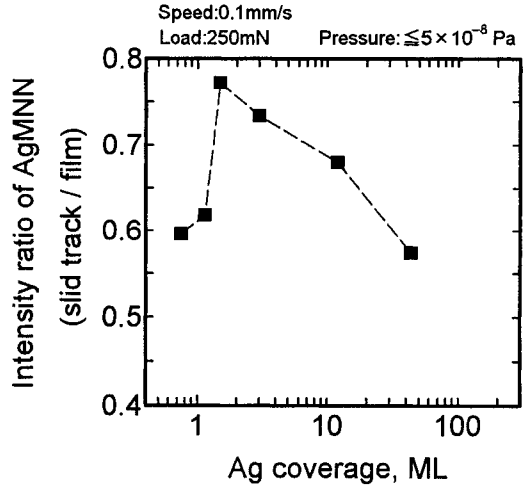


Figure 5. Ag MNN intensity ratio of the worn surface to the non-slid surface

In a reported study of the film thickness effect, the coefficient of friction gradually increased with increasing the film thickness, and this behavior was explained by the fact that the load-carrying capacity of the substrate decreased and there was an increase in the real contact area [3]. In this consideration, the material properties of the film were still constant, because the film thickness was in the bulk region. In this study, however, the coefficient of friction changed greatly between the thickness range of 8.4 nm and 39.5 nm. The gradient of the coefficient of friction versus the Ag film thickness was obviously steep compared to the conventional film thickness effect. To understand this anomalous phenomenon of nanometer layers acting as a solid lubricant, one must take in to account how the change in film material properties such as hardness, and yield strength affect the nano range. The details are discussed in the next section.

3.3. Nano-structure effect of Ag thin film on the tribological properties

Material properties of fine particles change according to the dimensions on an atomic scale and in the bulk region. In the case of Au fine particles, the melting temperature drops abruptly with a diameter less than 5 nm, which was studied by a high-energy electron diffraction technique [10]. The

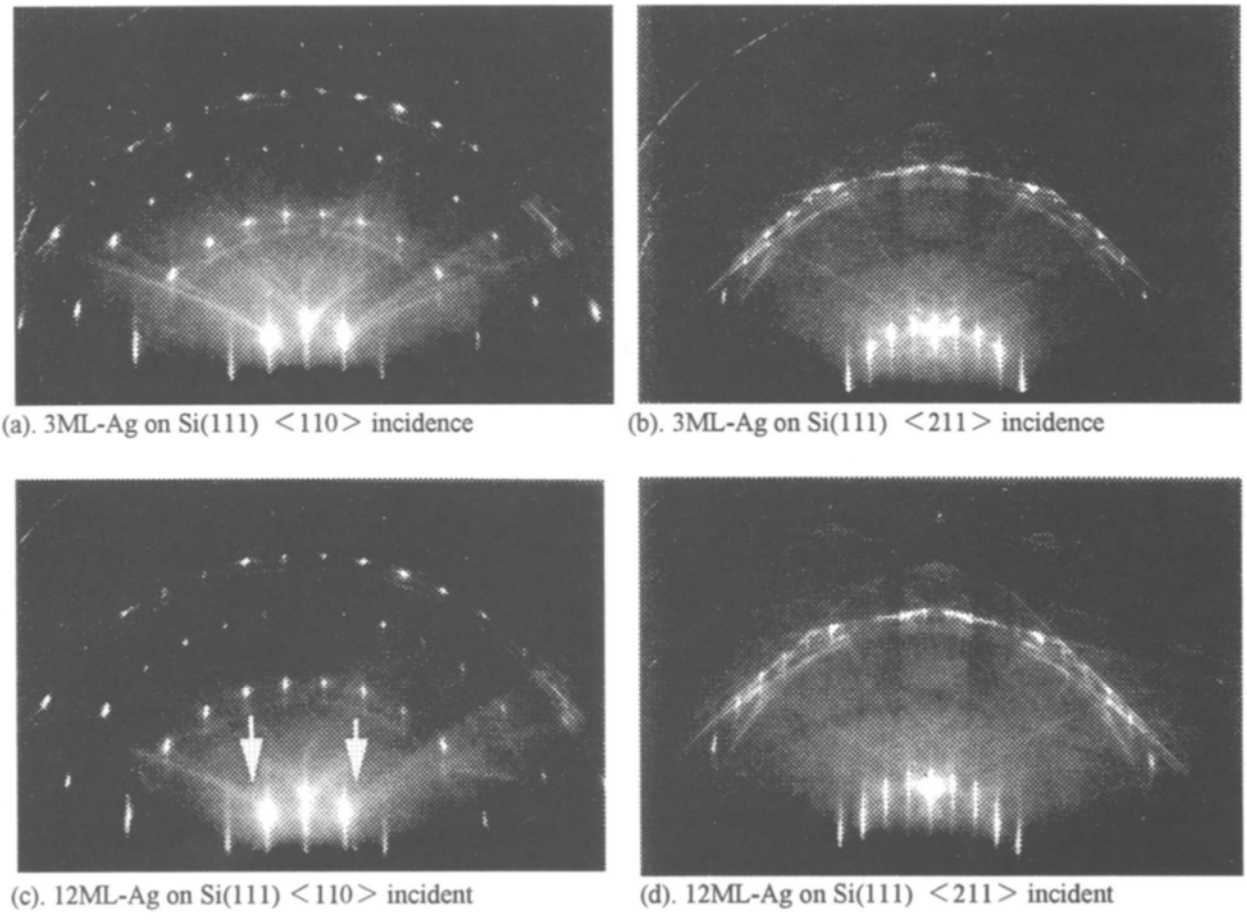


Figure 6. RHEED patterns of annealed Ag films
The diffraction streaks c from Ag crystal are observed on 1.6 nm Ag film (white arrow in (c)).

melting temperature decreases to 600 K at a diameter of 2 nm. The same phenomena have been reported for Pb and Ag [11, 12]. This melting point decrease may induce lower hardness of the material than bulk hardness at room temperature.

The Ag film deposited on the Si (111) 7×7 surface starts growth from its fine crystal structure, then gradually changes to epitaxial growth with increasing thickness, as discussed in 3.1. In Figure 4, the coefficient of friction increases at film thickness greater than 8.4 nm. This is within the same range as the observed melting point drop of Au fine particles [10]. As the melting point of Ag nano-film might be decreased, so the hardness of the film might be changed in each step of the initial growth process of the Ag film. The softening of the Ag film hardness in nano scale might be responsible for the low friction phenomena of Ag nano films.

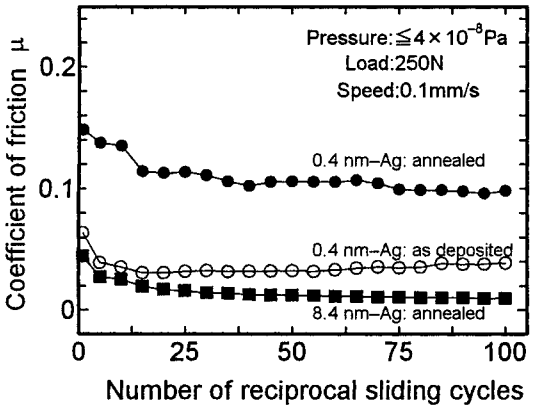


Figure 7. Results of frictional experiments on 3 types of surface; 0.4 nm annealed Ag film, 0.4 nm as-deposited Ag film, and 1.6 nm annealed Ag film

On the other hand, a Ag film on Si (111) surface is changed to $\sqrt{3} \times \sqrt{3}$ domain and Ag islands by annealing at approximately 500℃ [13]. On a $\sqrt{3} \times \sqrt{3}$ surface, Ag atoms of 1 ML are fixed on Si atoms having long scale periodicity, whereas the remaining Ag atoms form islands that have a relatively large diameter compared to as-deposited Ag fine crystals. The RHEED patterns of 0.4 and 1.6 nm Ag films after 500℃ annealing are shown in Figure 6 (a), (b) (c) and (d), respectively. The patterns in the left column are obtained in a $< 110 >$ azimuthal incidence, and the patterns shown in right column are also obtained in a $<211>$ azimuthal incidence. The diffraction spots originating from only $\sqrt{3} \times \sqrt{3}$ structure are observed on 0.4 nm Ag film ((a) and (b)), whereas the diffraction streaks c from Ag crystal are observed on 1.6 nm Ag film ((c) and (d)) with weak intensity. This means that a relatively large area of $\sqrt{3} \times \sqrt{3}$ exists on annealed 0.4 nm Ag film compared to Ag island regions, and the domain of Ag islands was increased on 1.6 nm Ag film as shown in Fig. 6 (c). Figure 7 shows the results of frictional experiments on 3 types of surface; 0.4 nm annealed Ag film, 0.4 nm as-deposited Ag film, and 1.6 nm annealed Ag film. By annealing at 500℃, the coefficient of friction increased from 0.02 to 0.1 on 0.4 nm Ag film, whereas a one-order lower coefficient of friction of 0.01 was obtained on 1.6 nm annealed Ag film. The high coefficient of friction was attributed to sliding on $\sqrt{3} \times \sqrt{3}$ surface that have long range periodicity, whereas the low coefficient of friction was caused by deformation of Ag fine crystals with short range periodicity. However, the average size of these crystals were different between 0.4 nm as-deposited Ag film and 1.6 nm annealed Ag film, as shown in Figure3(c) and Figure 6(c). Detailed investigation concerning mechanical properties such as hardness and shear strength is needed to clarify the tribological properties in a nano region.

4. CONCLUSIONS

We have investigated the tribological properties of Ag films ranging from 0.4 to 170 nm on a Si (111) 7×7 surface using a pin-on-plate type tribotester under the pressure of 10^{-8} Pa. We demonstrate that the film thickness effect in the nano region of Ag films was different from conventional behaviour of thicker regions. The coefficient of friction drops abruptly as the film thickness of Ag decreases from 39.5 to 8.4 nm. The coefficient of friction of nano-range thick Ag films shows a value approximately one-order of magnitude smaller than for thicker Ag films. This phenomenon was discussed in terms of the Ag film structure depending on the film thickness.

ACKNOWLEDGEMENT

The authors express their gratitude to Hino Motors, Ltd. for financial support. We also thank Professor Ino of the Utunomiya University for his constructive

suggestions and the RHEED techniques. We also acknowledged Professor Ichikawa of Meiji University for useful discussion about RHEED analysis.

REFERENCE

[1] F. P. Bowden and D. Tabor, The Friction and Lubrication of Solids, Oxford University Press, London, 1964.
[2] R. D. Arnell and F. A. Soliman, Thin Solid Films, 53 (1978) 333-341.
[3] T. Spalvins and B. Buzek, Thin Solid Films, 84 (1981) 267-272.
[4] M. Goto, F. Honda, and M. Uemura, Proc. 25th Leeds-Lyon Symp. Trib. (1998) P313
[5] M. Goto, F. Honda, and M. Uemura, WEAR, Vol. 252, No. 9-10 (2002) 777-786.
[6] E. J. van Loenen, J. E. Demuth, R. M. Tromp, and R. J. Hamers, Phys. Rev. Lett. Vol. 58, No. 4 (1987) 373.
[7] J. Kraft, S. L. Surnev, and F. P. Netzer, Surf. Sci. 340 (1995) 36-48.
[8] Y. Gotoh and S. Ino, Jpn. J. Appl. Phys. Vol. 17, No. 12, (1978) 2097-2109.
[9] Y. Gotoh and S. Ino, Thin Solid Films, 109 (1983) 255-261.
[10] Ph. Buffat and J-P. Borel, Phys. Rev. A, Vol. 13, No. 6, (1976) 2287-2298
[11] M. Takagi, J. Phys. Soc. Jpn. 9, (1954) 359.
[12] N. T. Gladkich, R. Niedermayer, and K. Spiegel, Phys. Status Solidi 15, (1966) 181.
[13] G Le Lay, Surf. Sci. 132 (1983) 169.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Published by Elsevier B.V.

Sliding wear behaviour of nanocrystalline Cr_3C_2-25(Ni20Cr) coatings at elevated temperature

Manish Roy, Andreas Pauschitz and Friedrich Franek

Institute for Micro Technique and Precision Engineering, Tribology and Engineering Design Section
Vienna University of Technology, A 1040 Vienna, Austria

Cermet coatings such as WC-Co, Cr_3C_2-NiCr obtained by thermal spraying are well known for their tribological application. The strength and toughness and in turn the wear resistance of these coatings can be increased by depositing them in nanocrystalline form. Thus the objective of the present investigation is to evaluate the wear performances of the Cr_3C_2-25(Ni20Cr) coating having a nanocrystalline grain size. Towards that purpose, nanocrystalline Cr_3C_2-25(Ni20Cr) powder was deposited on steel substrate by High Velocity Oxy Fuel (HVOF) spraying. The micro-structural features and mechanical properties of the coating were determined. The sliding wear behaviour of the said coating was evaluated as a function of temperature with the help of an elevated temperature wear test rig. The morphology of the worn surfaces and the regions beneath the worn surfaces were examined using SEM. The results of this study show that the friction coefficient of the coating decreased with increasing test temperature up to 873 K. The study also identified three different wear mechanisms.

1. INTRODUCTION

Cermet coatings obtained by thermal spraying are well known for their tribological application. Among these cermet variety of coatings, WC-Co and NiCr-Cr_3C_2 layers find application in large variety of industrial usages [1-3]. However, WC-Co coatings decarburises to W_2C, W_3C or even to elemental W on exposure to high temperature [4]. Such decarburisation degrades the performance of this coating and this leads to application of this coating limited to 450°C. In contrast, NiCr-Cr_3C_2 coating exhibits improved performances even up to a temperature of 900°C [5] by forming Cr_2O_3 scale on oxidation. The hardness and hence the wear performance of NiCr-Cr_3C_2 coating can be improved further by depositing the layer with nanocrystalline grains. The nanocrystalline grains of Ni base coatings exhibits excellent thermal stability against grain growth [6] up to 1200 K. Nano-crystalline Cr_3C_2-25(Ni20Cr) coating also exhibits formation of fine Cr_2O_3 precipitates [7] which helps to retain coating's hardness and should retard the grain growth. Further, this coating forms Cr_2O_3 scale on oxidation and the kinetics of the oxidation for nanocrystalline coatings are altered during oxidation giving better protection [8,9]. Thus Cr_3C_2-25(Ni20Cr) coating with nanocrystalline grains are expected to exhibit improved wear performances particularly at elevated temperature.

The tribological behaviour of Cr_3C_2-25(Ni20Cr) coating has been studied by several investigator [10-15]. Li and Lai [10] noted that at elevated temperature the friction and wear of this coating is initially high due to a stick-slip phenomenon and as the sliding distance increases these values decrease. Further, friction coefficient increases with decrease of sliding speed. An extensive work on room temperature sliding wear of this coating by Mohanty *et al.* [11] indicated decohesion between splats, controls the wear rate. Wear rate is proportional to applied load but the coefficient of friction decreases with increase of sliding velocity. Matoes *et al.* [12] reported the presence of two distinct zones on the wear vs. sliding distance plot. Other investigations have been primarily concerned either with abrasive wear [13] or with the elevated temperature erosion of the coating [14,15]. Hence, room temperature wear of Cr_3C_2-25(Ni20Cr) coating with conventional grain size has been studied for certain conditions.

No work was found on the wear behaviour of Cr_3C_2-25(Ni20Cr) coating having nanocrystalline grain size. In view of the above a comprehensive work was undertaken to evaluate the wear of Cr_3C_2-25(Ni20Cr) coating with nanostructure and with conventional grain at room and elevated temperature. The present work which is part of the comprehensive investigation, concerns with the elevated temperature wear behaviour of Cr_3C_2-25(Ni20Cr) coating having nanocrystalline grain size compared to the wear behaviour of a Cr_3C_2-25(NI20Cr) coating having a conventional grain size.

2. EXPERIMENTAL DETAILS

2.1. Materials

Cr_3C_2-25(Ni20Cr) powders were used for this study. The powder with nanocrystalline grains was procured from WOKA GmbH, Germany. The powder was a 75 wt% Cr_3C_2 and 25 wt% Ni-20Cr alloy. This powder was manufactured by agglomeration through spray drying followed by plasma densification. The substrate material used in this work was a special variety of stainless steel, 253 MA. The mating material was 100Cr6 steel, which is commonly used as a bearing material.

2.2. Thermal spraying

Thermal Spraying of the coating was carried out using a TAFA HVOF System. A JP-5000 HVOF system was employed for that purpose. The system is equipped with a gun, a control console, a powder feeder and an X-Y manipulator. The gun had a nozzle (100 - 200 mm long and 6.4 mm diameter) and a combustion chamber at right angles to the nozzle.

A schematic representation of the gun is made in Fig. 1. Oxygen and a fuel gas kerosene were combusted under high pressure in the chamber and the combustion product was allowed to exit into the atmosphere. A pilot flame was used to ignite the main flame. The powder was fed radially into the hot stream of gases. Air was used as a carrier gas to transport the powder from the feeder to the gun. The powders were heated in the gas stream and then got deposited on the substrate at a very high speed. The spraying parameters are summarised in Table 1. The surface of all the specimens were grit blasted using # 13 alumina grit and were ultrasonically cleaned in acetone prior to spraying.

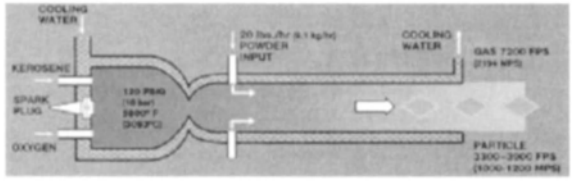
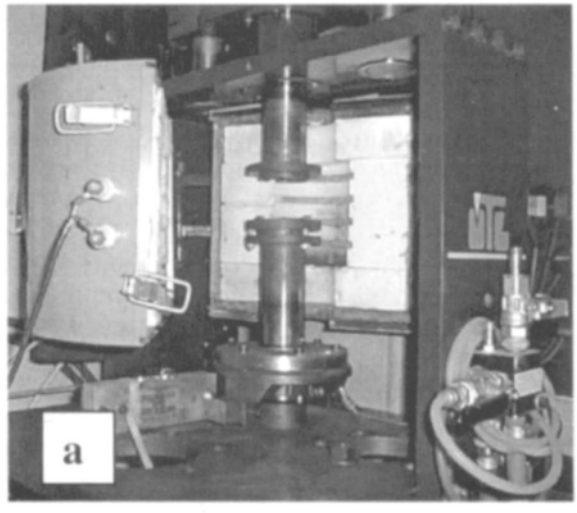


Figure 1. The schematic representation of the HVOF system.



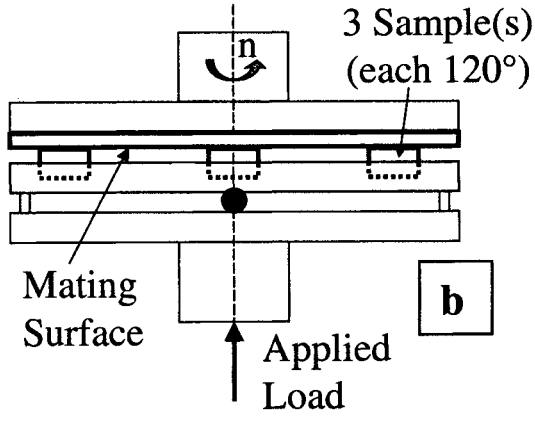


Figure 2. a) High temperature wear test rig; b) loading situation.

Table 1
The HVOF spraying parameters

Parameters	Liquid Fuel	Gaseous Fuel	Fuel Flow Rate	Fuel Flow Pressure	Powder Carrier Gas	Gas Flow Rate	Gas Pressure	Powder Feed Rate	Spray Distance
Values/ names	Kerosene	Oxygen	23 m^3/h	22 - 25 bar	Air	1 m^3/h	7 bar	2 Kg/h	260 mm

2.3. Wear testing

Wear test is carried out using the wear test rig manufactured by the Austrian Tribology Society as shown in Fig. 2. The rig is essentially a disc on disc configuration and modified for the present work as pin on disc configuration with 3 pins. It consists of a driving system whose speed is controlled by a frequency converter. The load is applied by a dead weight kinematic system the test chamber is an electrically heated furnace which is powered by a variable transformer and the temperature is monitored by a thermo couple. Wear is measured by a Linear Variable Displacement Transducer (LVDT). The friction force and hence the friction coefficient is measured by strain gauges. The output of the LVDT and the strain gauges are connected to a computer with a data logging software (HBM Catman 3.0 Rel4). The wear test conditions are listed in Table 2.

2.4. Characterisation of the coating and the worn surfaces

The coated samples were sectioned and polished and then observed under Scanning Electron Microscope (SEM) to characterise the micro-structure. The worn samples were also examined with SEM and an Energy Dispersive Spectrometer (EDS) to see the morphology of the worn surfaces. The worn samples were sectioned using a low speed saw and polished. The polished transverse section of the worn surfaces was seen in SEM again to understand the mechanism of material loss during wear.

3. RESULTS AND DISCUSSION

3.1. Microstructure of the coating

Low magnification micrographs of the sectioned surface of the coating are shown in Fig.3. The coating was 250 μm thick and free from cracks. A striated structure indicating that coating is built up by depositing splat after splat. Such a structure is characteristic of HVOF sprayed coatings. The high magnification Back-Scattered Electron (BSE) micrographs of the coating is shown in Fig. 4. It is similar to HVOF sprayed coating with conventional grains; in that there are three distinct zones. The first zone is dark and represents primarily orthorhombic Cr_3C_2 phase. The second zone is grey in appearance and contains mixture of orthorhombic Cr_3C_2 phase and cubic NiCr phase. The third zone was white and contains essentially Face Centred Cubic (FCC) NiCr phase. Detailed micro structural study of this coating with nanocrystalline grains has been given elsewhere [16] but a representative micrograph observed by Transmission Electron Microscopy (TEM) and a corresponding diffraction pattern is shown in Fig. 5. It is evident from Fig. 5 that the grains of the coating are around 80 to 100 nm.

The X-ray Diffraction (XRD) pattern from the as received powder and the coated surface is shown in Fig. 6. The change of phases due to spraying could not be traced. This is similar to what was observed in case of the Cr_3C_2-25(Ni20Cr) coating with conventional grains [17]. However, reports of other investigators [11] suggests that the Cr_3C_2 phase transform to Cr_7C_3 or $Cr_{23}C_6$. Comparison of the diffraction patterns from coatings with nanocrystalline grains and that with conventional grains shows that powders with nanocrystalline grains undergo amorphisation to a much less extend. It is worthwhile to mention that powders with nanocrystalline grains undergo amorphisation to a greater extend when sprayed with a OSU SIS gun compared to when sprayed with JP-5000 HVOF system [16].

Table 2
The wear test conditions

Parameter/materials	Pin	Mating material	Normal load	Sliding speed	Test temperature
Values/names	NiCr-Cr_3C_2 coating on 253 MA alloy	100Cr6	70 N	0.25 m/s	Room temperature to 1073 K

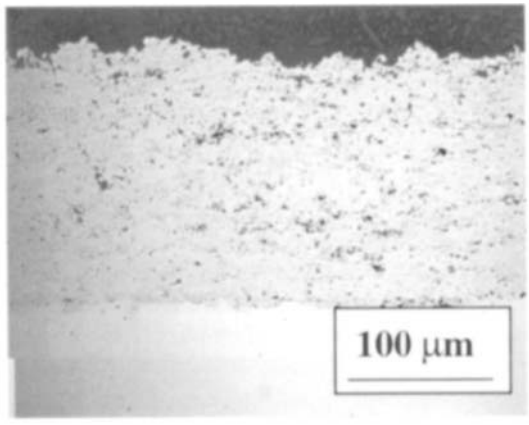


Figure 3. The optical micrograph of the sectioned surface of the coating with nanocrystalline grains.

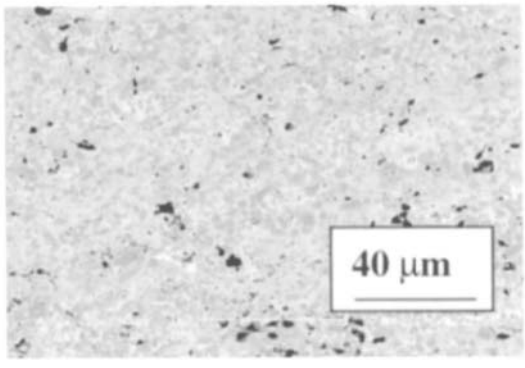


Figure 4. The BSE image of the sectioned surface of the coating with nanocrystalline grains.

3.2. The wear behaviour

The influence of test temperature on the friction coefficient vs. sliding distance plot is presented in Fig. 7. The friction coefficient increases from an initial low value to the steady state value irrespective of the test temperature. This increase in the initial stage is quite slow at ambient temperature and is rather sharp at elevated temperature. This initial increase is to overcome the initial adhesion between the absorbed contaminated coating layer and the mating steel. The influence of the absorbed contaminated layer on initial adhesion is decreased substantially as the temperature is increased. The friction coefficient decreases with increase of test temperature and attains a minimum at 873 K. Beyond 873 K the friction coefficient again increases. In contrast, the Cr_3C_2-25(Ni20Cr) coating with conventional grains exhibited minimum friction coefficient at 673 K [17]. Further, at all test temperatures the friction coefficient of the coating with nanocrystalline grains was higher than the friction coefficient of the coating with conventional grains. The exception is at 873 K where the reverse trend is followed. It is also to be noted that the friction coefficient of the Cr_3C_2-25(Ni20Cr) coating with nanocrystalline grains is observed to be lower than that of the same coating with conventional grains at ambient temperature when the coating was obtained using OSU-SJS HVOF system.

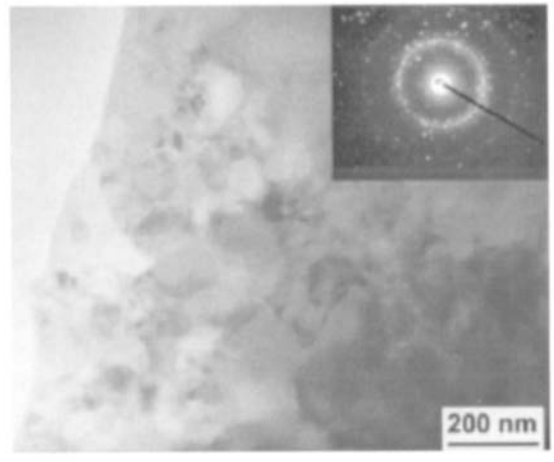


Figure 5. TEM micrograph and corresponding diffraction pattern of a nanocrystalline coating.

The influence of test temperature on the thickness loss is given in Fig. 8. The nanocrystalline coatings show a net thickness loss at ambient temperature and at 473 K. As the temperature is increased, there is net thickness gain due to transfer of material from the mating surface.

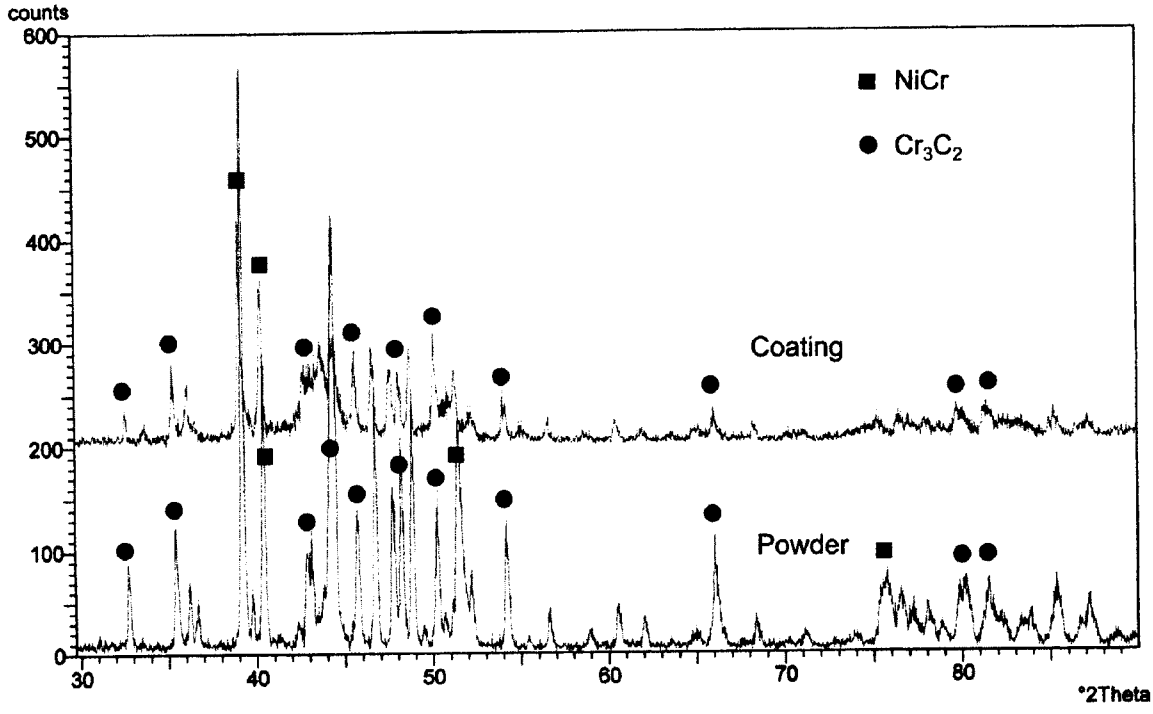


Figure 6. The X-ray diffraction pattern from the powders and the coated surface with nanocrystalline grains.

The thickness gain is maximal at 873 K. In contrast to this, the Cr_3C_2-25(Ni20Cr) coating with conventional grains exhibited maximum thickness gain at 1073 K. It also had thickness gain even at 473 K [17].

3.3. Wear mechanism

The morphologies of the worn surfaces are given in Fig 9. Fig. 9a, Fig. 9b, Fig. 9c and Fig. 9d represent the surfaces worn at ambient temperature 673 K, 873 K and 1073 K respectively. The surface worn at ambient temperature reveals a highly deformed surface. The worn surface appears to be much smoother than the as received coated surface. This also indicates a significant amount of a transfer layer. The surfaces, which are worn at 673 K and 873 K are relatively featureless. There were some scattered white patches where deformed material was transferred from the mating surface. Formation of the smooth surface is also reflected on the friction curves. Fig. 9d indicates an area where a part of the smooth surface was removed by a wearing action.

The EDS spectrums from the worn surfaces shown in Fig. 9 are presented in Fig. 10a to Fig. 10d. At all the four temperatures, the presence of a transferred Fe layer from the mating surface is evident. Quantitative analysis shows that extent of oxidation of the surface layer is maximum at 1073 K It is intermediate at ambient temperature and it is minimum at 673 K. In contrast, the maximum Fe is transferred at 873 K and the transfer was minimum at ambient temperature. This fact indicated that all the Fe, which got transferred does not get oxidised especially at 873 K. The reason for this is that a very high amount of Fe gets transferred at 873 K and there is not enough thermodynamic force to oxidise the whole amount of Fe. On the contrary, at ambient temperature relatively low flash temperatures prevent extensive oxidation of the transferred layer. At high temperature (1073 K) less material was transferred because high flash temperatures caused oxidation.

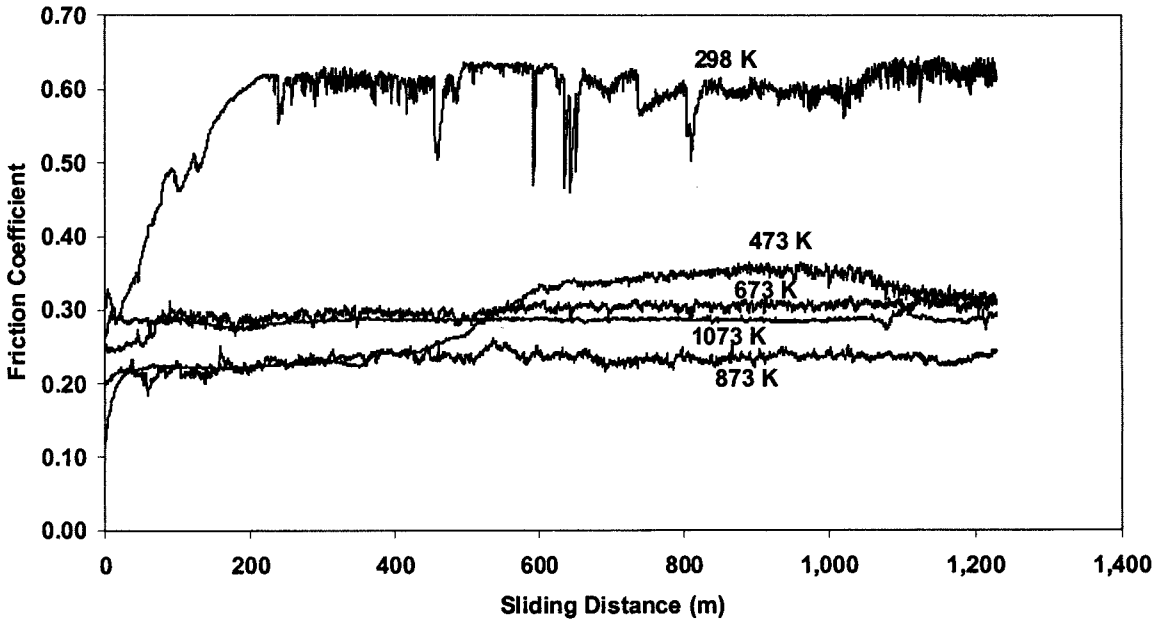


Figure 7. The variation of friction coefficient as function of sliding distance at different temperature for Cr3C2-25(Ni20Cr) coating with nanocrystalline grains.

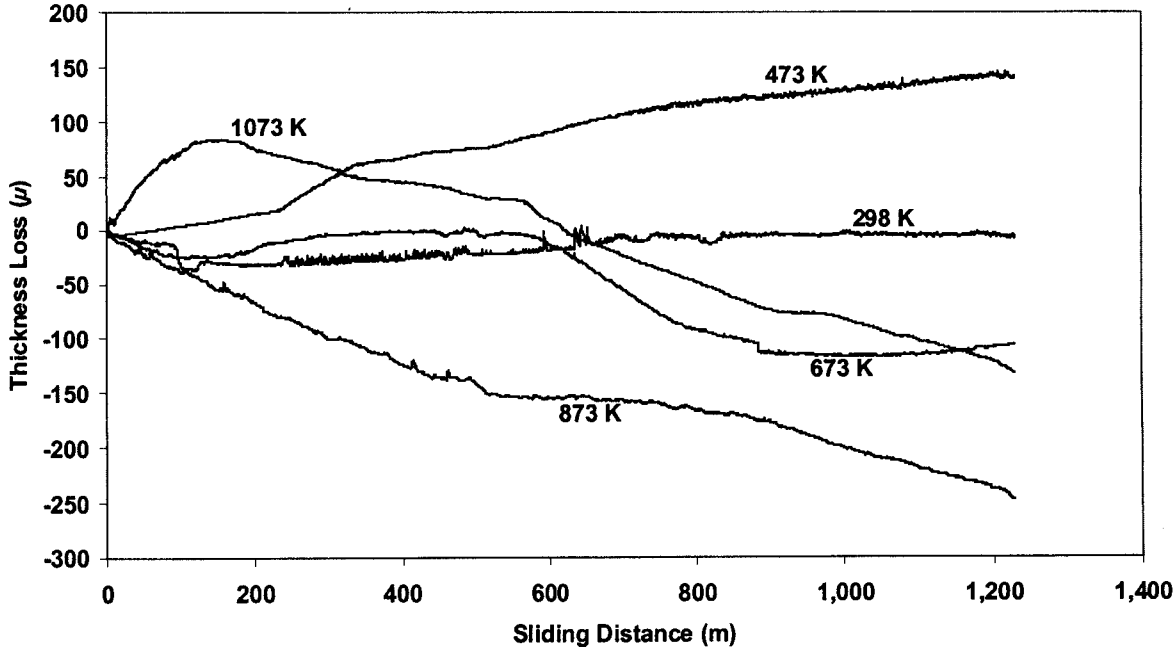


Figure 8. The variation of thickness loss as a function of sliding distance for a Cr_3C_2-25(Ni20Cr) coating with nanocrystalline grains.

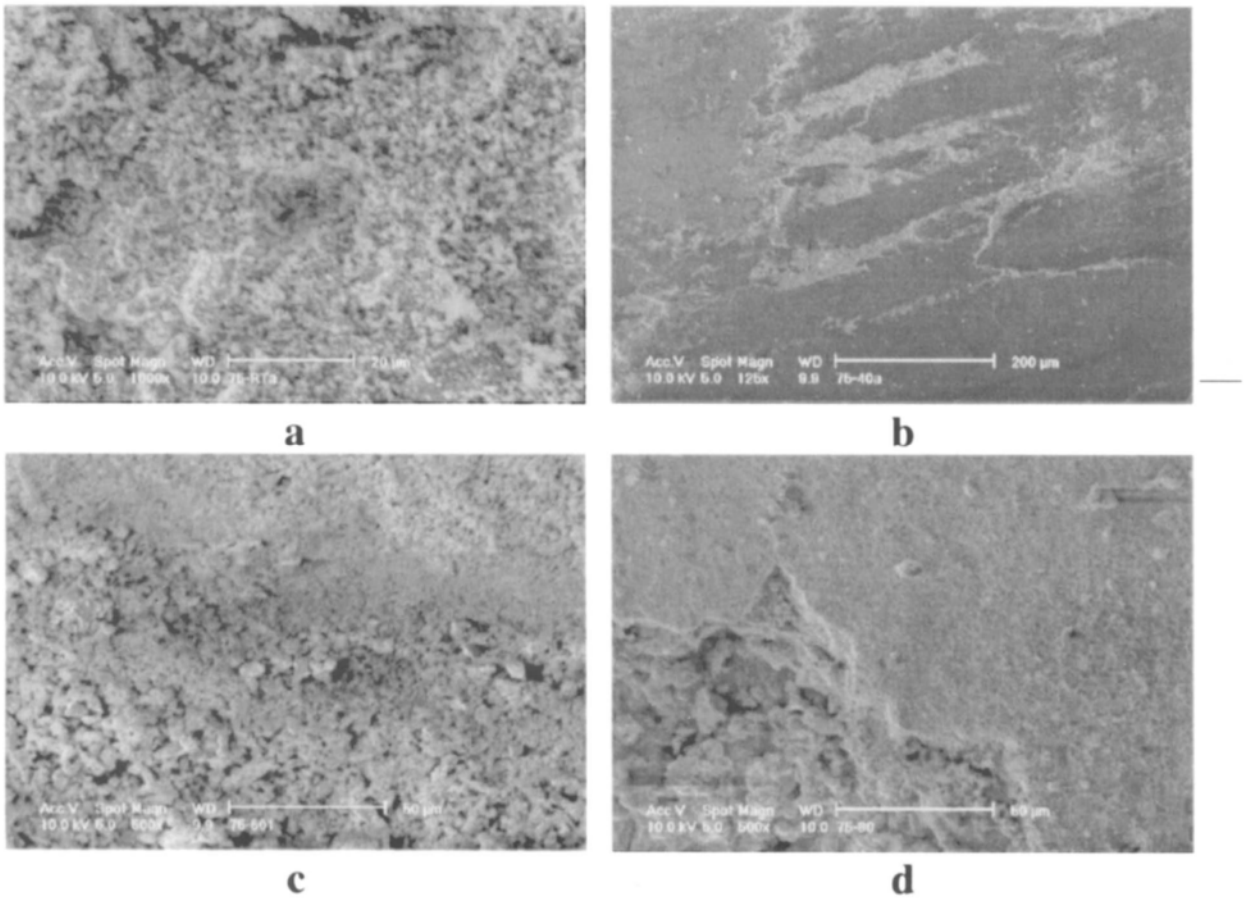


Figure 9. The morphologies of the worn surfaces at a) 298 K, b) 673 K, c) 873 K and d) 1073 K.

Cross sections of the surfaces shown in Fig. 9 are illustrated in Fig 11. Fig. 11a to Fig. 11d represent the transverse section of the surfaces worn at 298 K, 673 K, 873 K and 1073 K. At ambient temperature, the transferred oxidised layer is very thin. Consequently it is able to deform in the same fashion as that of the coating. Thus delamination wear is responsible for material loss. Though no such delamination cracks can be seen on the worn surface, such cracks are evident in the subsurface. As the temperature is increased the thickness of the top layer, which consists of oxides of the transferred Fe, alloy of Fe, is increased. But the presence of composite layer reduces the friction coefficient. This composite layer is not adherent to the substrate and thickens during sliding. Cracks were generated in the subsurface regions particularly at the interface between the composite layer and the coating. This process is very slow and the result is continuous thickness gain of the surface. The cracks at the subsurface, particularly at the interface between the coating and the composite layer, can be clearly seen. At high temperature (1073 K), the composite layer gets sintered and compressed to a greater extent under the action of multi axial compressive stress. Wear takes place from the adherent and compressed layer that forms due to transfer of Fe from counter surface. Fig. 11d shows one such composite layer formed at 1073 K.

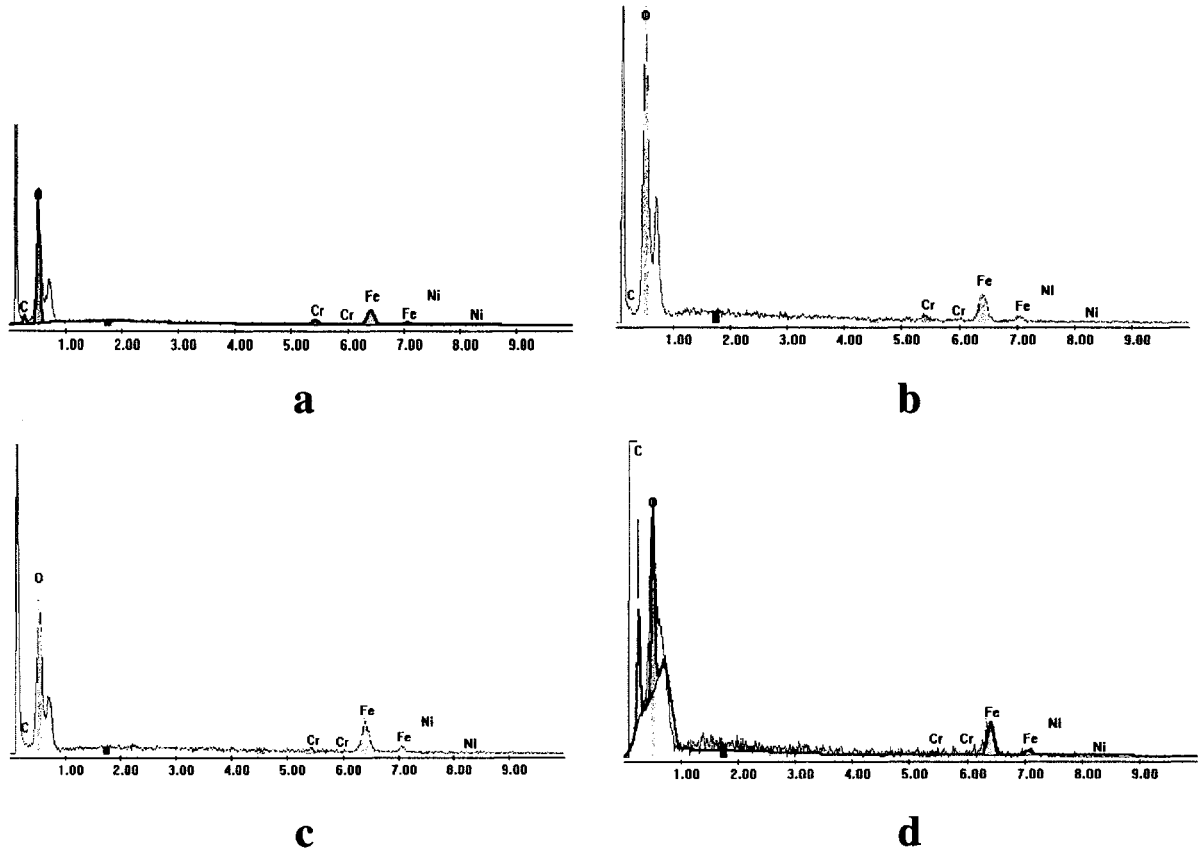


Figure 10. The EDS spectrum from the selected area of the worn surfaces at a) 298 K, b) 673 K, c) 873 K and d) 1073 K.

4. CONCLUSIONS

1) The friction coefficient of the Cr_3C_2-25(Ni20Cr) coating with nanocrystalline grains is minimum at 873 K and maximum at ambient temperature

2) The coating undergoes net thickness loss at 298 K and 473 K and it undergoes net thickness gain at 673 K, 873 K and 1073 K due to transfer of material from the mating surface.

3) There are three different wear mechanisms that prevail at different test conditions of the present investigation.

5. ACKNOWLEDGEMENT

The authors are grateful to Austrian Science Fund (FWF) for supporting this work through their project M 592. The authors also acknowledge the support of WOKA GmbH (Germany) for providing the coating powders and Terolab Services (Austria) for permitting access to the HVOF coating facilities for this work.

REFERENCES

1. L. Moskowitz and K. Trelewicz, J. Thermal Spray Technol., 6 (1997) 294.
2. D.J. Nolan and M. Samandi, J. Thermal Spray Technol, 6 (1997) 422.
3. J.D. Reardon, R. Mignogna and F.N. Longo, Thin Solid Films, 83 (1981) 345.

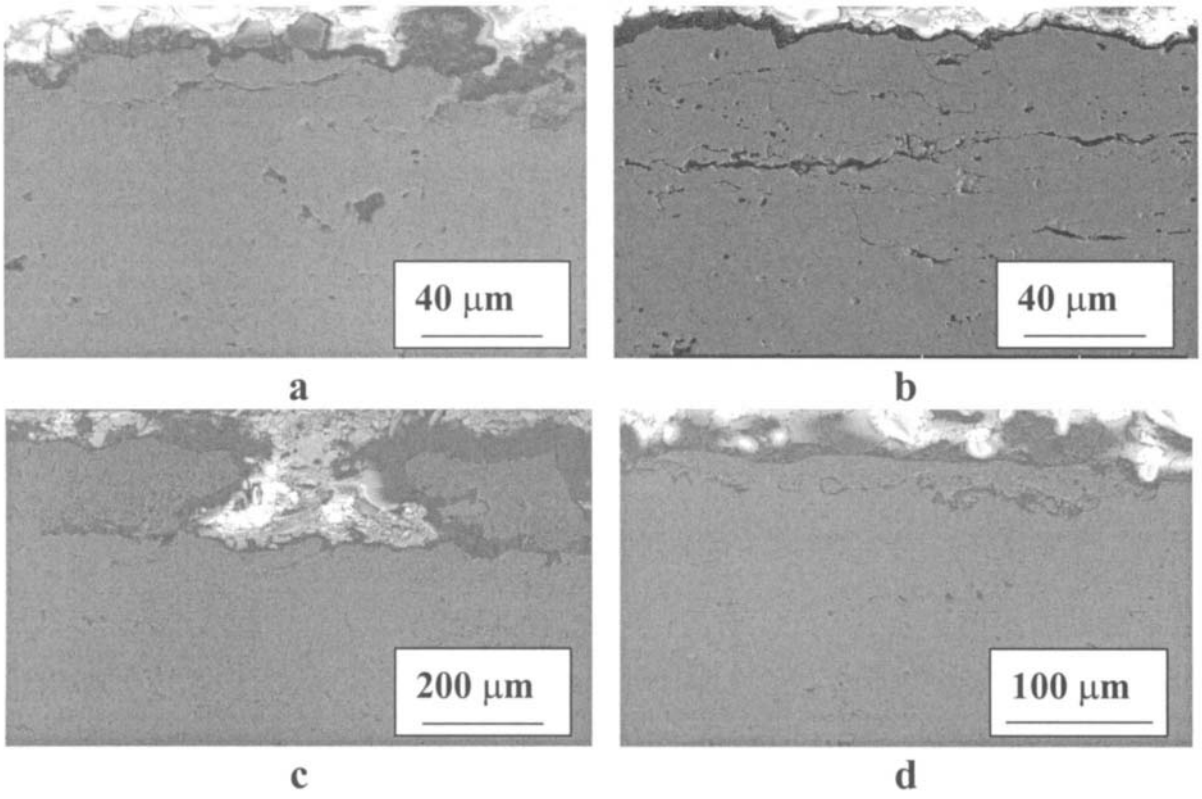


Figure 11. Cross sections through worn surfaces. Surfaces worn at a) 298 K, b) 673 K, c) 873 K and d) 1073 K.

4. L. Russo and M. Dorfmann in Thermal Spraying, Current Status and Future Trends, A. Ohmori, ed., High Temperature Society of Japan, Osaka, Japan, 1995, p 681.
5. O. Knotek, R. Elsing and H.R. Heintz, J. Vac. Sci. Technol., A3 (1985)2490.
6. H.G. Jiang, M.L. Lau and E.J. Lavernia, NanoStruct Materials, 10 (1998) 169.
7. J. He and E.J. Lavernia, Mater. Sci and Engg., A301 (2001) 69.
8. G. Chen and H. Lou, Oxid. of Metals 54 (2000) 155.
9. L. Ajdelsztajn, J. A. Picas, G. E. Kim, F.L. Bastian and J. Schoenung, Mater. Sci and Engg, 338 (2002) 33.
10. C.C. Li and G.Y. Lai, Proc. Int. Conf. On Wear of Materials, Dearborn, Michigan, April 16-18, ASME, Newyork, 1979.
12. J. Matoes, J.M.Cuetoes, R. Vivande and E. Fernandez, tribol. Int. 34 (2001) 345.
13. S. Wirojanupatump. P.H. Shipway, D.G. McCartney, Wear 249 (2001) 829.
14. B.Q. Wang and K. Luer, Wear ,174 (1994) 177.
15. Y. Fukuda and M. Kumon, Thermal Spraying, Current status and future trends, ed. A. Ohmori, High Temperature Society of Japan, Osaka, Japan, 1995, 107.
16. M. Roy, A. Pauschitz, J. Banerdi, T. Koch and F. Franek, Communicated to J. of Material Science.
17. M. Roy A. Pauschitz and F. Franek, Proc. Int. Corr. Cong. Granada Spain, September 22nd to 26th, 2002.

This Page Intentionally Left Blank

SESSION V
TRIBOLOGICAL STUDIES OF DRY AND LUBRICATED SURFACES (1) - DRY

Chairman : Dr A Kapoor

Paper V (i) **Modeling of Variable Friction in Contact with Fretting**

Paper V (ii) **Micro-abrasive Wear of Semi-crystalline Polymers**

Paper V (iii) **Wear Testing of Acetal for the Prediction of Polymer Rolling Bearing Durability**

Paper V (iv) **Self-Lubricating Dry Rubbing Bearings - Fundamentals and Methods of Calculation**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Modeling of variable friction in contact with fretting

M. Cheikh, S. Quilici and G. Cailletaud

Centre des Matériaux P.-M. Fourt, Ecole de Mines de Paris,
UMR CNRS 7633,
BP 87 F-91003 Evry, France.

Abstract In this paper, a new phenomenological model is developed to account variable coefficient of friction (COF) in space and time. The COF is no longer considered as a global value valid for the whole contact area. A local value is introduced instead, which evolution is governed by the local history of the contact and the amount of slip. The framework is inspired from elastoplasticity. The evolution of the COF depends on two variables: an isotropic evolution related to cumulated slip and a kinematic component computed from the actual relative position of the bodies.

1. INTRODUCTION

The fretting is the consequence of low oscillations between two bodies in contact [1]. The cycles can result from small relative displacements or stress variations in the bodies. It is most commonly found in all kinds of press fits, spline connections, leaf springs, riveted and bolted joints, nuclear fuel elements and also in wire ropes. The most significant factor in a contact with fretting is the coefficient of friction [2]. This factor is influenced by the surface quality, the wear induced by the fretting, the behavior of the third body, etc. Many studies have been completed for characterizing friction in fretting tests [2]. In the numerical studies dealing with fretting, the COF is generally presented as a constant of the contact [3, 4]. In fact, the COF varies according to the modification of the surface quality and the lubricant and it is never constant [5]. There is an absence of friction model that can accurately simulate the temporal and spatial variations of the COF during fretting fatigue. In this paper, a new phenomenological model is developed to account variable COF in space and time. The COF is no longer considered as a global value valid for the whole contact area. A local value is introduced instead, which evolution is governed by the local history of the contact and the amount of slip. The framework is inspired from elastoplasticity. The evolution of the COF depends on two variables: an isotropic evolution related to cumulated slip and a kinematic component computed from the actual relative position of the bodies. Other authors have already proposed an elastoplastic framework [6, 7], nevertheless, to the best knowledge of the authors, the combination of an isotropic and a kinematic variable is original.

In the following, section 2 shows typical experimental results for a titanium alloy, and section 3 presents the elementary components of the friction model. Section 4 shows the finite element implementation of the model in the computer code ZéBuloN (see `http://www.mat.ensmp.fr/Produits` or `http://www.nwnumerics.com` for more full informations about this code). The last section presents the fretting cycles obtained by this model compared to those obtained by experiment in the case of cylinder–plan problem.

2. EVOLUTION OF THE COF IN FRETTING TEST

The present study is performed in connection with experiments work made at Ecole Centrale de Lyon, for characterizing the fretting behavior of a titanium alloy. The experimental setup used is presented in [8]. It is composed of a cylinder, subjected to a normal load P, sliding on a plan. The displacement amplitude δ is applied on the plan. In this test, the cylinder is coated and lubrificated, the plan is shot peened without specific coating. Figure (1) shows a typical evolution of the COF during a fretting test with a displacement amplitude of $\delta^* = 50\mu$m.

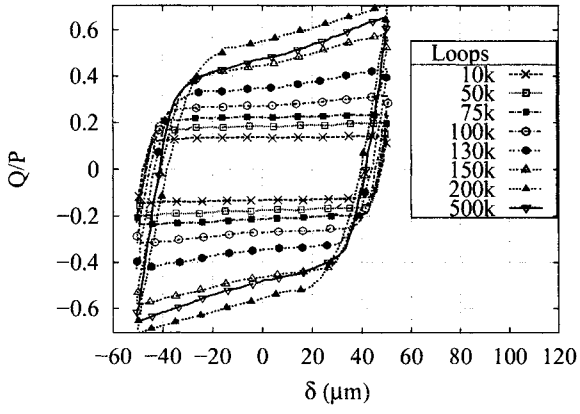


Figure 1. Examples of fretting cycles for $\delta^* = 50\mu$m. The unit k represents 1000 cycles.

Each cycle of fretting represents the evolution of the ratio Q/P according to displacement amplitude δ with Q the tangential force. The COF μ is identified in the phase corresponding to the gross slip. By choosing a law of Coulomb $|Q|/P <= \mu$, the horizontal sides of these cycles represent the evolution of the COF $\mu = |Q|/P$ according to displacement amplitude δ.

It can be noted in this figure a continuous evolution of the COF whatever the direction of slip. It is about an isotropic evolution equivalent to an isotropic hardening in plastic behavior. The isotropic evolution is induced by the increase in the roughness of surfaces of contact with time. An evolution of μ versus slip is also found. A large amount of slip is accompanied by an increase in the COF and the opposite slip is always facilitated compared to the current slip. It is about the kinematic component of the evolution of the COF. This evolution is similar to kinematic hardening for plastic behavior. The kinematic evolution is the result of the degradation of the contact by abrasive wears, by the modification of the form of the contact surface, by the behavior of the third body, etc.

The cycles presented in figure (1) have a parallelogram shape. The first cycles have horizontal sides. With the evolution of the fretting, these sides become oblique. The first cycles can be modeled with a constant μ. What amounts to using a perfect evolution of the COF equivalent to a perfect plastic behavior. The cycles with oblique sides cannot be modeled with a constant COF. To solve this problem, the numerical studies classically model this evolution by carrying out several calculations with different COF [9, 10]. The model presented in this study will allow us to derive a full modeling of the friction history during the whole life of the contact.

3. DESCRIPTION OF THE MODEL

3.1. Modeling of the isotropic COF

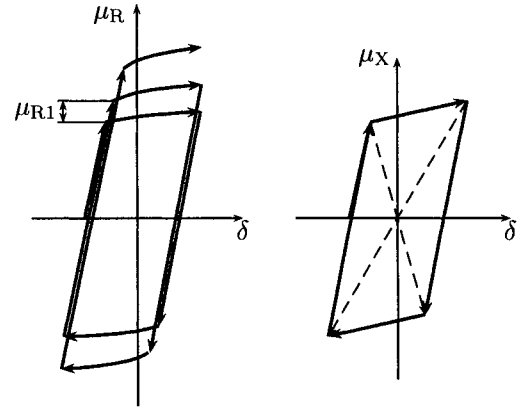


Figure 2. Definition of the kinematic and the isotropic COF.

The isotropic COF μ_R represents the expansion which undergoes the cycle of fretting with time. Between two cycles of fretting, μ_R can be represented by the variation of the initial COF μ_0 between these two cycles (see Fig. 2). The isotropic COF is independent of the slip, it can thus be parameterized by a scalar variable like the cumulated energy of slip $\overline{E}_d$ or the cumulated slip $\overline{\delta}$. In this study we chose to model the isotropic COF μ_R according to the cumulated slip $\mu_R = f(\overline{\delta})$. Figure 3 presents some examples of evolution of μ_R according to the cumulated slip. To model the experimental results of μ, the function f must have a nonlinear form in $\overline{\delta}$ and must have a threshold of saturation. We used the following form:

$$\mu_R = \mu_1(1 - \exp(-b_R\overline{\delta})), \qquad (1)$$

with μ_1 and b_R the parameters of the fretting of used material. The coefficient μ_1 represents the threshold of saturation and b_R a factor of acceleration of the isotropic evolution of the COF. The cumulated slip is defined by

$$\overline{\delta} = \int_0^t |d\delta|. \qquad (2)$$

Figure (3), with the experimental results, presents the μ_R obtained by this model.

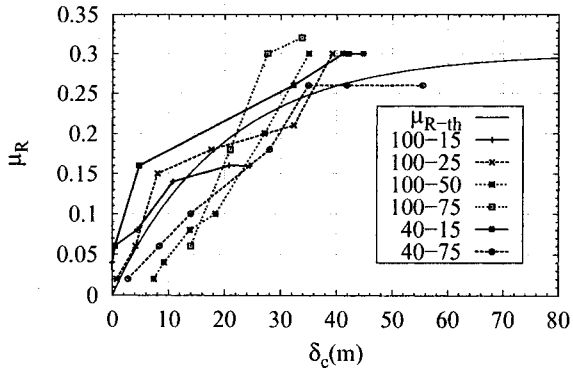


Figure 3. Examples of evolution of μ_R according to the cumulated slip $\delta_c = \overline{\delta}$ obtained by experiment.

3.2. Modeling of the kinematic COF

As mentioned above, the kinematic COF represents the component of the evolution of the COF dependent on the current slip δ. For low values of μ the slip is generally full and the cycles have the shape of a parallelogram. Before the COF does not reach significant values, so that the slip is transformed into partial slip, the kinematic COF is linear in δ hence: $\mu_X = H\delta$. The modulus H represents the slope of slip. This modulus is generally variable and depends on the cumulated slip $\overline{\delta}$ that is $H = g(\overline{\delta})$. The function g represents the evolution of the slope of climbing in the phase of slip. If the modulus H is constant, our model will be similar to the plastic model of Prager.

In the same way that the isotropic COF, and in order to model the experimental results, the function g must have a nonlinear form in $\overline{\delta}$ and must have a threshold of saturation. The evolution of the function g is modeled by:

$$g(\overline{\delta}) = \mu_2'(1 - \exp(-b_X\overline{\delta})^\beta), \qquad (3)$$

with b_X and β the constants of fretting of the selected material. The coefficient μ_2' represents the slope after saturation of evolution. This coefficient is not constant and depends on the maximum displacement amplitude of the test. This difference is explained by the fact that weak displacement amplitude is more favorable to the formation of pad, from where a great opposition to the slip and thereafter a more pointed form of the cycle of fretting. This difference is also explained by the importance of the role of the displacement amplitude in the behavior of the third body in a contact with fretting. The slope with saturation evolves in the form of a reverse function $\mu_2' = \mu_2/\delta_{max}$ (see Fig 4). The kinematic COF takes the following form then:

$$\mu_X = \frac{\mu_2}{\delta_{max}}(1 - \exp(-b_X\overline{\delta})^\beta)\delta. \qquad (4)$$

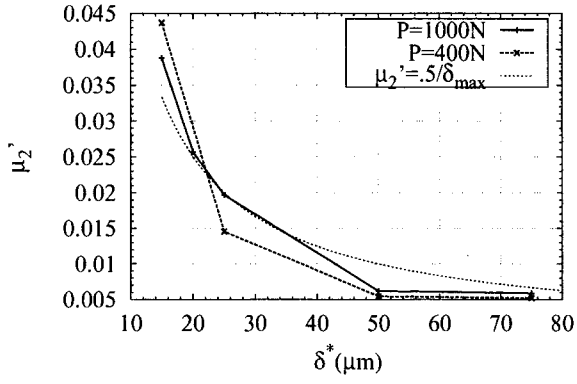


Figure 4. Evolution of the slope of slip after saturation for the loadings P=1000N and P=400N and for theoretical value with μ_2=0.5.

Figure (5) represents some examples of evolution of

$$\mu_{X\max} = \mu_2(1 - \exp(-b_X\bar{\delta})^{\beta}) \qquad (5)$$

obtained by experiment and by the used model.

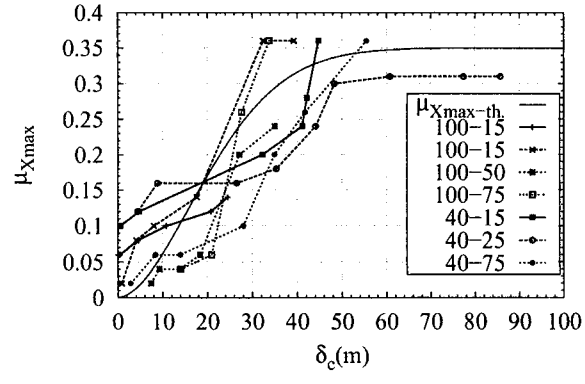


Figure 5. Examples of evolution of $\mu_{X\max}$ according to the cumulated slip $\delta_c = \bar{\delta}$ obtained by experiment.

3.3. Modeling of the total COF

The total COF μ is the sum of the kinematic and isotropic components with the initial COF μ_0:

$$\mu = \mu_0 + \mu_1(1-\exp(-b_R\bar{\delta})) + \frac{\mu_2}{\delta_{\max}}(1-\exp(-b_X\bar{\delta})^{\beta})\delta. \qquad (6)$$

It is a model with 6 coefficients whose values depend on selected material.

With this model, Coulomb law is introduced, so that:

$$|Q| <= \mu P, \qquad (7)$$

$$\text{if } |Q| < \mu P \text{ stick;} \qquad (8)$$

$$\text{if } |Q| = \mu P \text{ slip.} \qquad (9)$$

By choosing the variables $Q_y = \mu_0 P$, $Q_R = \mu_R P$ and $Q_X = \mu_X P$ the law of Coulomb (7) takes the form of a *function of load*:

$$f(Q, Q_X, q_R) = |Q - Q_X| - Q_y - Q_R, \qquad (10)$$

with $f <= 0$ and the conditions (9)become:

$$\text{if } f < 0 \text{ stick;} \qquad (11)$$

$$\text{if } f = 0 \text{ slip.} \qquad (12)$$

The model of the COF (6) is general and can have simpler forms according to the conditions of fretting. We present in the figures (6) and (7) the curves of typical evolution of the COF. Figure (6) presents the case of a coated cylinder and a shot peened plan. In this figure, the evolution of the COF is progressive under the effect of the coating and the lubricant. The progressive damage of the lubricant is accompanied by a progressive increase in the COF. In the case of a cylinder without coating, the COF has a more significant value and almost constant along the test (Fig. 7). In this case there is no isotropic evolution of the COF and the function g (Eq. 3) is constant. Kinematic evolution depends only on the slip δ. The model (6) takes a simpler form then:

$$\mu = H\delta \qquad (13)$$

similar to the linear kinematic model of Prager. The modulus H is constant representing the slope of curry up in the phase of slip in the cycle of fretting (see Fig. 8).

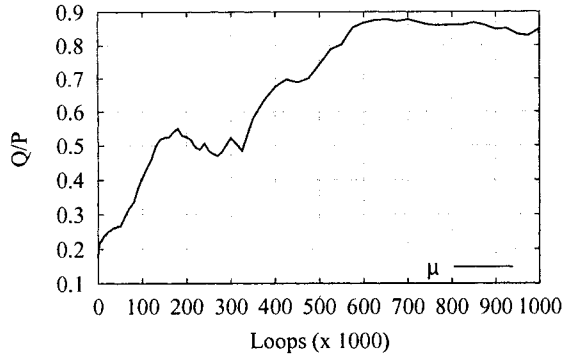


Figure 6. Evolution of the COF for coated and lubricated material.

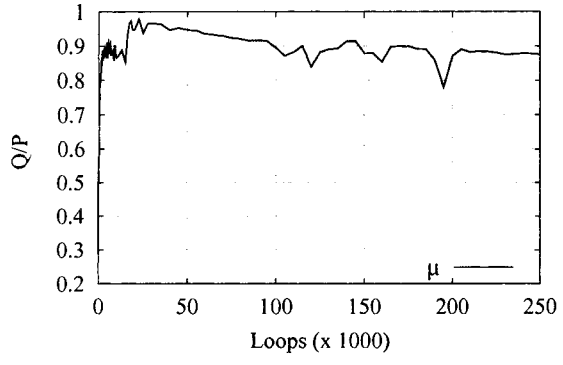


Figure 7. Evolution of the COF for shot peened dry material.

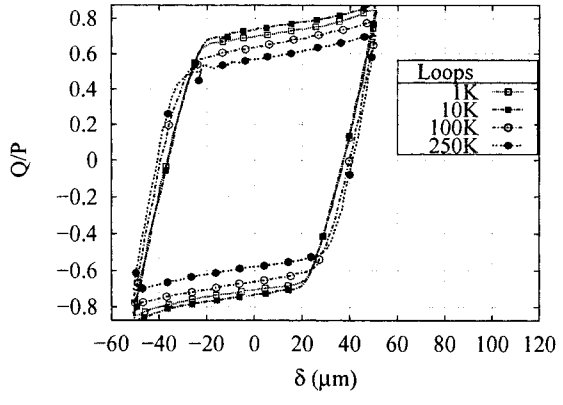


Figure 8. Fretting cycles for the shot peened dry material.

4. NUMERICAL IMPLEMENTATION

This model applies locally on the scale of the contact. In the case of the finite elements for example, each element of contact has its COF dependent on the history, the state of stress, the slip, the state of the contact, etc. We implemented this model with its components in `ZéBuloN` . To model the contact `ZéBuloN` uses the technique of the impactor and the target. The impactor is composed of a set of nodes whereas the target is made up of a set of lines (or a set of faces for the 3D case). An element of contact is composed by a particle (a candidated node of the impactor) and the target. The initial COF μ_0 and the others 5 constants of the model are global characteristics of the zone of contact. Each element of contact inherits (in the meaning of the object programming) the same values of these constants. On the other hand, each element of contact has its cumulated slip $\overline{\delta}$ and its current slip δ. Each element has thus, its kinematic component and its isotropic component of the COF.

4.1. Identification of the model

The identification in fretting of a material is carried out after a great number of cycles which can reach the million. In order to simulate numerically these tests with an acceptable computing time, we use a factor of acceleration of the evolution of the coefficients of the model. In model (6) one replaces the coefficients b_R and b_X by respectively b'_R and b'_X with:

$$b'_R = ab_R \qquad \text{and} \qquad b'_X = ab_X. \tag{14}$$

The coefficient a is the acceleration factor of the model.

A second possible technique to accelerate calculation is the technique of jumps of cycles. Instead of modifying the coefficients b_R and b_X, one increments the value of the cumulated slip periodically. The coefficient a in this technique represents the value of increment of the jump of cycle. Considering the weakness of displacement amplitude in front of the cumulated slip, the two techniques are equivalent. In the applications presented thereafter we used the technique of acceleration.

4.2. The contact algorithm

To deal with the problems of contact ZéBuloN uses a direct method known as method of flexibility [11, 12, 13, 14]. With this method the conditions of contact are not introduced explicitly into the variational formulation of the problem. The reactions of contact are calculated a priori, then added in the equilibrium equations like known additional external forces at the deformed configuration $C^{t+\Delta t}$:

$$R^{t+\Delta t} = F_{\text{int}}^{t+\Delta t} + F_{\text{ext}}^{t+\Delta t} + R_{\text{c}}^{t+\Delta t}, \qquad (15)$$

with R the residue of equilibrium; F_{int} and F_{ext} respectively the internal and external forces.

Displacements u_{c} due to the contact are related to the contact reactions r_{c} on a local frame of the element of contact by the relation

$$u_{\text{c}} = T^{\text{T}} K^{-1} T r_{\text{c}} = W r_{\text{c}} \qquad (16)$$

with $W = T^{\text{T}} K^{-1} T r_{\text{c}}$ the matrix of flexibility condensed on the scale of the contact. The matrix T is a linear operator allowing to express the local variables in the global frame such as $R_{\text{c}} = T r_{\text{c}}$.The system (??) represents the equilibrium system of the structure subjected to the reactions of contact only. In order to define displacements u_{c} and the reactions r_{c} the law of Coulomb ?? and the condition of impenetrability (condition of Signorini) are imposed. The condition of Signorini is expressed by $r_n = \text{Proj}_{\text{R}^+}(r_n - \rho_n x_n)$ where Proj_{R^+} denotes the orthogonal projection on the positive real set $\mathbf{R}^+$; $\rho_n > 0$ and x_n the gap between the particle and the target. The law of Coulomb can be expressed by: $r_t = \text{Proj}_C(r_t - \rho_t \delta)$ where C is the disk $\{r_t \in \mathbf{R}^2 : ||r_t|| \leq \mu r_n\}$ and $\rho_t > 0$ [15]. To define the local unknown factors $(u_{\text{c}}, r_{\text{c}})$ the system is solved

$$\begin{aligned} W r_{\text{c}} &= u_{\text{c}} \\ r_n &= \text{Proj}_{\text{R}+}(r_n - \rho_n x_n), \quad \rho_n > 0 \\ r_t &= \text{Proj}_C(r_t - \rho_t \delta), \quad \rho_t > 0 \\ \delta &= \delta_{\text{lib}} + u_{ct} \\ x_n &= x_{\text{lib}n} + u_{\text{c}n} + x_n^0 \end{aligned} \qquad (17)$$

with $x_{\text{lib}n}$ and δ_{lib} respectively the gap and free slide when no contact reactions are applied and x_n^0 is the initial altitude of the particles. Once definite the locally reactions of contact r_{c} , one calculates the reactions of contact on a global scale $R_c = T r_{\text{c}}$ and solves the system (15).

4.3. Viscoplastic constitutive equations

As law of mechanical behavior we use the viscoplastic model with internal variables of Chaboche [Chaboche, 1977; Lemaitre and Chaboche, 1988]. The strain is partitioned into an elastic and plastic part:

$$\varepsilon = \varepsilon^e + \varepsilon^p. \qquad (18)$$

The elastic domain is defined by means of a von Mises equivalent, introducing both kinematic X and isotopic R hardening, so that :

$$f(\sigma, X, R) = (3/2(\sigma - X)^{\text{dev}} : (\sigma - X)^{\text{dev}})^{\frac{1}{2}} - R - \sigma_{\text{y}} \qquad (19)$$

with σ_{y} the initial yield stress. The viscoplastic flow is governed by the normality rule, so that :

$$\dot{\varepsilon}^p = \dot{p} \frac{\partial f}{\partial \sigma} \qquad (20)$$

with p the cumulated plastic strain. The hardening depends on the state variables α and p :

$$X = \frac{2}{3} C \alpha \quad \text{with} \quad \dot{\alpha} = \dot{\varepsilon}^p - \frac{3}{2} \frac{C}{D} X \dot{p} \qquad (21)$$

and

$$R = Q(1 - \exp(-bp)). \qquad (22)$$

The parameters Q and b define the variation of the isotropic hardening R. The parameters C and D characterize the kinematic hardening X.

5. APPLICATION

In order to validate the suggested model we present thereafter an application on the case of the cyclindre-plan and a comparison with the experimental results of the Ecole Centrale de Lyon.

5.1. Mesh and bondary conditions

The cylinder-plan set is meshed by using Zmaster (the graphical pre- and post-processor of Z-SeT/ZéBuLoN code). Figures 9 and 10 represent this mesh. Considering the small size of the contact area compared to the dimension of

the set, we meshed the half of the cylinder only. To show the influence of the size of the mesh on the developed model we use two meshes (see figure 10): Coarse mesh with an elements of $50\mu m \times 25\mu m$ size on the level of the contact, that is tipically 1/10.2 of the total contact size for the element size. Finer mesh characterized by an elements of $8\mu m \times 4\mu m$ size on the level of the contact, hence tipically 1/63.75 of the total contact size for the element size. The first mesh allows a fast calculation of the cycles of fretting, whereas the second allows, in addition to the simulation of the cycles, a finer calculation of the variables of contact. The first mesh is discritised with 1050 elements 2D with linear interpolation and 761 nodes, hence 1522 DOFs for the problem. For the second mesh, the set is discritised with 4744 elements 2D with linear interpolation and 3796 nodes, hence 7592 DOFs for the problem.

In experiments at the beginning of each test a load P is applied in indentation. The test is composed in continuation by cycles of displacement amplitudes $\pm\delta^*$ applied to the plan (see figure 11). The normal load is applied to the cylinder in the form of an impsosed displacement in order to reproduce the desired loading. This way of imposing the loading is in conformity with the tests of the ECL. In these tests the imposed loading is not controlled and its value can vary during the test.

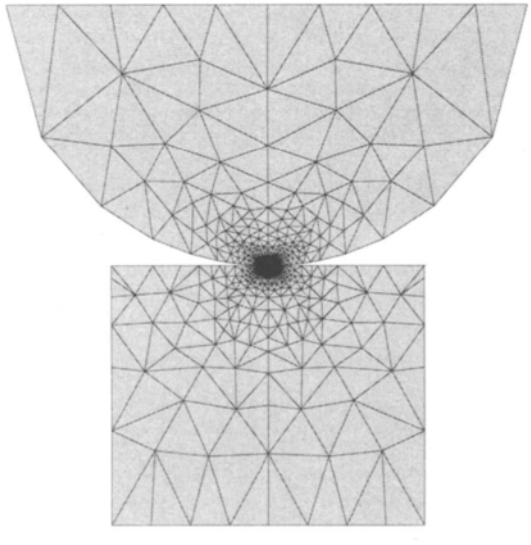


Figure 9. Cyclinder-plan set mesh

The ends of the plan are blocked along the axis y. The displacement amplitudes $\delta(t)$ is applied to the plan along the x axis. The higher extremity of the cylinder is blocked along the x axis.

Due to the small size of the contact area in front of the transverse dimension of the set, the mechanical behavior is supposed to be in plane deformation.

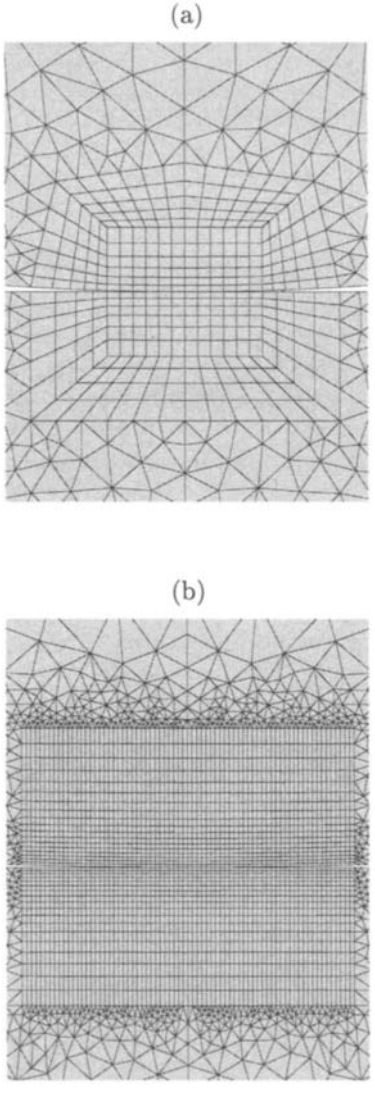


Figure 10. Zoom of the contact zone : a) coarse mesh; b) finer mesh.

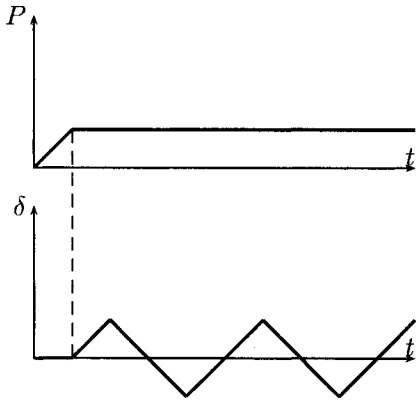


Figure 11. Load and displacement amplitude evolution.

5.2. Results

We present thereafter a sample of fretting test simulations by the suggested model in comparison with the corresponding experimental results. For the loading P=1000N we present the results with displacement amplitude $\delta^* = 20\mu$m in the figure (12). For the loading P=400N we present the results with displacement amplitude $\delta^* = 25\mu$m and $\delta^* = 50\mu$m respectively on the figures (13) and (14). The cycles with a legend LN represent the numerical cycles. The cycles with a legend LE represent the experimental cycles. For the experimental cycles the unit K represents 1000 cycles. In each figure we present the fretting logs and some samples of cycles representing the evolution of the COF according to time.

Figure (12) shows that for the displacement amplitude $\delta^* = 20\mu$m, the slip regime is mixed. The shape of the log is rectangular in the beginning characteristic of gross slip, and at the end of the test the shape of the log becomes elliptic characteristic of partial slip. Cycles presented in this figure show the same evolution for the experimental and simulated results. In this figure, until $325000th$ cycle , the cycles are parallelipidic with more pointed form at the end. Afer this cycle, the cycles become elliptic and the slip regime is partial. For the load P=400N, the slip regime is gross for the displacement amplitude $\delta^* = 25$ and $\delta^* = 50$ as the figures (13) and (14) show it. The fretting logs have parallelipidic shape and the cycles keep a rectangular form along the test.

The whole of the results presented on these figures shows a similarity between the evolution of the numerical cycles and the evolution of the experimental cycles. The small discrepancies between simulation and the experiment are due to the discrepancy between the experimental results of μ_R and μ_X as one notices it on the figures of identification of the coefficients of the model (Fig. 3 and Fig. 5).

Dispersion in the experimental results is due to the problems of reproducibility like noticed by Fridrici in [16]. The discrepancy between simulations and the experimental results generally remains lower then 0.1 along calculations.

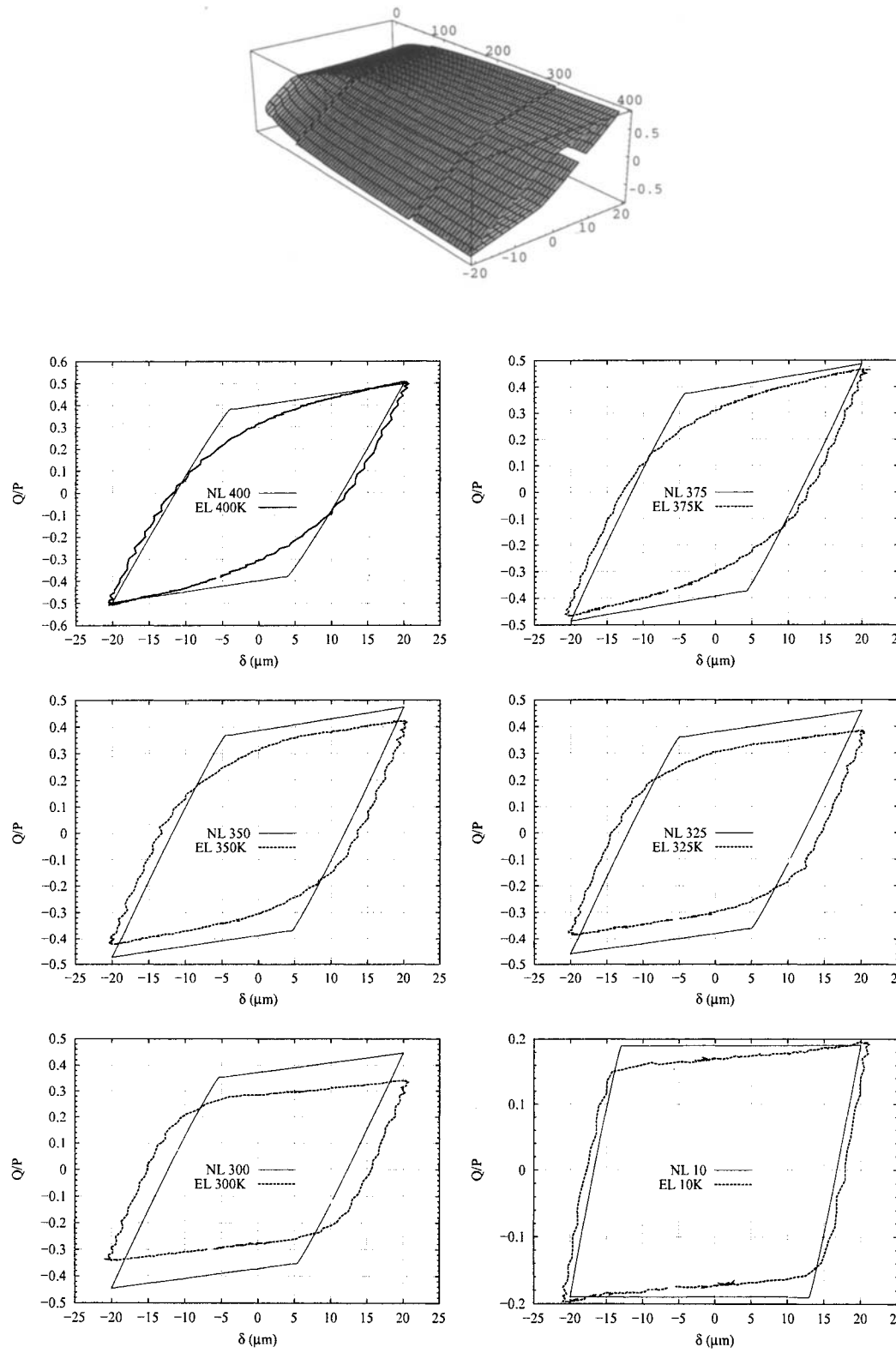


Figure 12: Numerical and experimental cycles evolution for $P = 1000\text{N}$ and $\delta^* = 20$.

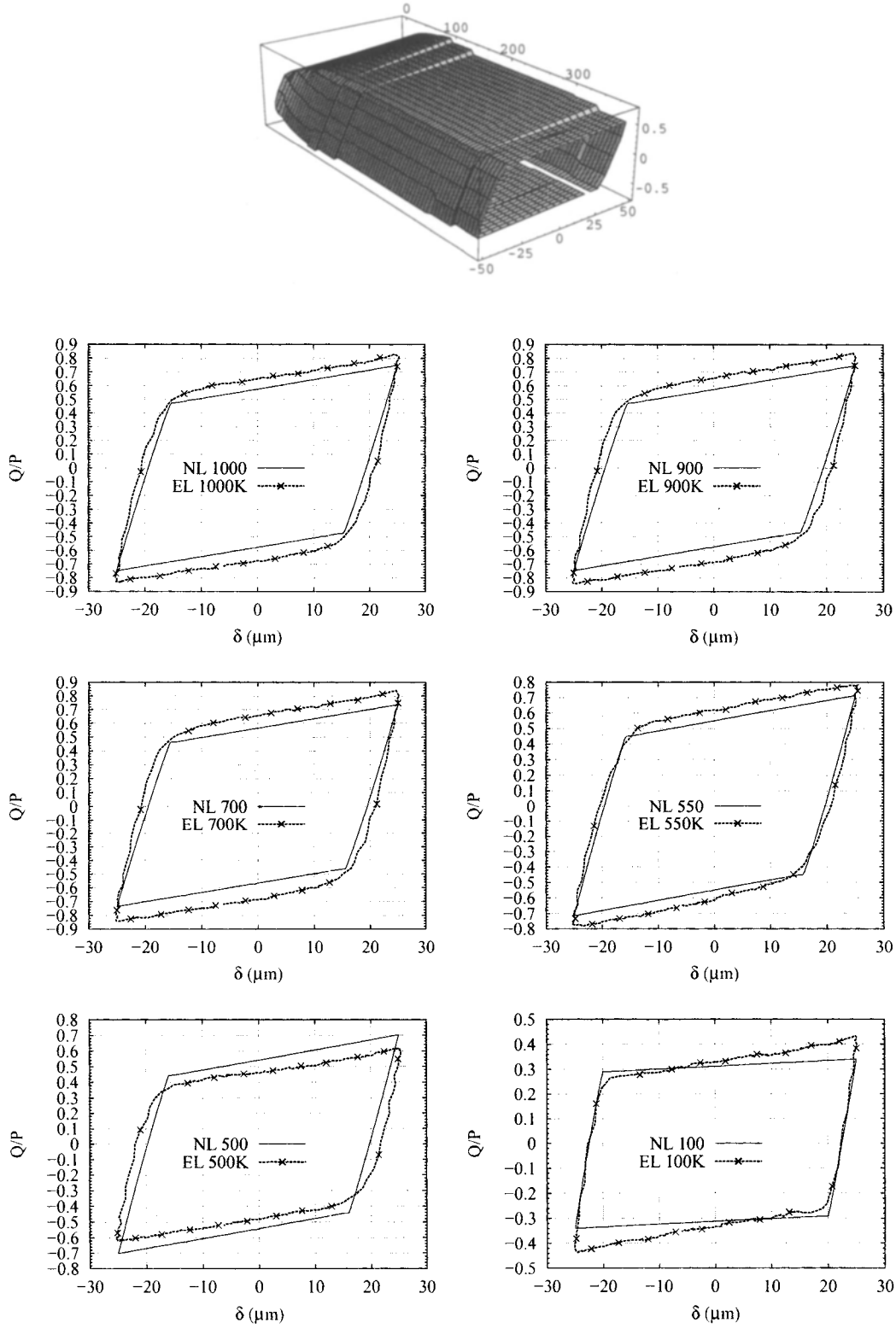
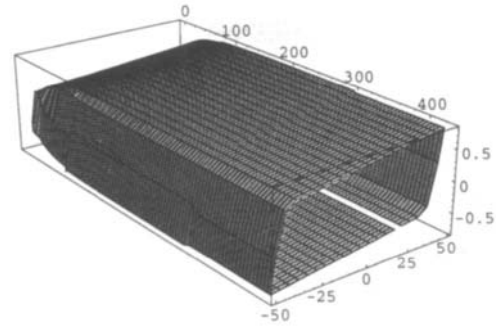
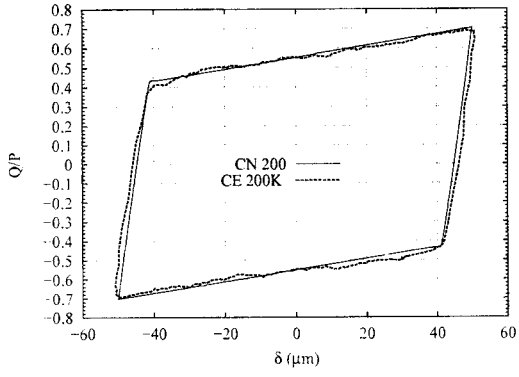
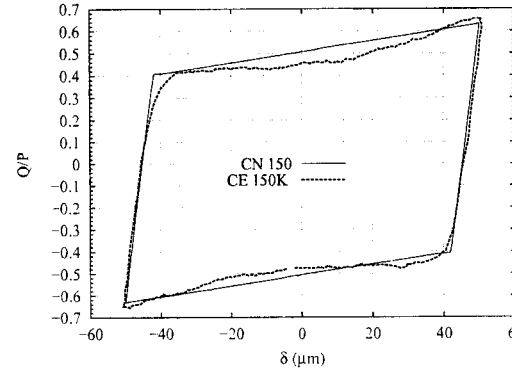
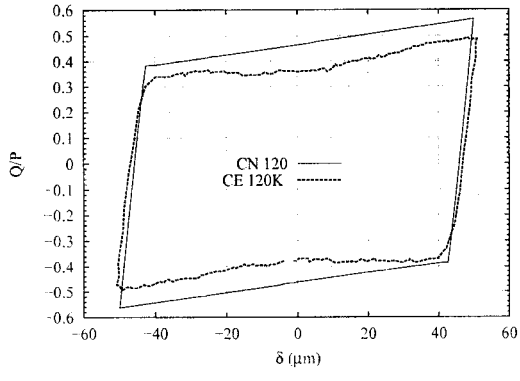


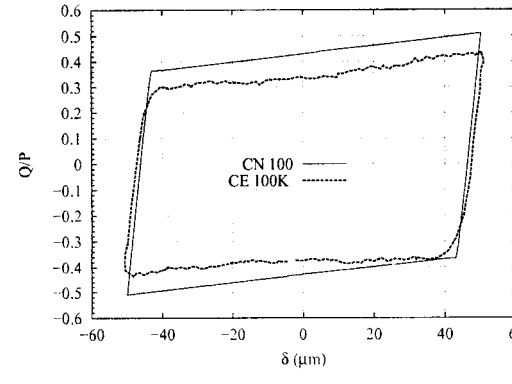
Figure 13: Numerical and experimental cycles evolution for $P = 400$N and $\delta^* = 25$

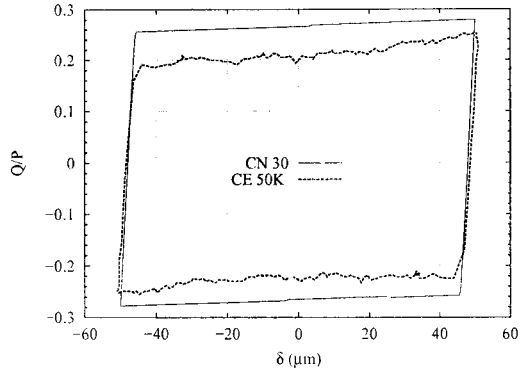

0
100
200
300
400
0.5
0
-0.5
50
25
0
-25
-50

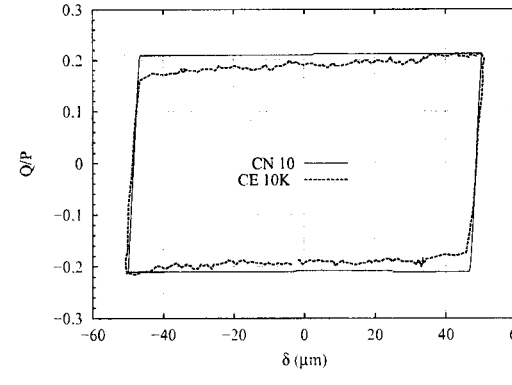

Q/P
δ (μm)
CN 200
CE 200K


Q/P
δ (μm)
CN 150
CE 150K


Q/P
δ (μm)
CN 120
CE 120K


Q/P
δ (μm)
CN 100
CE 100K


Q/P
δ (μm)
CN 30
CE 50K


Q/P
δ (μm)
CN 10
CE 10K

Acknowledgements

This study is financially supported by SNECMA Moteurs. The authors wish to thank P. Perruchaut and S. Mosset (from SNECMA Moteurs, centre Villaroche) for the helpful discussion and permission to publish. The authors are gratiful to V. Fridrici and S. Fouvry (from LTDS, Ecole Centrale de Lyon) for experimental data and useful discussions.

References

[1] R. B. Waterhouse. Fretting fatigue. *Inter. Materials Reviews*, 37(2):77–97, 1992.

[2] M. P. Szolwinski and T. N. Farris. An experimental study of fretting fatigue crack nucleation in airframe alloys: a life prediction and maintenance perspective. In *Proceedings of 1st joint DoD/FAA/NASA conference on aging airfact*, Ogden, UT, 1997.

[3] N. Strőmberg. A Newton method for three-dimensional fretting problems. *Int. J. Solids and Structures*, 36:2075–2090, 1999.

[4] C. Petiot, L. Vincent, K. Dang Van, N. Maouche, J. Foulquier, and B. Journet. An analysis of fretting-fatigue failure combined with numerical calculations to predict crack nucleation. *Wear*, 181-183:101–111, 1995.

[5] D. R. Swalla and R. W. Neu. Influence of coefficient of friction on fretting fatigue crack nucluation prediction. *Tribology international*, 34:493–503, 2001.

[6] A. Curnier. A theory of friction. *Int. J. Solids Structures*, 20(7):637–647, 1983.

[7] R. Michalowski and Z. Mroz. Associated and non-associated sliding rules in contact friction problems. *Arch. Mech.*, 30:259–276, 1978.

[8] V. Fridrici, S. Fouvry, and P. Kapsa. Effect of shot peening on the fretting wear of Ti-6Al-4V. *Wear*, 250:642–649, 2001.

[9] A. E. Giannakopoulos and S. Surech. A three-dimesional analysis of fretting fatigue. *Acta mater.*, 46(1):177–192, 1997.

[10] C. T. Tsai and S. Mall. Elasto-plastic finite element analysis of fretting stresses in pre-stressed stip in contact with cylindrical pad. *Finite elements in analysis and design*, 36:171–187, 2000.

[11] A. Francavilla and O. C. Zienkiewicz. A note on numerical computation of elastic contact problems. *Int. J. Numer. Methods Eng.*, 9(4):913–924, 1975.

[12] T. D. Sachdeva and C. V. Ramakrishnan. A finite element solution for the two-dimensional elastic contact problems with frictions. *Int. J. Numer. Methods Eng.*, 17:1257–1271, 1981.

[13] M. Jean and G. Touzot. Implementation of unilateral contact and dry friction in computer codes dealing with large deformations problems. *Journal de Mécanique théorique et appliquée*, 7(supplement n: 1):145–160, 1988.

[14] M. Wronski. Couplage du contact et du frottement avec la mécanique non linéaire des solides en grandes déformations. Thèse d'université, Université de Technologie de Compiègne, 1994.

[15] F. Jourdan, P. Alart, and M. Jean. A gauss-seidel like algorithm to solve frictional contact problems. *Computer Methods in Applied Mechanics and Engineering*, 155:31–47, 1998.

[16] V. Fridrici, S. Fouvry, and P. Kapsa. Etude du comportement en fretting d'un alliage de titane Ti-6Al-4v: Application au dimensionnement du contact aubedisque de fan. *Rapport d'avancement N:3*, Octobre 2000.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Micro-abrasive wear of semi-crystalline polymers

Y. J. Mergler[a,b] and A. J. Huis in 't Veld[b]

[a]Dutch Polymer Institute, P.O. Box 902, 5600 AX Eindhoven
[b]TNO Industrial Technology, Surface Technology, De Rondom 1, 5612 AP Eindhoven, The Netherlands. Tel. +31 40 2650376, Fax. +31 40 2650302. E-mail: y.j.mergler@ind.tno.nl

The abrasive wear of a series of four semi-crystalline polymers was investigated with a micro-scale abrasive wear tester. The wear of four semi-crystalline polymers: PA6.6, PBT, POM and PTFE was measured under specific test conditions. The silicon carbide slurry concentration and the load were varied. The wear coefficients of the polymers were compared with the (mechanical) properties.

Scanning Electron Microscopy images showed that high slurry concentrations and low loads lead to three-body wear, while low slurry concentrations and high loads result in predominantly two-body grooving. Depending on the experimental conditions, several relations between wear and polymer properties can be found. Experimental conditions should, therefore, be carefully chosen to avoid different wear regimes while measuring a series of polymers. In order to compare the wear coefficient of polymers and obtain relations between wear and polymer properties one should be aware of the dominant wear mechanism.

Keywords: Abrasive wear, grooving wear, rolling wear, two-body wear, three-body wear, Micro-scale abrasion, Polyoxymethylene, Polyamide 6.6, Polybuthylene terephthalate, Polytetrafluoroethylene.

1. INTRODUCTION

A micro-scale abrasion test was recently developed by Hutchings et al [1, 2] for measurements of the wear resistance of the outermost (less than 30 μm) surfaces of a material. Although originally developed for studying hard coatings, progressive studies appeared on bulk materials, for instance glass [3, 4], hard and soft metals [5, 6] and polymers [7]. A micro-scale abrasion test consists of rotating a hard steel ball against a sample in the presence of abrasive slurry, usually silicon carbide. The rotational speed, slurry concentration and load can be varied.

Shipway et al found ridges within the wear scar when using the micro-abrader technique on glass [3, 4]. The formation of those ridges, or ribbons, is dependent on the applied experimental conditions. Ridges form at a relatively high load and low slurry concentration. The formation is explained by the hindered, or prevented, entering of SiC particles in the contact region [7]. Low loads and high particle concentration will tend to promote slurry entrainment and thus reduce the ribbon formation [6, 4]. Allsopp et al found that the surface condition of the ball influences the magnitude and reproducibility of the measured wear coefficient. A ball with slight surface pitting is recommended [5].

Apart from the care that should be taken to ensure reproducible results, another important phenomenon that should be carefully considered is the possible wear transition from two- to three-body wear. Generally, low loads and high slurry

concentration lead to three-body wear with the characteristic image of a heavily deformed and multiple indented wear surface. On the other hand, high loads and low slurry concentration lead to predominantly two-body wear, characterised by grooving with groove sizes comparable with the size of the particles [6]. This finding is of utmost importance when comparing a range of materials. Optimum conditions for three-body wear might be different throughout a series of materials.

The three-body wear behaviour of the following polymers is compared: Polyoxymethylene (POM), Polyamide 6.6 (PA6.6), Polybuthylene terephthalate (PBT) and Polytetrafluoroethylene (PTFE). The optimum conditions are defined by measuring the wear rate at different concentrations and applied normal loads. Attention was paid to ensure that the surface condition of the ball was constant within our series of measurements.

2. EXPERIMENTAL

2.1. Micro-scale abrasion test

Experiments were performed with a commercially available micro-scale abrasion tester from Plint (TE-66). Details about the configuration of the micro-scale abrasion tester can be found elsewhere [1, 2]. Basically, it consists of a stationary specimen, which is pressed onto a rotating 100Cr6 steel ball by a dead weight. The ball rotates about a horizontal axis, parallel to the plane of the specimen surface. An abrasive slurry is dripped onto the ball and its feed is coupled to the rotational speed of the ball. Slurry is entrained into the gap between the ball and specimen, resulting in wear of the specimen.

The wear volume can be calculated from the diameter of the wear scar.

$$V \cong \pi b^4 / (64 R) \qquad (1)$$

V is the wear volume, b the diameter of the wear crater and R the radius of the ball. The penetration depth, h, is determined by the radius of the ball via:

$$h \cong b^2 / (8 R) \qquad (2)$$

The wear volume can be related to the total distance of sliding, S, and the normal load on the contact, N, by:

$$V = k S N \qquad (3)$$

where k is the specific wear rate in mm^3/Nm.

The wear experiment was interrupted periodically to measure the wear scar diameter. The wear scar was measured using an optical microscope. The wear coefficient was determined according to equation 3, assuming that only one abrasive mechanism or one particular combination of mechanisms prevails during the test [6].

A 100Cr6 hardened steel ball, 25.4 mm in diameter, was used. All experiments were carried out with the same ball, without any surface pre-treatment. Within our range of experiments, no effects of possible changes in the ball-surface were detected. The experimental programme with different loads, polymers and slurry concentrations was performed in a random order.

The abrasive slurry was made from commercially available SiC with a particle size smaller than 4.5 μm. Slurry concentrations were varied from 175 g/l, 350 g/l to 700 g/l. Polyoxymethylene was also measured at slurry concentrations slightly higher and lower than 350 g/l. The slurries were stirred vigorously during testing to prevent settling of the particles. The normal load was varied from 0.5, 1, 2 to 4 N. All tests were performed at a rotational speed of 80 rpm.

Examination of some of the wear scars after testing was done with a Philips XL-30 E-SEM and a Mitutoyo profilometer.

2.2. Materials

Four polymers, POM, PA6.6, PTFE and PBT were tested. PA6.6 and PBT were injection moulded. POM was obtained as extruded sheet. PTFE was sintered. Measurements were performed with a sliding direction parallel to the injection moulding direction. The sliding direction was arbitrarily chosen in case of the extruded POM and sintered PTFE. Table I shows some characteristics of POM, PA6.6, PBT and PTFE [8].

3. RESULTS AND DISCUSSION

3.1. Wear versus load

The change of the wear volume per sliding distance with normal load was studied and is shown in figure 1 to 4 for POM, PA6.6, PBT and PTFE respectively. The wear volume is linear with the normal load for all concentrations measured, as was expected according to equation 3.

Trezona et al [6] found that at low slurry concentrations (50 g/l) the wear mechanism of tool steel is predominantly 2-body wear. At low slurry concentrations, the wear volume was non-linear with the normal load and could be described by a power law function with an exponent of load of 0.62. An increasing load will lead to an increasing area. Consequently, more particles will be needed to wear the increasing wear area of the wear scar. Furthermore, the slurry concentration can affect the actual concentration of abrasive particles in the wear contact [6] and thus the load per abrasive particle. This implies that higher slurry concentrations are needed at higher loads to cover the total wear area with particles. Probably, in Trezona's case, a minimum slurry concentration is required to cover up the total wear area with particles. Below this minimum concentration wear will decrease with an increasing load.

3.2. Wear versus slurry concentration

The wear coefficient varies with slurry concentration. Figure 5 shows the variation of the wear coefficient with slurry concentration for POM. A linear relation exists for 2 N and slurry concentrations below 600 g/l, whereas at a normal load of 0.5 N the relation is clearly non-linear. Generally, the erosive wear rate increases with an increase in particle concentration (or the number of particles). However, for polymers, this linear relation of erosive wear with particle concentration is less obvious [9-11].

Probably, the change in wear coefficient with slurry concentration can be explained by the ease of particle entrainment into the contact area and as a result of that, by the transition in wear mechanism from three- to two-body wear. Trezona et al showed that a minimum concentration of abrasive particles is needed to ensure three-body rolling wear [6]. The abrasive particle concentration in the contact area will decrease with increasing load and with decreasing slurry concentration. This means that on the left-hand side in figure 5, with a normal load of 0.5 N, wear will be predominantly two-body grooving.

Table I
Properties of POM, PA6.6, PBT and PTFE [8]

Polymer	Density (kg/dm^3)	Stress-to-break (MPa)	Strain-to-break (%)	Young's Modulus (GPa)
PTFE	2.17	25	200	0.5
POM	1.41	65	40	2.7
PBT	1.29	50	50	2.5
PA6.6	1.13	80	200	2.0

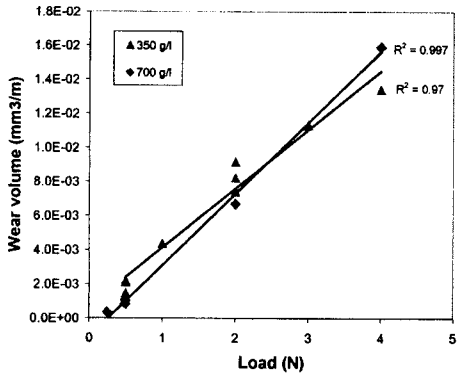


Figure 1. Wear volume per sliding distance of POM as function of the normal load.

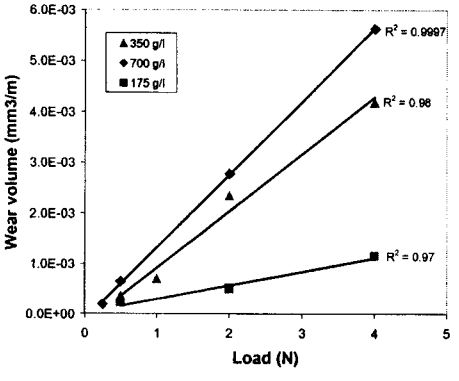


Figure 2. Wear volume per sliding distance of PA6.6 as function of the normal load.

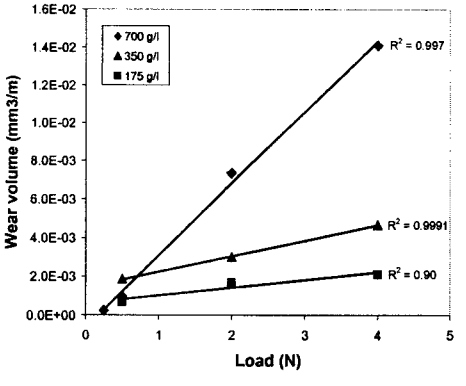


Figure 3. Wear volume per sliding distance of PBT as function of the normal load.

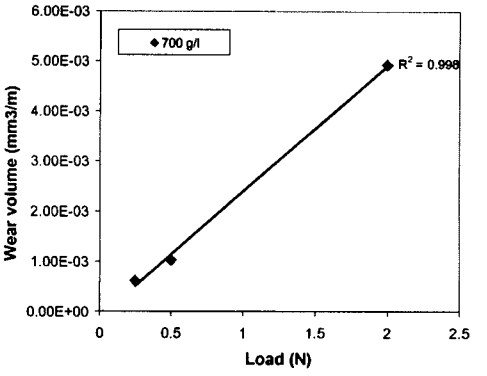


Figure 4. Wear volume per sliding distance of PTFE as function of the normal load.

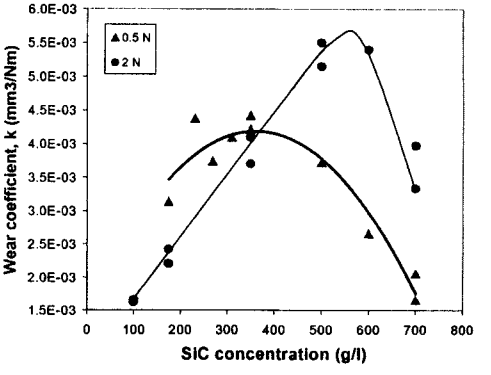


Figure 5. Wear coefficient (mm^3/Nm) of POM as function of the SiC slurry concentration. The drawn lines are a guide to the eye.

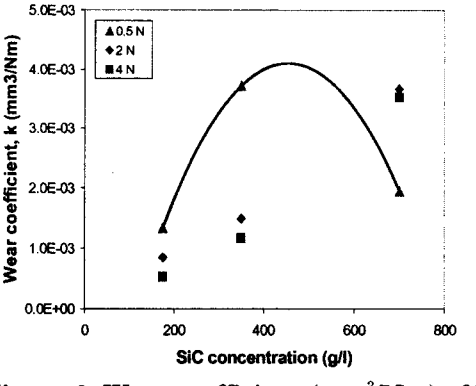


Figure 6. Wear coefficient (mm^3/Nm) of PBT as function of the SiC slurry concentration. The drawn line is a guide to the eye.

An increasing slurry concentration will lead to an increased wear rate of POM, still by predominantly two-body wear processes. However, at a certain slurry concentration, the wear mechanism changes towards predominantly three-body wear, which results in a lower wear coefficient than in two-body wear. This is shown on the right-hand side in figure 5. The results, displayed in figure 5, also imply that at the higher loads of 2 or 4 N, wear will be predominantly two-body grooving. A transition in wear at higher loads will occur at even higher slurry concentrations.

According to Allsopp et al [5] the rather large spread in the wear coefficients of POM with concentration should be explained by the use of an untreated steel ball, as they found that slight surface pitting of the metal ball lead to better reproducibility of the wear results. As was already mentioned in the experimental section, we used the same ball for our complete series of experiments described here without any pre-treatment of the ball surface. We assumed that our measurements are not influenced by the surface condition of the ball. This is confirmed by the following observations: the random order of the measurements of the four polymers under specific experimental conditions, still lead to clear trends of wear with normal load or slurry concentration as displayed in the figures 1 till 7.

Figure 6 shows the relation between wear coefficient and slurry concentration for PBT. The behaviour of PBT is similar to that of POM. A combination of high slurry concentrations and low loads results in a relatively low wear coefficient. A maximum in the wear coefficient is found at a slurry concentration of about 450 g/l with a normal load of 0.5 N. As explained above, this maximum occurs at the transition from predominantly two into predominantly three-body wear.

A different behaviour is seen in figure 7, which displays the results on PA6.6. Here, even at low loads, wear increases with an increasing concentration. This implies that no transition from predominantly two- to three-body wear is found for PA6.6 under these measuring conditions. Therefore, all the polymers were also measured at a slurry concentration of 700 g/l and a normal load of 0.25 N. SEM analysis confirmed predominantly three-body wear of PA6.6 by using a normal load of 0.25 N.

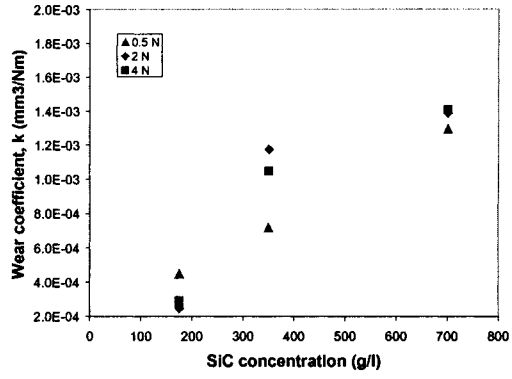


Figure 7. Wear coefficient (mm^3/Nm) of PA6.6 as function of the SiC slurry concentration.

3.3. Wear mechanisms

Figure 8 shows scanning electron microscopy images of the worn surface of POM. The characteristic pattern of three-body wear [6] was found at a load of 0.5 N and a SiC concentration of 700 g/l (figure 8a). Predominantly two-body grooving is seen at lower slurry concentrations (figure 8b) and higher loads (figure 8c). Closer inspection of the grooves reveals that, in some cases, SiC particles rolled in the groove. The size of the groove corresponds to the maximum size of the SiC particles of about 4.5 μm. This effect is illustrated in figure 9.

The colour of the wear scars after testing ranged from white or opaque (depending on the original colour of the sample) via grey to dark greyish-black. In some cases, ridges and grooving could already be discerned by optical microscope images during the periodic measurements of the wear crater diameter. When ridges and grooves were seen on the surface of the wear scar, the colour ranged from white to grey. No ridges or grooves were

seen in combination with a dark grey or black wear scar. SEM images showed that dark grey or black wear scars are a result of predominantly three-body rolling, while white or grey colours indicate two-body grooving of the surface. The colour was definitely not a result of the embedding of SiC particles into the polymer material, as hardly any SiC particles are detected with SEM (see figure 8 and 9). The colour after wear testing can be used as an indication that predominantly three- or two-body wear occurs.

3.4. Comparison of the wear rate of POM, PA6.6, PBT and PTFE

When comparing a series of materials, it is important that the measurements are performed under similar abrasive wear conditions. The results obtained in our measurements showed that a wear transition from two- to three body wear can occur depending on the applied normal load and slurry concentration. This is illustrated in figure 10.

According to Ratner and Lancaster, two-body wear resistance is related to the toughness of a polymer; where toughness (σ_b*ε_b) is defined as the area under the tensile stress-strain curve to break [12, 13]. The tougher the polymer, the higher the wear resistance against two-body wear is. Figure 11a shows that with an applied load of 2 N and a slurry concentration of 700 g/l the polymers wore in the two-body wear regime, as wear decreases with an increasing toughness of the polymers. This was confirmed by SEM photographs (figure 8b, 8c and 9) of the worn surfaces.

However, no correlation between wear and toughness was found when the polymers were measured by using a SiC slurry concentration of 700 g/l and a normal load of 0.5 N. In the latter case, the wear mechanism is predominantly three-body wear, as was shown by SEM (figure 8a). This leads to the conclusion that two-body wear resistance is related to the toughness of the polymer, while three-body wear resistance is not related to the toughness of the polymer.

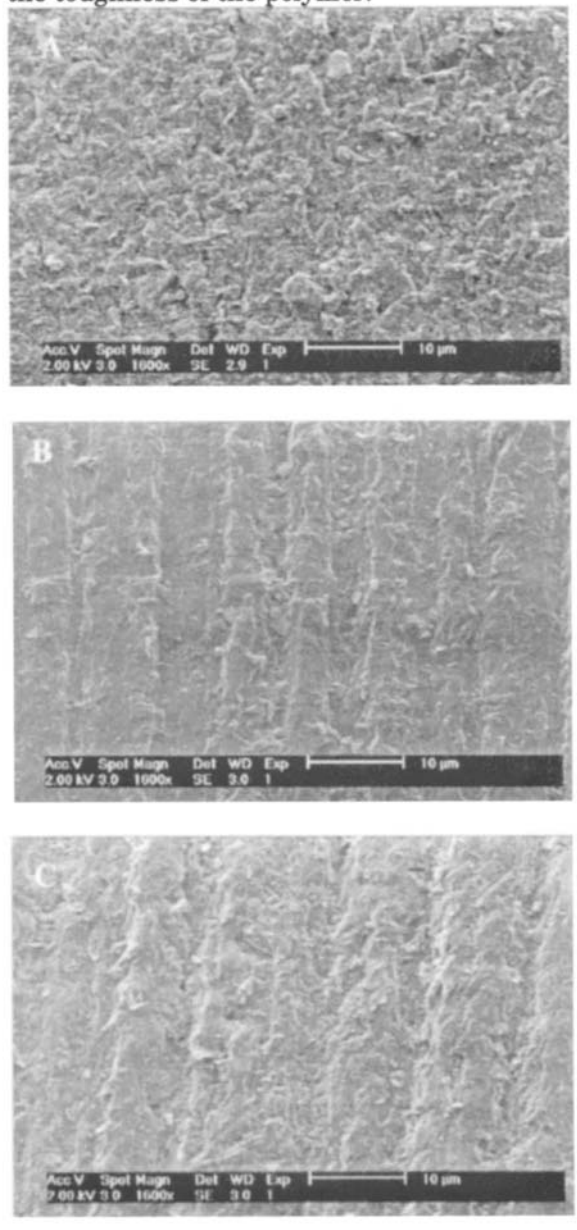


Figure 8. Scanning Electron Microscopy images of the abrasive wear of POM. The sliding direction in the photographs is from top to bottom. a) 3-body wear at a SiC concentration of 700 g/l and a load of 0.5 N. Predominantly 2-body grooving at b) 0.5 N and a SiC concentration of 175 g/l, c) 2 N and a SiC concentration of 500 g/l.

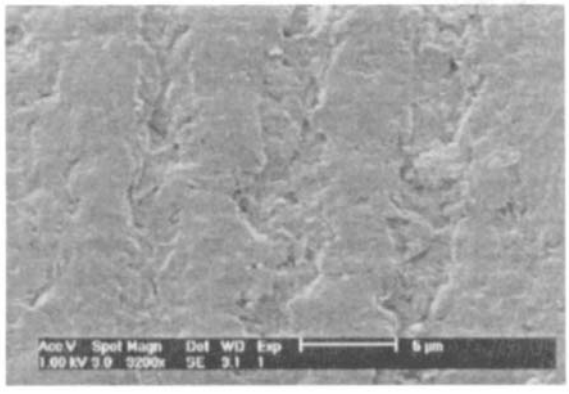


Figure 9. Scanning Electron Microscopy image of the abrasive wear of POM at a SiC concentration of 350 g/l and a load of 0.5 N. The sliding direction in the photograph is from top to bottom. The surface is worn by predominantly two-body wear.

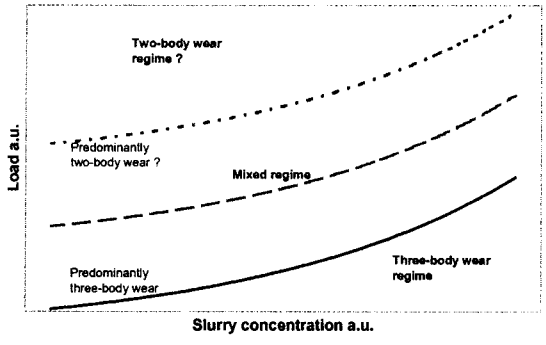


Figure 10. Schematic diagram of wear regimes in micro-abrasive wear measurements depending on experimental conditions.

Less is known about three-body wear. Recently, Tervoort et al and Visjager [14, 15] found a non-linear relation between three-body wear and the average number of entanglements per chain within a series of polyethylene polymers. An entanglement is an effective physical crosslink, or knot, of polymer chains. Above a critical molecular weight, M_c, polymer chains can entangle. An increase in the average number of entanglements per chain led to a decrease in three-body wear [14, 15]. The entanglements will prevent slippage of polymer chains, resulting in higher wear resistance. Visjagers results were obtained on a series of polyethylenes with molecular weights ranging from 2.000 to 2.000.000 g/mole. Above 10 to 20 times the critical molecular weight (Mc = 1250 g/mole [15]), wear resistance hardly increased with an increasing molecular weight or increasing average number of entanglements per chain [15]. We do not know the molecular weights of our, commercially available grades of PA6.6, POM, PBT and PTFE. Relatively low molecular weights can be expected for the injection moulded PA6.6 and PBT compared to the extruded POM and sintered PTFE. We expect the molecular weights of commercially available POM, PBT, PA6.6 and PTFE high enough to provide for enough entanglements per chain to prevent chain slippage. This also means that we assume that with POM, PA6.6, PTFE and PBT no effect of the amount of average number of entanglements per chain on wear is present. This implies that the wear behaviour of the polymers measured here cannot be explained by differences in the average number of entanglements per chain.

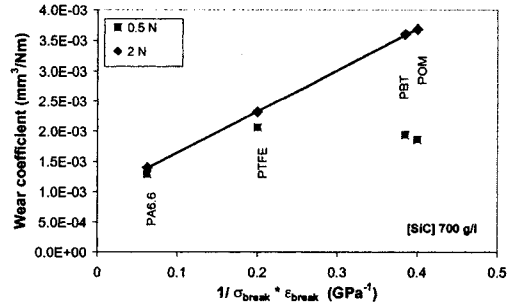


Figure 11. a) Relation between the toughness ($\sigma_{break} * \varepsilon_{break}$) and the wear coefficient (mm^3/Nm) of POM, PA6.6, PBT and PTFE measured with a slurry concentration of 700 g/l.

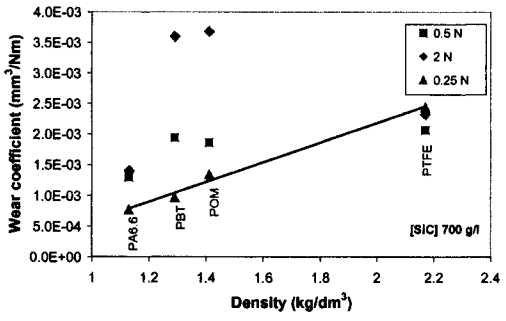


Figure 11. b) Relation between the density and the wear coefficient (mm^3/Nm) of POM, PA6.6, PBT and PTFE measured with a slurry concentration of 700 g/l.

Figure 11b shows that at a very low load of 0.25 N and a high slurry concentration of 700 g/l, a relation is found between polymer density and wear. Under these conditions, predominantly three-body wear occurs, as was shown by SEM. This leads to the conclusion that three-body wear is related to the density of the polymers.

No explicit relation between polymer wear and mechanical properties of the polymers was found for average slurry concentrations and average loads. This can be expected, since the wear mechanisms under those conditions for the four polymers are a mixture of two- and three-body wear.

4. CONCLUSIONS

1) Micro-scale abrasion measurements of POM, PA6.6 and PBT show that the wear volume per sliding distance increases linearly with an increasing normal load.
2) The wear coefficient varies with the abrasive slurry concentration, showing a maximum at a low load of 0.5 N and 2 N for POM and 0.5 N for PBT.
3) The variations in wear coefficient with slurry concentration can be explained by transitions in wear mechanism from predominantly two- to predominantly three-body wear. These results were confirmed by Scanning Electron Microscopy images of worn surfaces. As a rule of thumb, dark wear scars are indicative for three-body rolling wear.
4) Improper choice of the experimental conditions, e.g. normal load and slurry concentration, leads to a mix of abrasive wear mechanisms. In these cases, no relation could be found between wear and toughness of the polymers, or any other mechanical property of the polymers.
5) The wear coefficient of PA6.6, PBT, POM and PTFE decreases with increasing toughness of the polymers, at a load of 2 N and a slurry concentration of 700 g/l. Under these conditions, predominantly two-body wear occurs.
6) Predominantly three-body wear is found at a load of 0.25 N and a slurry concentration of 700 g/l. In this case, wear increases with an increasing polymer density.

REFERENCES

1 K.L. Rutherford and I.M. Hutchings, J. Testing and Evaluation, 25 (1997) 250-260.
2 K.L. Rutherford, R.I. Trezona, A.C. Ramamurthy and I.M. Hutchings, Wear 203-204 (1997) 325-334
3 P.H. Shipway, Wear 233-235 (1999) 191-199
4 P.H. Shipway and C.J.B. Hodge, Wear 237 (2000) 90-97
5 D.N. Allsopp, R.I. Trezona and I.M. Hutchings, Trib. Letters 5 (1998) 259-264
6 R.I. Trezona, D.N. Allsopp and I.M. Hutchings, Wear 225-229 (1999) 205-214
7 R.I. Trezona and I.M. Hutchings, Wear 233-235 (1999) 209-221
8 D. W. Van Krevelen, in: "Properties of Polymers", Third completely revised edition, Elsevier 1997
9 S.W. Zhang, W. Deguo, Y. Weihua, J. Mater. Sci. 30 (1995) 4561-4566

10 A. Yabuki, K. Sugita, M. Matsumura, M. Hirashima, M. Tsunaga, Wear 240 (2000) 52-58
11 J.J. Rajesh, J. Bijwe, U.S. Tewari, B. Venkataraman, Wear 249 (2001) 702-714
12 S.N. Ratner et.al. in: " Abrasion of rubbers", D.I. James eds. Mac Laren and Sons, London, 1967
13 J.K. Lancaster in: "Friction and wear" A Materials Science Handbook, A.D. Jenkins eds., North Holland Publishing Co., 1972
14 T. A. Tervoort, J. Visjager, P. Smith, Macromolec. 35 (2002) 8467-8471
15 J. Visjager, Dissertation ETH Zurich, No. 13978 (Available via www.eth.ch)

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Wear Testing of Acetal for the Prediction of Polymer Rolling Bearing Durability

A F Beke [1], R S Dwyer-Joyce [1], F R Jones [2], J Gregory [3]

[1] Department of Mechanical Engineering, University of Sheffield, Sheffield, UK
[2] Department of Materials Engineering, University of Sheffield, Sheffield, UK
[3] Sarnatech BNL, Knaresborough, UK

Polymer rolling bearings consist of steel balls running in acetal (POM) raceways. Their capacity is often limited by increased diametric clearance caused by wear. The aim of this work is to understand the wear processes with a view to developing a model for bearing life. Wear tests with dry flat acetal specimens in sliding contact with a steel ball were carried out on a reciprocating rig. The aim is to produce data transferable to a rolling bearing scenario. The data gives an understanding of bearing wear throughout the expected life cycle of an acetal rolling bearing. Wear was found to take place by a combination of adhesion and fatigue cracking. Wear rates and surface damage in the model test work were found to be qualitatively similar to those experienced in rolling bearings.

1. INTRODUCTION

A common design of polymer rolling bearing consists of steel balls held in a polymer cage running between polymer raceways. Polyoxymethylene (POM or acetal) is the material of choice in this application since it is dimensionally stable during manufacture and has good tribological properties. The polymer parts are injection moulded in the finished state and require no machining of the critical surfaces. This lends itself to mass production and so these bearings are relatively low cost.

The integrity of the bearings depends on the durability of the ball raceway contacts. Whilst the load capacity will certainly be less than that of a conventional steel ball bearing, the reduced cost and possibility of running without lubricant, mean that these components have many important applications. The key to finding these applications is a full understanding of the durability of the bearings, under various operating regimes. Life testing of bearings is a time consuming and costly process. In this work a simple ball-flat model contact tester has been used to understand the process of failure in a steel ball on acetal flat contact. This data is to be used to assist in the development of life theories for acetal ball bearings.

2. BACKGROUND

2.1. Bearing Operating Conditions

Figure 1 shows a picture of a standard polymer rolling bearing. For this design a maximum load of 100 N and operating speeds of 1500 rpm are commonly specified. If it is assumed that this load is shared by the balls according to the Stribeck equation [1] then a Hertz contact analysis predicts a maximum contact patch of semi-elliptic radii 1.92 and 0.123 mm and a maximum contact pressure of 110 MPa. The rolling contact speed is 0.99 m/s. The micro-slip within the contact patch varies from – 29% to + 0.4%. The low modulus of the raceway material results in a large contact area between ball and raceway. This is why there is inherently more micro-slip in a polymer bearing than in a conventional steel bearing.

The yield stress of acetal is approximately 65 MPa. It is not normal to specify operating loads, which will cause stresses in excess of this yield stress. However, it has been observed that the polymer bearings can operate satisfactorily for long periods at load some 5 times that which would cause first yield.

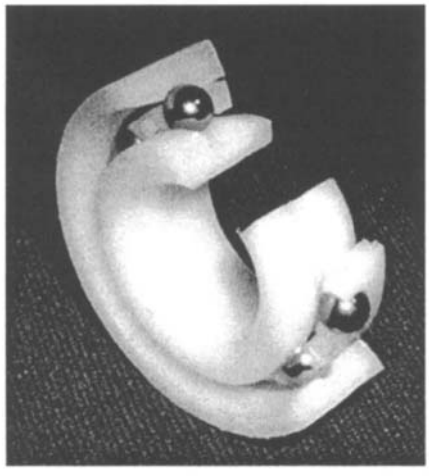


Figure 1 - Sectioned standard polymer rolling bearing

2.2. Modes of Polymer Rolling Bearing Failure

There are two main modes of failure that have been linked to acetal rolling bearings. These modes can occur in conjunction with each other or by themselves. These modes of failure are the result of wear. Wear occurs by the following mechanisms; adhesion between the smooth steel ball and the polymer raceway surface; and material loss associated with fatigue pits on the surface. Generally the steel balls are so smooth that there is little wear caused by surface asperities abrading the polymer surface.

A bearing has failed when the diametric clearance (the amount of movement found between the raceways and balls) exceeds a specified amount. Diametric clearance increases as material from both inner- and outer-raceways has been worn away. Figure 2a shows the wear scar of a failed inner-raceway from a polymer rolling bearing.

Upon closer examination of the wear scar (figure 2b), a rough pitted surface is seen. This wear scar and loss of material has been attributed to the combination of both wear types; fatigue due to the cyclic loading of the raceway as the balls roll around the raceway; adhesion where small particles of the raceway material have stuck to the balls and have been plucked out as the balls roll over.

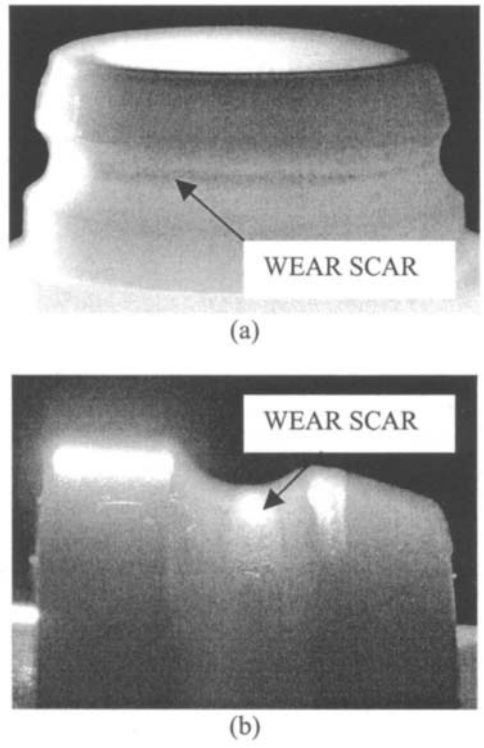


Figure 2 - Wear scar from a failed inner-raceway

A more severe bearing failure mode occurs where the bearing rolling action is no longer smooth and unhindered. This happens when severe scuffing of the raceway surface has taken place. Figure 3 shows two failed inner raceways as a result of severe surface damage. This was the result of the operating conditions and the environment in which the bearings where used. Both bearing raceways were lubricated with a silicone based grease, in addition the bearing raceway on the left in figure 3 was subjected to contamination with water. Fatigue has occurred and the presence of a fluid being forced into the micro defects at the points of contact has accelerated crack growth.

The raceway shown on the left in figure 3 had greater severity of surface damage due to water contamination. It is possible that the water, with its lower viscosity, is more easily entrained into fatigue cracks and can therefore promote crack growth more readily.

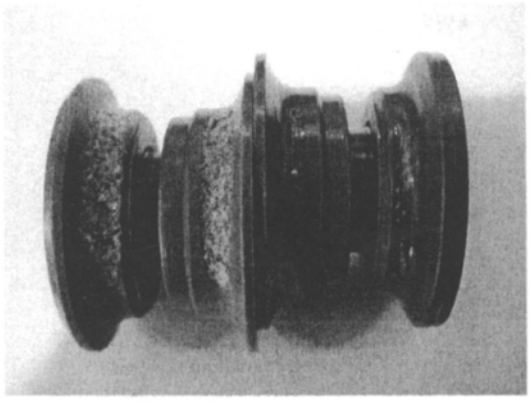


Figure 3 - Failed inner-raceways tested with a lubricant

2.3. Acetal Wear Mechanisms

At mild contact conditions (i.e. low load, speed, and slip), the wear of unlubricated acetal against acetal is thought to take place by a surface fatigue process [2]. Small surface pits are formed which propagate and debris is removed in flakes. Under more severe contact conditions there is an increase in frictional heating and wear is generally associated with softening or melting. In this regime surfaces crack in the transverse direction [2] and large areas of the surface become detached.

Over a range of contact conditions (i.e. varying load, speed and slip) the wear of dry, unlubricated acetal against steel, is thought to be predominated by adhesion rather than abrasion [3].

Lancaster [4] suggests that the wear of polymers against steel is dominated by three processes; abrasion, adhesion, and fatigue. The wear process involved is dependent on the surface topography and the material characteristics. Abrasive wear depends on the metal surface roughness and two polymer material characteristics (its breaking strength and elongation to break). Whereas adhesive and fatigue wear are considered as interrelated and dependent on how the surface topography changes during sliding. Similar observations were made by Santer et al. [5] who suggests that at low surface roughness values, adhesion is considered the dominating factor, whereas for high values abrasive wear prevails.

Under increasing severe contact conditions (i.e. rising temperature) the wear rate of acetal sliding against steel reaches a minimum at 110°C, before rising sharply as the temperature approaches the acetal softening point [6]. In materials it is generally thought that the ratio of indentation hardness (H) to elastic modulus (E) is related to wear rate. The lower the H/E, the greater the wear rate. However the relationship between wear rate and varying temperature cannot be explained because the ratio of H/E seems to be independent of temperature till nearing the softening point [6].

2.4. Rolling Bearing Wear

There is little data available on the wear and failure of polymer rolling bearings. The steel balls are very smooth so two body abrasion is unlikely to occur. This means that surface damage is probably associated with adhesion or fatigue. To understand these failure mechanisms in a rolling bearing is difficult since life testing is very time consuming. The purpose of this work is to establish a durability test method based on a single reciprocating contact. Loading cycles can be accumulated rapidly so this approach provides a convenient method for material screening tests.

3. APPARATUS & PROCEDURE

3.1. Apparatus

A reciprocating rig was used throughout this investigation, shown schematically in figure 4. The rig consisted of a Ling Dynamic Systems V200 Series electro-dynamic vibration generator connected to a function generator and amplifier. The vibrator was driven by a signal from the function generator passing through an amplifier. A double-beam oscilloscope and a voltmeter were used to monitor the signal throughout the experiment.

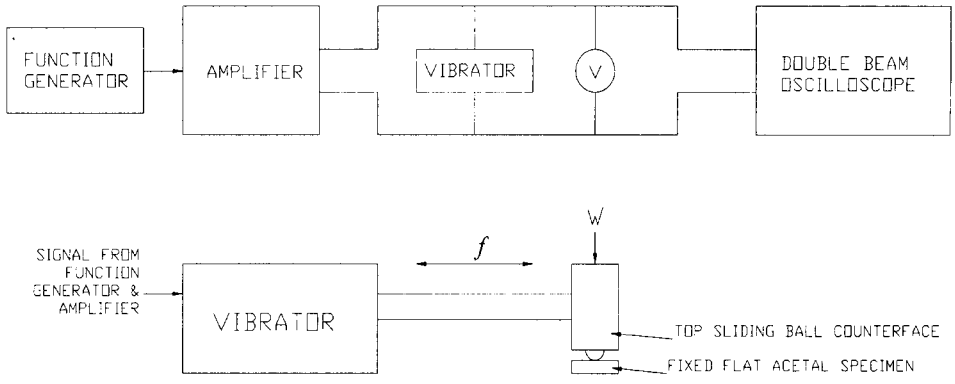


Figure 4 - Schematic of the reciprocating rig

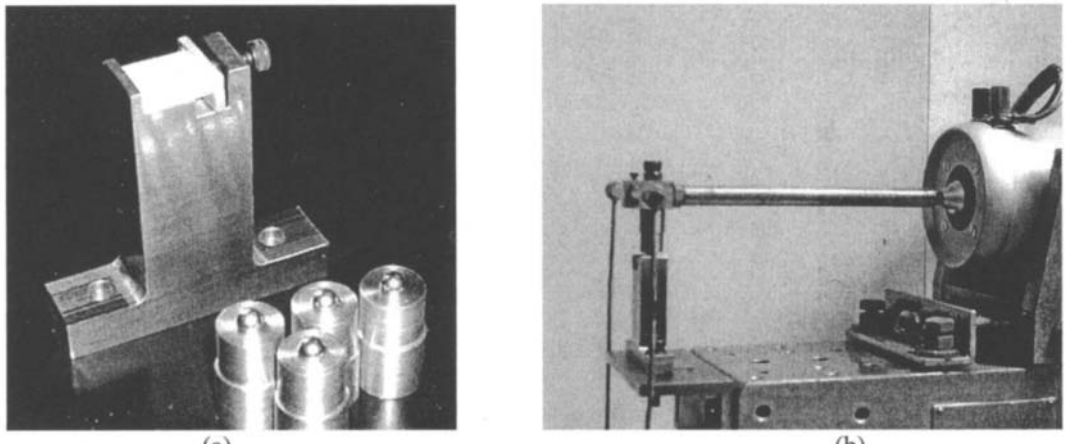


Figure 5 - (a) Specimen holders, (b) Vibrator and contact loading assembly

uncoloured, injection-moulded plaques were used as the bottom specimen. The acetal plaques where injection-moulded at the same mould tool temperatures. This ensured that the fine surface crystal structure of the plaques resembled that of the bearing raceway material. Stainless steel balls (as used in the polymer rolling bearing-diameter 4.7625mm) mounted into cupped holders were used as the counter-face. Figure 5 shows the acetal holders.

A mounted ball bearing was attached by an arm to the vibrator. Dead weight load was applied by use of a hanger connected to the ball holder.

3.2. Procedure

Tests were carried out under the following conditions:

- Reciprocating rig-operating frequencies of 100, 120, and 140 Hz.
- Peak contact pressures of 80, 100, 115, 127, 136, and 145 MPa.
- Sliding distance of 1.525 ± 0.325 mm.
- Test duration 10 minutes.

4. RESULTS

4.1. Wear Scar Appearance

The images depicted in figure 6 show the wear scars obtained at an operating frequency of 140 Hz for a range of normal loads. Similar results were observed at the other operating frequencies. Under close examination of the images in figure 6, three different effects of surface wear become evident. Polishing of the surface, growth of transverse cracks and surface melting. These effects accumulate as the loading increases. Under low load, figure 6a, polishing was found to be the dominant surface damage of the wear scar. Material has been removed but there is not severe surface damage. As the loads were increased the formation of transverse cracks, starting in the maximum Hertzian stress region, along the whole length of the wear scar were observed. The higher the respective loads the longer the crack lengths, (see figure 6b-6e). At a loading of 127 MPa and above, figure 6d-6f, surface melting was found to take place in the mid-regions of the wear scar. Furthermore, wear debris formation was observed as particle build up at the edge of the wear scar, figure 6d-6f, and as build up of polymer particles of similar size on the ball counter-face. The particles could be wiped of with ease from the surface.

4.2. Wear Rate

The wear rate has been expressed as the average depth of material removed per sliding cycle. One cycle being defined as two passes along the wear scar length. The depth worn has been determined approximately from the dimensions of the wear scar and the ball counter-face. It is assumed that the acetal specimen is the only surface to exhibit wear, because of the short time span of the experiments and the material characteristics of the interfaces. This definition was chosen since, generally the wear rate of bearings are determined by the increase in diametric clearance, i.e. depth wear of the raceways. Figure 7 shows how the wear rate varies with the maximum contact pressure at three different oscillating frequencies.

In general, it can be seen that as the pressure increases, the wear rate increases. Furthermore it is noted that the lower the operating frequency, the greater the wear rate at similar loads. There also appears to be a wear rate transition at a maximum contact pressure of around 115 MPa.

5. DISCUSSION

5.1. Wear Characteristics

At low loads the wear is probably adhesive in nature. As the contact stress increases, fatigue cracks are observed growing in the transverse direction. The threshold for the initiation of fatigue cracks occurs at around 115 MPa, which is 1.8 times the yield stress of acetal. It is of note that the fatigue cracks occur across the whole wear track length and not just at regions of maximum velocity. This suggests the fatigue is stress rather than velocity (or temperature) controlled. The onset of fatigue has resulted in material removal and a transition in the wear rate. It should be noted however that in this test geometry the contact is fully sliding. In the rolling bearing case the contact is nominally rolling with a degree of micro-slip. In that case the velocity may have a more significant effect on cracking.

At the highest loads, there is evidence of surface melting having occurred (see figures 6d-6f). This is associated with frictional heating of the surface and is observed to be most severe in regions of highest velocity. Furthermore it is probable that the characteristic wear rate temperature of 110°C, as observed by Lancaster [6], has been attained and surpassed in the maximum contact pressure region of 115-127 MPa. This at present is undetermined since no temperature measurements have been made. However, with a transition in wear rate noted in this

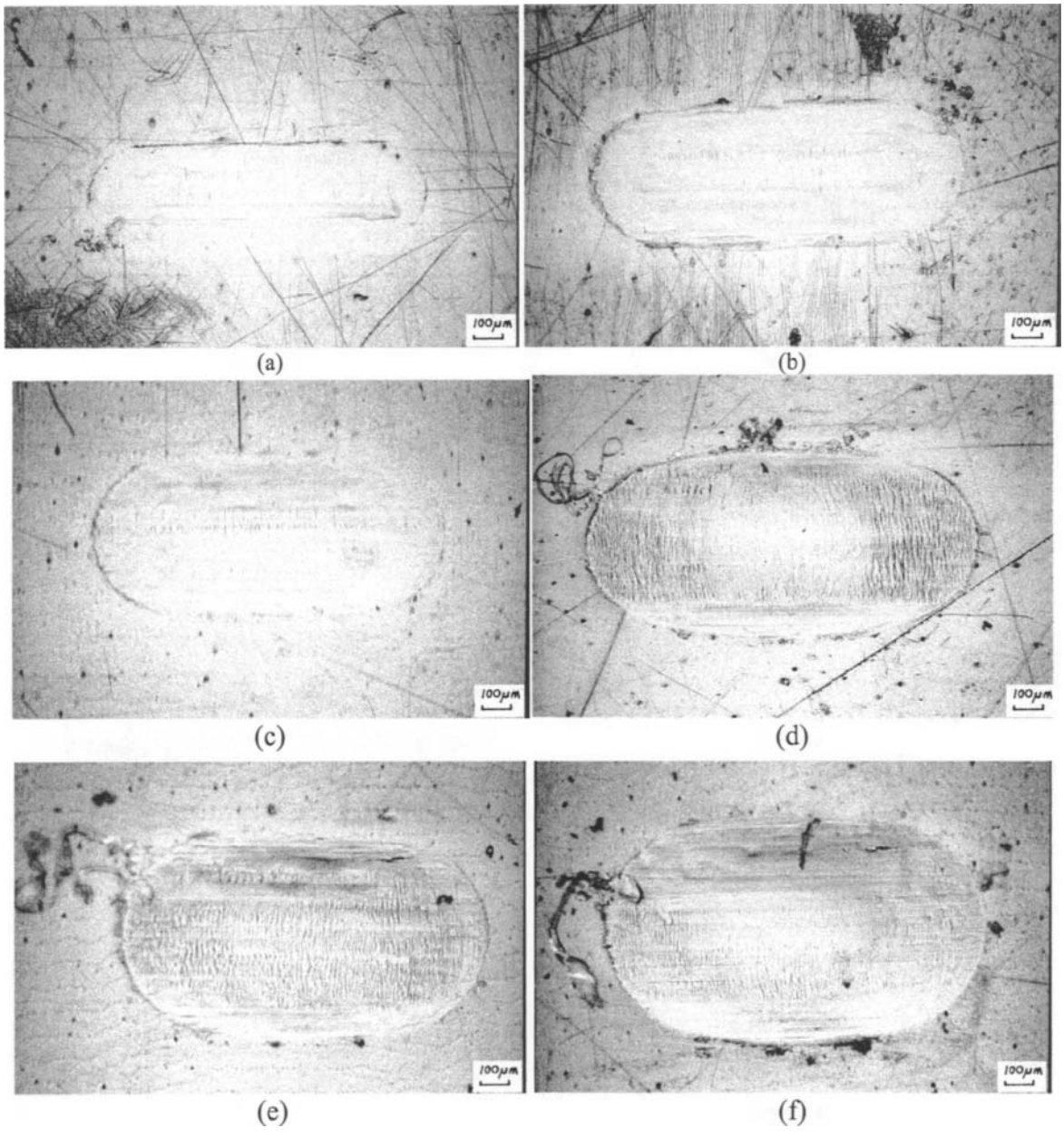


Figure 6 - Wear scars on the acetal specimens obtained at 140 Hz and contact pressures of; (a) 80 MPa, (b) 100 MPa, (c) 115 MPa, (d) 127 MPa, (e) 136 MPa, (f) 145 MPa

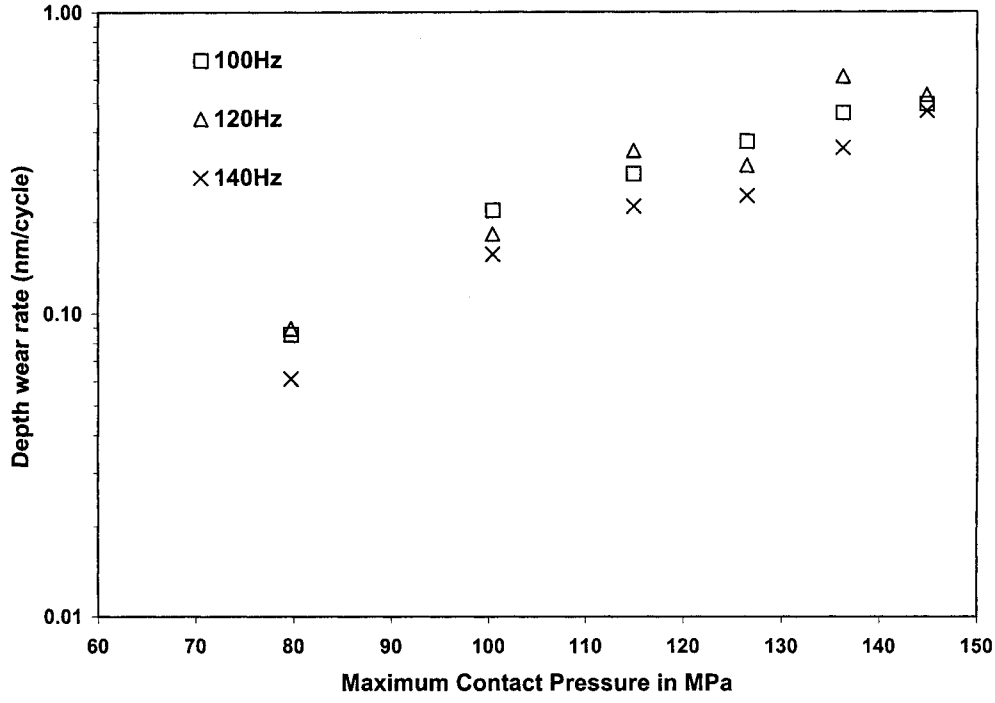


Figure 7 - Wear rate against contact pressure at three oscillating frequencies

pressure region, it is likely that this is the case.

As the frequency decreases (i.e. as sliding speed decreases), there is a small increase in the wear rates for all loads. The wear rate of the sliding system usually increases with speed as frictional heating takes place. At this stage it is unclear why the wear rate in this case decreases slightly with increasing the sliding frequency.

5.2. Application of the Test Results to Acetal Rolling Bearings

The test set-up consists of a stainless steel ball sliding across a flat acetal specimen. Whereas in the rolling bearing scenario, the balls roll within the raceways as the bearing rotates. Due to the conformed nature of the ball raceway contact there is a degree of micro-slip within the nominally rotating contact. The average slip magnitude over the contact can be calculated (see for example [7]) and is typically in the region 3 - 10%.

A simple approach, therefore to compare wear rates between the sliding model contact and the ball bearing case is to scale the data according to the slip ratio. This assumes that wear is proportional to sliding distance (i.e. following an Archard [8] approach) and so is strictly only appropriate for adhesive or abrasive wear and not fatigue associated material removal.

In rolling bearing life tests two wear regimes are apparent. A low wear regime exhibited under mild operating conditions (<80 MPa) where a characteristic wear rate is typically 0.001 nm per ball cycle. A high wear regime exhibited under severe operating conditions (>100 MPa) where a characteristic wear rate is typically 0.013 nm per ball cycle. In the former case the average micro-slip across the contact is 5.3%, in the latter higher loaded case the average slip is 9.8%. In both wear regimes a similar transition in wear rate was observed as exhibited in the test results.

From the reciprocating sliding test results, at a similar low wear regime to the bearing operating conditions (80 MPa), a wear rate of 0.086 nm/cycle was obtained. The high wear rate at higher loading (115 MPa) was found to be 0.226 nm/cycle. This test data corresponds to 100% slip. If the data is scaled by the mean bearing microslip values at low and high loads, (5.3% and 9.8%) then the wear rates are 0.005 nm/cycle and 0.02 nm/cycle respectively.

In both corresponding high and low wear regimes the wear rates are of a similar order. Albeit the scaled test wear rates were slightly higher than that of the actual bearing life data, a qualitative analysis and a starting point for a quantitative analysis has been achieved.

6. CONCLUSIONS

A wear test has been developed using a reciprocating rig to attain wear rates from a large number of cycles in a short space of time. The test set-up consisted of a simple ball sliding over a flat acetal specimen. A range of loads and operating frequencies were used. It was found that as the load was increased the wear rate increased with a transition occurring at the onset of fatigue cracking. This wear rate is believed to be influenced by the temperature attained by frictional heating. Above this load regime the wear rate increased sharply again. This rise in wear rate was believed to be the result of the temperature approaching the softening/melting point of the acetal.

The wear mechanisms of an unlubricated sliding contact between acetal and a smooth steel ball is a combination of adhesion and fatigue.

A qualitative comparison of the test results and bearing wear data was undertaken. The comparison indicates similar wear rates occur in the two cases.

REFERENCES

1. Harris, T. A., 2001, "Rolling Bearing Analysis," Fourth Edition, Wiley-Interscience publication, Chichester.
2. Kukureka, S. N., Chen, Y. K., Hooke, C. J., and Liao, P., 1995, "The wear mechanisms of acetal in unlubricated rolling-sliding contact," Wear 185, pp 1-8.
3. Kar, M. K., and Bahadur, S., 1974, "The wear equation for unfilled and filled polyoxymethylene," Wear 30, pp 337-348.
4. Lancaster, J. K., 1973, "Basic mechanisms of friction and wear of polymers," Plastics and Polymers 41, pp 297-306.
5. Santner, E., and Czichos, H., 1989, "Tribology of polymers," Tribology International 22, pp 103-109.
6. Lancaster, J. K., 1968-9, "Relationships between the wear of polymers and their mechanical properties," Proc. Inst. Mech. Engineers 183, pp98-106.
7. Johnson, K. L., 1989, "Contact mechanics," Cambridge University Press, Cambridge.
8. Archard, J. F., 1953, "Contact and rubbing of flat surfaces," J. Appl. Phys. 24, pp981-988.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Self-Lubricating Dry Rubbing Bearings - Fundamentals and Methods of Calculation

L. Deters[a], F. Mueller[a], M. Berger[b]

[a] Otto-von-Guericke-University Magdeburg, Institute of Machine Design, P.O. Box 41 20, D-39016 Magdeburg, Germany

[b] GLACIER GARLOCK BEARINGS, P.O. Box 18 62, D-74008 Heilbronn, Germany

Following a general discussion about the construction and the lubrication mechanisms of compound bearings made of a steel backing with a porous bronze layer filled with a PTFE/solid lubricant mixture, problems in conjunction with the use of pU diagrams are pointed out. Instead of pU values parameters such as the overlap ratio, the compressive strain or the shear stress ratio for example which are important for a reliable design of dry rubbing bearings are represented then. As basis for the friction calculation a relationship between the shear stress ratio and the quotient of the compressive strain and the overlap ratio was experimentally found and the temperature-dependent curves of the Young's modulus and of the shear strength of PTFE are shown. After that, the contact angle influenced by deformation, wear and temperature rise is examined and an approximation equation for determining the contact angle is introduced.

A method for calculating temperature, friction and working life of dry rubbing bearings that is based on theoretically and experimentally established data is offered afterwards. In addition the positive effect of a wear particle trap is demonstrated. Finally, calculation examples show the mode of operation of the calculation method.

1. INTRODUCTION

Self-lubricating dry rubbing bearings are mainly used in applications where no lubricants are desirable or permitted, loads are high and/or low sliding velocities need to be achieved, bearing dimensions should be small, and maintenance should be kept on the lowest level possible. They are employed for example in the aircraft and aerospace industries, in motor vehicles, agricultural and building machines, office and household devices, and in the food industry.

In general, the operating performance of self-lubricating dry rubbing bearings is described by using the pU value. This value is the product of the specific surface pressure $\overline{p} = F/(B \cdot D)$ with the normal load F, the bearing width B, the inner bearing diameter D and the sliding velocity U.

However, any design of dry rubbing bearings based on the pU value must be considered to be an approximation only. The pU graph provides information only about the performance limits of bearings established in the test machine where the bearings were examined, but cannot be simply applied to other bearings in general. Hence, sufficient accuracy can only be ensured in bearing design when the frictional behaviour and the conditions of heat transfer are included in addition to the loading parameters.

2. CONSTRUCTION AND OPERATION OF COMPOUND BEARING MATERIALS

Figure 2.1 depicts a blown-up section of the surface layer in a compound bearing material.

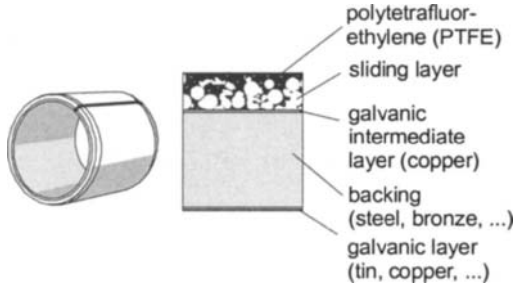


Figure 2.1. Construction of a compound bearing material

The basis of the compound bearing is a 1 to 3 mm thick backing most frequently consisting of steel or bronze. A galvanic layer of copper or tin on the outer side of the backing ensures good removal of heat to the surrounding housing, and serves as a corrosion protection. An intermediate layer is applied to the inner side of the backing to ensure good bonding with the sliding layer above. The latter consists of a porous bronze (70 to 80% v/v) and is 0.2 to 0.4 mm thick. A mixture of a polymer, in this case polytetrafluorethylene (PTFE), and solid lubricant additives (lead, MoS_2, etc.) are rolled into the pores of the bronze. The volume percents amount to about 80% for PTFE and 20% for the solid lubricant. Finally, the material is covered by a 5 to 30 µm thick running-in layer consisting of the same polymer/solid lubricant mixture as in the sliding layer.

As a result of the strongly varying material characteristics of the materials contained in the compound bearing, the compound bearing material exhibits very specific properties. The resulting performance mechanisms are discussed below.

In the sliding process the running-in layer which is on the sliding layer in the beginning, is transferred onto the surface of the metallic counterface and fills its roughness valleys with PTFE and the solid lubricant partly or completely. Here, the surface structures and the dimension of the sliding surface of the mating component play an essential role. To fill the roughness valleys completely, the volume of the running-in layer of the bearing must be designed such that a sufficient volume of the PTFE/solid lubricant mixture is applied to the overall sliding surface of the counterbody. WALDHÜTER [1] suggested an optimum range for the roughness of the mating steel surface. This roughness must not be too low in order to ensure mechanical interlocking of the polymer/solid lubricant mixture on the surface. However, if it is too high the volume of the running-in layer would not be sufficient to fill the roughness valleys and form a uniform friction-reducing layer.

The actual sliding layer consists of a compound of two materials exhibiting physical and mechanical properties which strongly differ from each other. The compressive strength of the compound bearing material is ensured by the thin porous sintered bronze layer, whereas the good sliding properties result from the PTFE/solid lubricant mixture. To ensure its friction-reducing effect, the mixture needs to be transferred into the contact zone. This is achieved by two different mechanisms which may operate simultaneously or in succession.

When locally high specific loads are applied to the bearing material, the PTFE/solid lubricant mixture (referred to as PTFE+SL below) is plastically deformed and squeezed out from the compound material due to its low strength. This process is supported by the porous structure of the bronze layer and can be compared to squeezing out a sponge (Figure 2.2).

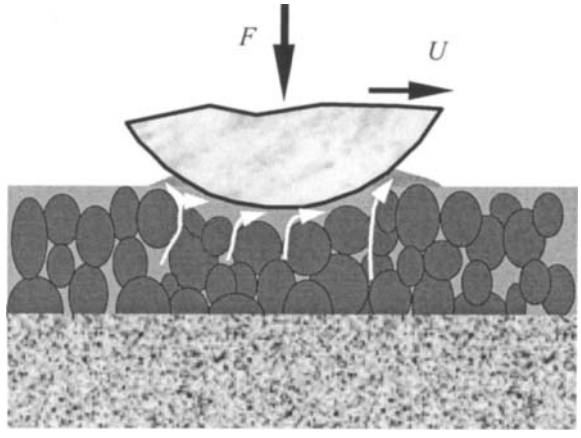


Figure 2.2. PTFE/solid lubricant mixture displaced as a result of high local surface pressure

The second way in which the lubricant can get into the contact zone, is based on the different coefficients of thermal expansion of the components bronze and PTFE+SL. Their ratio is about 1 : 5.5. When the sliding layer warms up, expansion of the PTFE+SL contained in the compound is 5.5 times higher compared to bronze. As a result of this high expansion rate, the mixture is displaced from the pores in the sintered bronze and gets into the contact zone (refer to Figure 2.3). The PTFE+SL displaced is sheared off by the opposing surface, carried away and distributed as a thin layer on the sintered bronze and/or on the mating running surface.

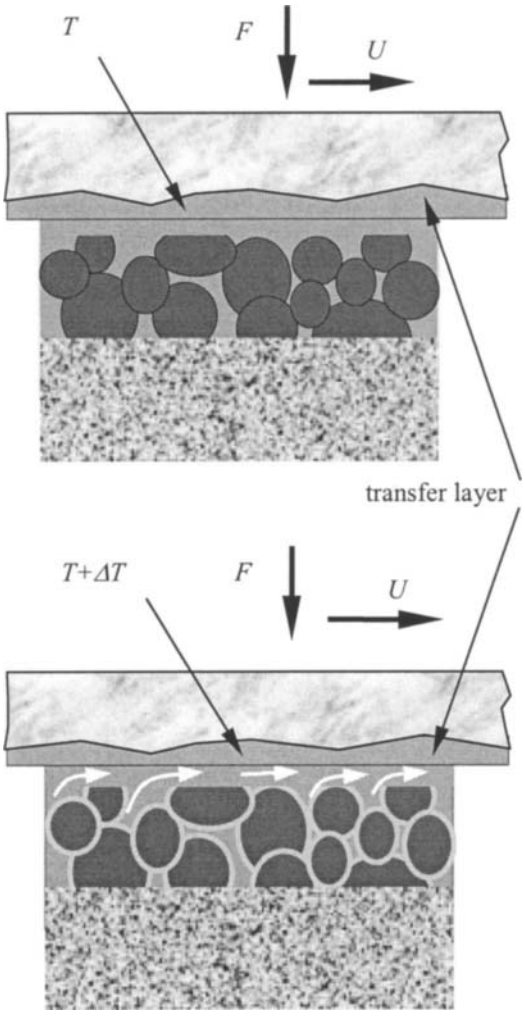


Figure 2.3. PTFE+SL displaced as a result of temperature rise

In most of the applications - also in the stationary operating state under examination- it is this mechanism supplying the bearing with the lubricant. Each temperature rise is linked with a displacement of PTFE+SL.

3. pU DIAGRAM

Today the design of dry rubbing bearings is most frequently based on the pU value. However, the pU value cannot promote the understanding of the function of such bearings. Below it is explained how pU diagrams are drawn up and which problems are resulting therefrom.

3.1. Generating a pU diagram

The first stage for generating a pU diagram is the determination of the static value of the loading limit, i.e. the determination of $\overline{p}_{max}$ at $U = 0$. After that, as described by LANCASTER [2], the specific surface pressure $\overline{p} = F/(B \cdot D)$ is gradually increased from zero at a number of sliding speeds U. In each loading stage the stationary state of temperature is waited for. This means that a constant operating temperature must be established for this point. If this is no longer the case at a definite amount of loading, the limit value of loading capacity for this sliding speed will be reached.

The interconnection line of all pU limit values [$(pU)_{lim}$ values] results in the pU limiting curve.

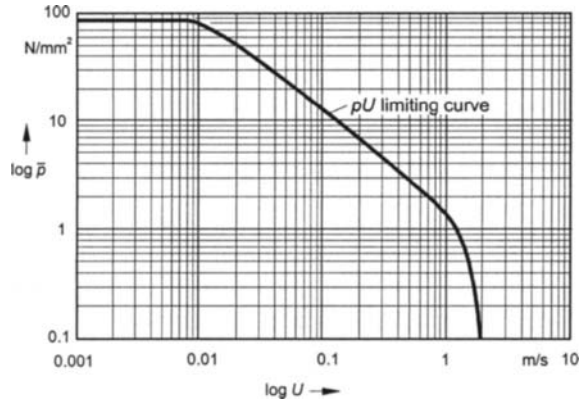


Figure 3.1. pU diagram of a dry rubbing bearing (example)

3.2. Problems in conjunction with the use of pU diagrams

The determination of the admissible pU value is based on the establishment of the heat balance for the bearing. The friction power P_f generated in the contact zone is equated with the quantity of heat carried off per time unit. This can be expressed by

$$P_f = f \cdot \overline{p} \cdot U \cdot B \cdot D = k_{rsl} \cdot A_{rsl} \cdot (T - T_{amb}) \qquad (3\text{-}1)$$

with the
resultant heat transmission coefficient k_{rsl} [W/(m²K)] and the
resultant heat emitting area A_{rsl} [m²].

For $T = T_{lim}$ the limit value $(\overline{p}U)_{\lim}$ may be determined from eqn. 3-1 so that

$$(\overline{p}U)_{\lim} = \frac{k_{rsl} \cdot A_{rsl}}{f \cdot B \cdot D} \cdot (T_{\lim} - T_{amb}). \quad (3\text{-}2)$$

If $(\overline{p}U)_{\lim}$ should be system-independent, the condition

$$\frac{k_{rsl} \cdot A_{rsl}}{f \cdot B \cdot D} = const. \quad (3\text{-}3)$$

has to be fulfilled. However, in most cases this requirement cannot be met.

The product of k_{rsl} and A_{rsl} is normally different from system to system. Among others, even the friction coefficient f is different from application case to application case. The friction coefficient f is a function of the load F, the material data, the amount of wear h_w and the relative bearing clearance ψ, even in case the amounts of B, D, T_{lim} and T_{amb} are equal to the values from the original test. The amount of wear h_w is among others dependent on the surface roughness R_z of the counter body and the operating temperature T. The bearing clearance is furthermore influenced by the operating temperature and the ambient bearing conditions (installation situation). The material data of bearing material like the linear coefficients of thermal expansion ($\alpha_{th;\ PTFE}$, $\alpha_{th;\ Bz}$), the modulus of elasticity (E_{Bz}, E_{PTFE}) and the Poisson's ratio (ν_{Bz}, ν_{PTFE}) are temperature-dependent as well.

The facts mentioned before make clear that a safe design of dry rubbing bearings is hardly possible by means of pU diagrams, and specifically not possible due to the fact that essential information about the system used as the basis for generating the pU diagram is missing.

Below a practicable calculation method is presented which has been based on similarity relations so that a comparison of different systems becomes possible.

4. CHARACTERISTICS OF COMPOUND BEARINGS

The polymer/solid lubricant mixture is transferred from the sliding layer into the contact zone and can exert its friction-reducing effect only when the lubricant layer is sufficiently thick. The height of the lubricant layer depends on the volume of the polymer/solid lubricant mixture V_S and the storage capacity of the opposing surface. Lubricant accumulation by the metallic counterface results from the area A_{sl} multiplied by the roughness of the mating metal surface (R_z value). The contact zone is subjected to the compressive stress $\overline{p}_C$, and the polymer/solid lubricant mixture is sheared off as a result of the relative velocity U and subjected to shear stress. The maximum stress a material can withstand is the shear strength τ_S regarding shear and the compressive strength σ_p regarding compression. On the basis of these observations the following relevance list was compiled to describe the frictional force:

$$F_f = f(\overline{p}_C, U, \tau_S, \sigma_p, E, h, V_S, A_C, A_{sl}, R_z)$$

The following parameters can be derived from the characteristic values included in the relevance list:

a.) $$A_C^* = \frac{A_C}{A_{sl}} \quad (4\text{-}1)$$

The parameter A_C^* is the ratio of the contact area A_C to the sliding area A_{sl}. For journal bearings A_C^* is equal to $\hat{\gamma}_C / (2\pi)$ with the radian measure of the contact angle γ_c and indicates the so-called „overlap ratio" of a tribological system. The overlap ratio should be as high as possible. PTFE+SL is available only to a limited extent and needs to be distributed on the overall opposing sliding surface to be able to exert its friction-reducing effect.

b.) $$V_S^* = \frac{V_S}{A_{sl} \cdot R_Z} \quad (4\text{-}2)$$

The parameter V_S^* describes the ratio of the PTFE+SL volume (running-in layer and sliding layer) contained in the bearing material to the volume resulting from the possible lubricant storage by the mating metallic surface. This volume needs to

be filled up by the PTFE+SL contained in the compound bearing material. The lubricant accumulation by the mating surface is approximately proportional to the product of the sliding area A_{sl} and the roughness R_z of the opposing surface and for journal bearings V_S^* can be determined by the equation $V_S^* = V_S / (\pi \cdot B \cdot D \cdot R_z)$.

One precondition for a smooth operation of a bearing built up with a compound bearing material is: $V_S^* > 1$. This ratio should be as high as possible. To need only little PTFE+SL, large sliding areas require a better surface quality (lower roughness). As a result the volume to be filled up remains low. However, it must be noted that roughnesses which are too low may have a negative impact on the formation of a transfer layer.

c.) $$\varepsilon_p = \frac{\overline{p}_C}{E(T)} \qquad (4\text{-}3)$$

The parameter ε_p is the quotient of the mean contact pressure $\overline{p}_C$ and the temperature-dependent Young's modulus $E(T)$ of the polymer/solid lubricant layer and can be calculated by $\varepsilon_p = F / [B \cdot D \cdot \sin(\gamma_C / 2) \cdot E(T)]$ for journal bearings. By analogy to the Hooke's law, this quotient indicates a compressive strain which can only increase to a critical value of $\varepsilon_p \leq \varepsilon_{p,crit}$. The critical compressive strain is determined by means of tests.

d.) $$\tau_f^* = \frac{\tau_f}{\tau_S(T)} = \frac{F_f}{\tau_S(T) \cdot A_C} \qquad (4\text{-}4)$$

The parameter τ_f^* expresses the ratio of the frictional shear stress to the shear strength $\tau_s(T)$ of the polymer/solid lubricant mixture and for journal bearings the parameter τ_f^* can be established by $\tau_f^* = 4 \cdot T_f / [\tau_s(T) \cdot \hat{\gamma}_C \cdot B \cdot D^2]$ with the friction torque T_f. As the allowable shear stress of the polymer/solid lubricant layer mainly depends on the shear strength, this parameter is essential. Below it is referred to as the shear stress ratio τ_f^*. When the shear stress ratio $\tau_f^* \geq 1$, the shear exceeds the strength of the PTFE+SL-mixture and destroys it. As a result the friction-reducing effect is diminished which becomes evident by an increase in friction force and wear. Figures 4.1 to 4.3 show this behaviour for various material combinations and testing arrangements. As can be clearly seen, the friction coefficient of compound bearings increases when the limit value of $\tau_f^* = 1$ is exceeded.

The friction force can be calculated when the shear stress ratio is known.

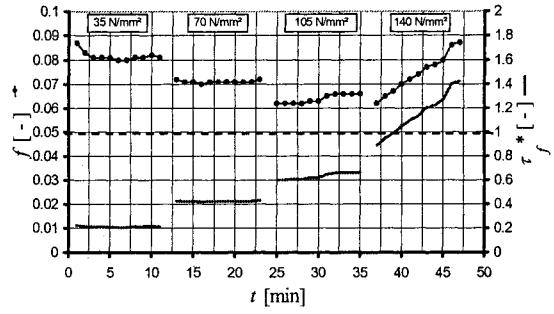


Figure 4.1. Progressive loading test on a pin-on-disk tribometer with U = 0.19 m/s; pin: compound material (sintered bronze/PTFE with lead additive); disk: steel

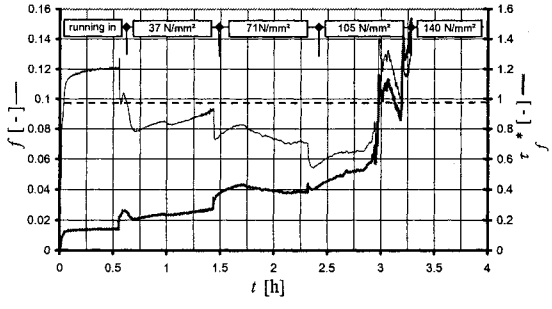


Figure 4.2. Progressive loading test with a journal bearing with U = 0.045 m/s; ψ = 2.28 ‰; bearing: compound material (sintered bronze/PTFE with lead additive); shaft: steel

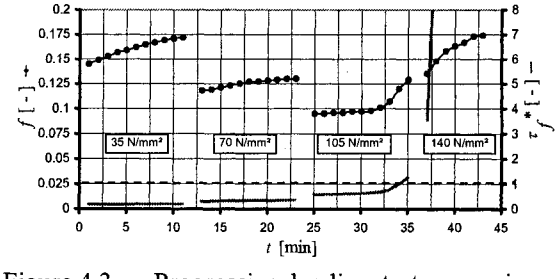


Figure 4.3. Progressive loading test on a pin-on-disk tribometer with U=0.19 m/s; pin: compound material (sintered bronze/POM); disk: steel

5. DETERMINATION OF THE SHEAR STRESS RATIO

Tribometer tests and tests on a bearing test machine with original bearings were performed and evaluated to establish a connection between the parameters described above. The following results were obtained [3]:

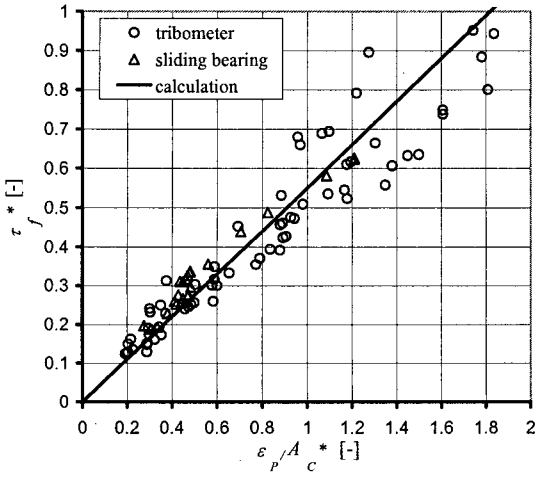


Figure 5.1. Shear stress ratio τ_f^* as a function of the ε_p-to-A_C^* ratio

The shear stress ratio is proportional to the quotient of compressive strain and overlap ratio. The proportionality factor established was 0.55.

$$\tau_f^* = 0.55 \cdot \frac{\varepsilon_p}{A_C^*} \qquad (5\text{-}1)$$

Hence, it is possible to compare the parameters determined by means of various testing units. Strikingly, the values determined on the tribometer exhibit the same relations as those of the journal bearing.

6. TEMPERATURE-DEPENDENT MATERIAL VALUES OF PTFE

6.1. Temperature-dependent Young's modulus of PTFE

The temperature-dependent Young's modulus of PTFE is a function of the degree of crystallinity of the thermoplast; also refer to DOMININGHAUS [4]. The Young's modulus at a degree of crystallinity of 60% was chosen. It can be approximated using the exponential function

$$E(T) = 23500 \cdot e^{-0.012 \cdot (T[°C]+273.15)} \text{ [N/mm²]}. \qquad (6\text{-}1)$$

The temperature T in °C is used to determine the Young's modulus in N/mm². Figure 6.1 shows a comparison between results reported in the literature and the results obtained from eqn. 6-1.

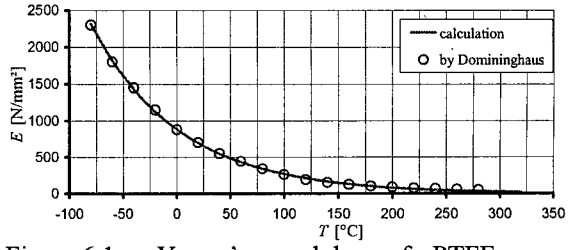


Figure 6.1. Young's modulus of PTFE as a function of the temperature

6.2. Temperature-dependent shear strength of PTFE

For the temperature-dependent shear strength of PTFE a continuous function course could not be found in the relevant literature. Only some isolated values are reported for various temperatures. Hence, measurements were performed to study the dependence of shear strength on temperature. The results of measurements were evaluated by means of a regression calculation and, hence, the temperature-dependent shear strength of PTFE can be expressed by means of the following equation:

$$\tau_S(T) = 1{,}24 \cdot e^{\frac{771}{T[°C]+273.15}} \quad \text{[N/mm²]} \qquad (6\text{-}2)$$

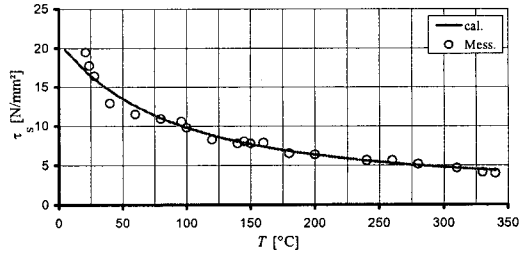


Figure 6.2. Shear strength of PTFE as a function of the temperature

As a result, the temperature-dependent material properties required for calculations are available. They can be used to determine the corresponding parameters to serve as a basis for discussing the following observations.

7. DETERMINATION OF THE CONTACT ANGLE

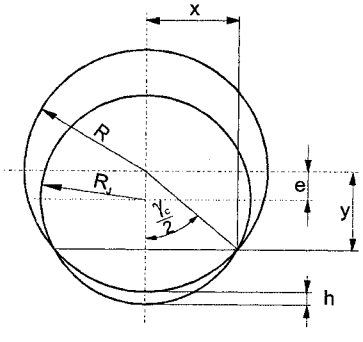
7.1. Contact angle as a result of deformation and wear

Following the approach as proposed by WENGER [5], the contact angle can be calculated on the basis of deformation $\gamma_{C,D}$ and wear $\Delta\gamma_{C,w}$ as follows:

$$\gamma_{C,D} + \Delta\gamma_{C,w} \approx 2 \cdot \arccos\left(\frac{\psi}{\psi + \zeta_{w,v}}\right) \text{ [deg]} \qquad (7\text{-}1)$$

with the relative bearing clearance $\psi = (R - R_J)/R$ and the relative shaft displacement $\zeta_{w,v}$.

It is assumed here that (i) the shaft is not subject to any deformation, (ii) shaft displacements as a result of deformation and wear can be superimposed, and (iii) the resulting shaft displacement in the direction of force is lower than half of the bearing clearance (Figure 7.1).



Trigonometric equations:

Bearing:
$x^2 + y^2 = R^2$

Shaft:
$x^2 + (y - e)^2 = R_J^2$

$h = h_{w,D} + h_w$

Figure 7.1. Circular equation approach as suggested by WENGER [5]

The relative shaft displacement is determined with the help of the relation

$$\zeta_{w,v} = \frac{2 \cdot (h_{w,D} + h_w)}{D} \qquad (7\text{-}2)$$

where h_w is the shaft displacement caused by wear. The shaft displacement caused by deformation $h_{w,D}$ results from

$$h_{w,D} = \left[R \cdot \cos\left(\frac{\gamma_{C,D}}{2}\right) - R + R_J\right] - \left[R^2 \cdot \cos^2\left(\frac{\gamma_{C,D}}{2}\right) - R^2 + R_J^2\right]^{\frac{1}{2}} \qquad (7\text{-}3)$$

with the deformation induced contact angle $\gamma_{C,D}$ which was examined in tests in dependence of load, bearing clearence and different PTFE/solid lubricant mixtures and can be calculated by the relationship

$$\gamma_{C,D} \approx 235 \cdot \sqrt[4]{\frac{F}{B \cdot D \cdot E_{rsl} \cdot \psi}} \approx 235 \cdot \sqrt[4]{\frac{\overline{p}}{\psi \cdot E_{rsl}}} \quad \text{[deg]}. \qquad (7\text{-}4)$$

with F in N, B and D in mm, $\overline{p}$ in N/mm² and E_{rsl} in N/mm². The resulting Young's modulus E_{rsl} can be obtained by using

$$\frac{1}{E_{rsl}} = \frac{1}{2} \cdot \left(\frac{1 - \nu_J^2}{E_J} + \frac{1 - \nu_B^2}{E_B}\right) \qquad (7\text{-}5)$$

with the Poisson's ratio of the journal ν_J and the bearing ν_B and the Young's modulus of the journal E_J and the bearing backing E_B.

7.2. Contact angle growth as a result of temperature rise

Results which were obtained from the contact angle and bearing temperature measurements performed on dry rubbing journal bearings revealed the following connection between the temperature-induced contact angle growth, the sum of thermal expansions of bronze and PTFE+SL and the relative bearing clearance:

$$\Delta\gamma_{C,\Delta T} \approx 210 \cdot \frac{\left(\sum \alpha_{th,S} \cdot \Delta T\right)^{0,94}}{\psi^{0,44}}$$

$$\approx 210 \cdot \frac{\left[\left(\alpha_{th,PTFE} + \alpha_{th,bronze}\right) \cdot \Delta T\right]^{0,94}}{\psi^{0,44}} \quad \text{[deg]} \quad (7\text{-}6)$$

with the thermal coefficent of expansion of PTFE $\alpha_{th,PTFE}$ and bronze $\alpha_{th,bronze}$ in 1/K and the temperature difference ΔT between the bearing temperature T and the ambient temperature T_{amb} in K. Hence, the total contact angle γ_C (Figure 7.2) is

$$\gamma_C \approx \gamma_{C,D} + \Delta\gamma_{C,w} + \Delta\gamma_{C,\Delta T}$$

$$\approx 2 \cdot \arccos\left(\frac{\psi}{\psi + \zeta_{w,\nu}}\right)$$

$$+ 210 \cdot \frac{\left(\sum \alpha_{th,S} \cdot \Delta T\right)^{0,94}}{\psi^{0,44}} \quad \text{[deg]} \quad (7\text{-}7)$$

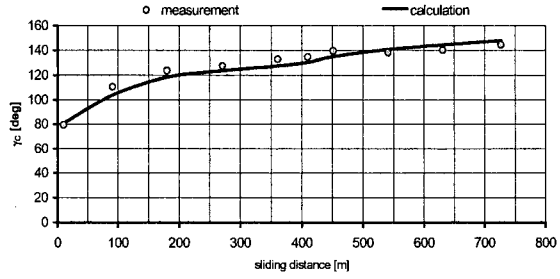


Figure 7.2. Behaviour of the total contact angle over the sliding distance (example)

8. TEMPERATURE CALCULATIONS

The frictional heat generated (frictional power) is removed from the contact zone by the shaft and the bearing bush and emitted to the environment through heat transfer and/or heat radiation. In most of the applications, the carrying-off of heat consists of a convective part and a radiation part. The influence exerted by heat radiation is quite significant, however difficult to measure in most of the cases. The radiation part is included in the approach when the empirically determined heat transfer coefficients k_i are used.

The heat balance can be descibed by the equation

$$P_f = F_f \cdot U = \left(k_1 \cdot A_1 + k_2 \cdot A_2\right) \cdot \left(T - T_{amb}\right)$$
$$= k_{rsl} \cdot A_{rsl} \cdot \left(T - T_{amb}\right)$$
$$= W_{rsl} \cdot \left(T - T_{amb}\right) \quad (8\text{-}1)$$

with

k_i heat transition coefficient [W/(m²K)]
A_i heat emitting surface area [m²]
T bearing temperature [K]
T_{amb} ambient temperature [K]
k_{rsl} resultant heat transition coefficient [W/(m²K)]
A_{rsl} resultant heat emitting surface area [m²]
W_{rsl} resultant heat conductance [W/K]

The $k_i \cdot A_i$ values are calculated following the common heat transfer approaches (also refer to DUBBEL [6], BAEHR/STEPHAN [7]) and are referred to as thermal conductivity coefficients W_i below. Another approach is based on establishing the W_{rsl} value through measurements.

Then, we obtain for the bearing temperature:

$$T = \frac{P_f}{W_{rsl}} + T_{amb} \quad (8\text{-}2)$$

9. CALCULATION OF FRICTION

The following equation can be used to determine the friction torque T_f of a journal bearing when the bearing geometry, the contact angle, the shear strength of the PTFE/solid lubricant mixture and the shear stress ratio are known.

$$T_f = \frac{1}{4} \cdot \tau_f^* \cdot \tau_S(T) \cdot \hat{\gamma}_C \cdot B \cdot D^2 . \quad (9\text{-}1)$$

As both the shear stress ratio and the contact angle as well as the shear strength are a function of the temperature and, consequently, also depend on friction and the prevailing heat transfer conditions, friction and temperature must be calculated iteratively. The calculation method is depicted in Figure 9.1.

Given:
F ; U ; A_{sl} ; B; D; T_{start}; $\alpha_{th,PTFE}$; $\alpha_{th,bronze}$; ψ ; W_{rsl}; *Lim*

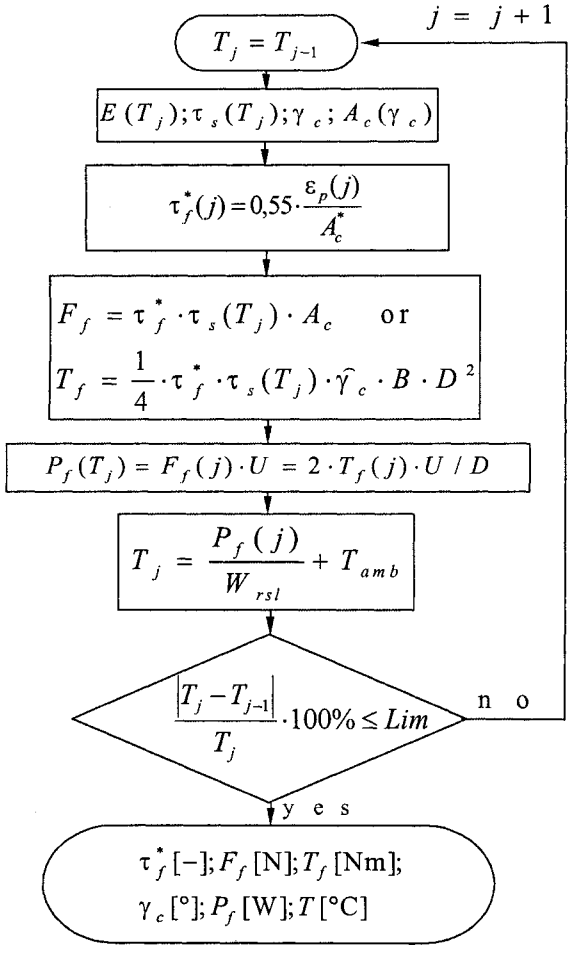


Figure 9.1. Calculation sheme for friction and temperature

10. WORKING LIFE CALCULATION

Not only friction forces but also working life is essential when designing a bearing. The calculation below follows the „basic energetic wear equation" as proposed by FLEISCHER [8]. It is based on the „wear-specific friction work" w_f. This characteristic is an expression of the friction work input into the friction material with reference to the worn volume. For the material combination examined here the wear volume is calculated only from the worn-off volume of the bearing material. The wear-specific friction work is calculated as follows:

$$w_f = \frac{\int_0^{s_f} F_f \, ds}{V_w} = \frac{W_f}{V_w}. \qquad (10\text{-}1)$$

Proceeding from a constant friction force F_f resulting in a linear increase in the friction work W_f with the friction path s_f, we can write that

$$w_f = \frac{W_f}{V_w} = \frac{F_f \cdot s_f}{V_w} = \frac{F_f}{A_C} \cdot \frac{s_f}{h_w} = \frac{\tau_f}{I_h} \qquad (10\text{-}2)$$

With I_h the linear wear intensity (also referred to as the wear-to-sliding distance ratio) is included. Transforming eqn. 10-2 yields

$$\tau_f = w_f \cdot I_h = \tau_f^* \cdot \tau_S(T). \qquad (10\text{-}3)$$

This equation shows that the wear intensity I_h is proportional to the shear stress τ_f generated by the friction force. The proportionality factor is the „wear-specific friction work" w_f.

To facilitate a non-dimensional presentation, it is helpful to transpose eqn. 10-3 to the wear intensity I_h:

$$\frac{\tau_f}{w_f} = I_h = \frac{h_w}{s_f} = \frac{h_w}{U \cdot t}. \qquad (10\text{-}4)$$

If the „wear-specific friction work" w_f is known, working life can be calculated as a function of the remaining parameters with the given $h_{w,crit}$ upon re-arranging eqn. 10-4 to the time t.

$$t_{crit} = \frac{w_f \cdot h_{w,crit}}{\tau_f \cdot U} = \frac{w_f \cdot h_{w,crit}}{\tau_f^* \cdot \tau_S(T) \cdot U} = L_H \qquad (10\text{-}5)$$

The critical wear height $h_{w,crit}$ can be determined by means of measurements. However, it must be noted that the lubricant (PTFE+SL) is available in the compound material to a limited extent only. As soon as it is used up, the bearing suffers from a lack of lubricant. Then, the friction force increases, wear intensity grows, the bearing is overheated by friction and, finally, fails. Hence, the total sliding layer height cannot be used as the allowable wear height. In journal bearings the critical wear height is determined through the wear height in the direction of

load. Consequently, the approach applies equally for thrust and journal rubbing bearings. The sliding velocity U is defined by the operating parameters and the frictional shear stress τ_f can be calculated with the help of the shear stress ratio. To this end, τ_f^* is multiplied by the temperature-dependent shear strength. The connection between the shear stress ratio and the operating parameters is discussed in Section 5. The wear-specific friction work w_f is obtained from measurements and, similarly to the friction work, depends on the overall system. Wear measurements on a tribometer yielded values for w_f of about $2.5 \cdot 10^5$ Nm/mm³ for compound materials with PTFE.

As can be seen from eqn. 10-5 working life is reciprocally proportional to the sliding velocity. High sliding velocities result in a short working life when they do not exert an essential influence on τ_f and w_f.

11. WEAR PARTICLE TRAP

One of the causes for failure of dry rubbing journal bearings is that wear particles remain in the bearing and the contact angle reaches a value > 180°, resulting in jamming (Figure 11.1). If the contact angle is limited by grooves, bores, pockets, etc. where wear debris is accumulated without exerting any pressure, jamming and overheating can be avoided (Figure 11.2).

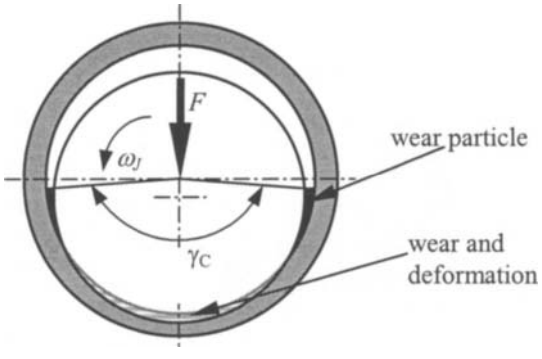


Figure 11.1. Enlargement of the contact angle by wear and wear debris



Figure 11.2. Bearing with bores as wear particle trap to limit the contact angle

Tests have revealed that, in contrast to a bearing without wear debris traps and consequently without any contact angle limitation, the friction torque remains constant although the bearing clearance is reduced by fifty percent. Such bearings exhibit a stationary operating state and a higher guiding accuracy than conventional bearings. In general, the bearing clearance there is increased to avoid jamming.

The volume of the wear particle trap V_R must be designed such that the wear particles from the overall contact zone can be collected. As a rule, the volume can be calculated from the radian measure of the contact angle, multiplied by the bearing width and the permissible wear height.

$$V_R \approx \hat{\gamma}_C \cdot \frac{D}{2} \cdot B \cdot h_{w,crit} \qquad (11\text{-}1)$$

In order not to reduce the load-bearing capacity of the bearing, the remaining bearing surface angle should be at least of the same size as the contact angle resulting from the deformation by external loading and wear.

12. CALCULATION EXAMPLES

12.1. New bearings

The following details are known about the location with a dry rubbing journal bearing with a PTFE/ solid lubricant mixture:

n = 50 min^{-1}
W_{rsl} = 0.6 W/K
D = 20 mm $\Rightarrow U$ = 0.052 m/s
B = 15 mm
F = 10 000 N $\Rightarrow \overline{p}$ = 33.4 N/mm^2

Recommendation for the relative bearing clearance: Ψ = 1.5 up to 6 ‰

Selected: Ψ = 2.5 ‰
$D_J = D\,(1-\Psi)$
= 20 mm (1 – 0.0025)
= 19.95 mm

In the following it is established whether the bearing is duly dimensioned and what the expected life and/or the resultant sliding distance are.

Required quantities:

τ_f^*	or	$1/\tau_f^*$ (Safety)
L_H [h]	or	s_f [m]

Bearing temperature, friction torque, contact angle and safety factor ($1/\tau_f^*$) are calculated following the iteration algorithm as shown in Figure 8.1. First, calculations are performed without including the influence exerted by wear (h_w = 0). As an initial value for the temperature-dependent values ($E(T)$, $\tau_S(T)$, γ_C) room temperature is used in the beginning (T_0 = 20 °C).

The iteration process without considering any influence by wear yields:

Bearing temperature	T	78 °C
Friction torque	T_f	6.6 Nm
Contact angle	γ_C	145 °
Safety	$1/\tau_f^*$	1 / 0.157 = 6.4

12.2. Worn bearings

In the next step a permissible wear height or a permissible shaft displacement caused by wear respectively of $h_{w,crit}$ = 0.1 mm is included, resulting in an additional increase in the contact angle γ_C. The results of this computation show the bearing behaviour at the end of the bearing life.

From eqn. 7-4 we obtain a value for the load-dependent contact angle at the initial temperature (T_0 = 20 °C) of $\gamma_{C,D}$ = 118 ° and from eqn. 6-3 the shaft displacement caused by deformation of $h_{w,D}$ = 0.02371 mm.

Hence, the relative shaft displacement of the bearing according to eqn. 7-2 is $\xi_{w,v}$ = 0.012371.

Considering wear, the total contact angle at the initial temperature T_0 = 20 °C according to eqn. 7-7 is γ_C = 161 °.

As shown in Figure 9.1, the iteration process is repeated with this contact angle. This procedure yields the following results for the usable life of the bearing (including the results of Section 12.1):

		New	*Worn*
Bearing temperature	T	78 °C	74 °C
Friction torque	T_f	6.6 Nm	6.2 Nm
Contact angle	γ_C	145 °	186 °
Safety	$1/\tau_f^*$	6.4	9

Throughout its usable life the bearing is operated within these limits. If the contact angle achieves a value > 180°, a bearing may fail through jamming. Hence, it might happen that such a bearing jams, is overheated by friction and, consequently, fails even before it has reached the permissible wear depth. Therefore, calculations of the life of bearings include uncertainties and should be based, for safety reasons, on the highest shear stress ratio.

Calculation of the life of the bearing following eqn. 10-5 gives L_H = 275 600 s = 76.6 h. The sliding distance results in $s_f = L_H \cdot U$ = 14,331.2 m.

12.3. Bearing with contact angle limitation

This section is aimed at demonstrating how life and the guiding accuracy can be increased by including a contact angle limitation with the help of a wear particle trap. The input values given in Sections 12.1 and 12.2 remain the same, apart from the relative bearing clearance. This will be reduced to Ψ = 1.5‰.

From eqn. 7-4 we obtain the load-dependent contact angle at an initial temperature of T_0 = 20 °C for a bearing with contact angle limitation $\gamma_{C,D}$ = 134 ° and from eqn. 7-3 the deformation influenced shaft displacement of $h_{w,D}$ = 0.02355 mm. Hence, the relative wear of the bearing according to eqn. 7-2 is $\xi_{w,v}$ = 0.012355.

Therefore, the total contact angle at an initial temperature of $T_0 = 20$ °C according to eqn. 7-1 is $\gamma_C = 168$ °.

The contact angle is to be limited to this value. Hence, the contact angle remains constant during iteration. Iteration as shown in Figure 9.1 yields the following results:

Bearing temperature	T	74 °C
Friction torque	T_f	6.2 Nm
Contact angle	γ_C	168 °
Safety	$1/\tau_f^*$	8.1

Calculation of the bearing life according to eqn. 10-5 yields $L_H = 342\,823$ s = 95.2 h. For the sliding distance we obtain $s_f = L_H \cdot U = 17{,}826$ m.

13. SUMMARY

As established, the performance of the compound bearing material is mainly influenced by the shear strength of the polymer/solid lubricant mixture. When the friction-induced shear stress exceeds the shear strength of the PTFE/solid lubricant mixture, the mixture is subjected to overstress and shears off. This causes an increase in friction and, consequently, wear. After a short period the bearing fails. Hence, the shear strength of the polymer/solid lubricant mixture marks the upper load limit of such bearings. This limit is reached at a shear stress ratio of $\tau_f^* = 1$. Irrespective of the testing facilities this parameter is valid and applies equally to both linear and journal bearings.

Due to the special contact-geometrical features of dry rubbing journal bearings - as opposed for example to thrust bearings - wear and wear particles cause permanent changes in the contact geometry within the 360° journal bearing. Hence, dry rubbing bearings cannot reach a stationary operating state. In addition, accumulation of wear particles in the bearing can lead to jamming and overheating.

If grooves, pockets, bores, etc. are included in the design to remove the wear particles and, hence, limit the contact angle, this angle remains constant and dry rubbing bearings may be improved with regard to their loading capacity and guiding accuracy.

Knowledge of the contact angle is essential to the calculations of dry rubbing bearings. This angle is required for determining the shear stress ratio which is not only indicative of the upper load limit of such bearings, but can be also used for calculating friction forces and/or friction torque. The shear stress ratio of various bearings can be determined from the compressive strain and the overlap ratio.

Calculating friction in a dry rubbing bearing is only possible iteratively, giving due consideration to all properties and parameters involved. Friction, temperature, material properties and the prevailing heat transfer conditions influence each other.

REFERENCES

[1] WALDHÜTER, W.: Selbstschmierende und wartungsfreie Gleitlager, Band 422. Verbundgleitlager-Werkstoff-System für den selbstschmierenden und wartungsarmen Betrieb von Gleitlagern S. 235-276. Ehningen bei Böblingen: Expert, 1993

[2] LANCASTER, J.K.: Dry bearings: a survey of materials and factors affecting their performance. Trobology, Dec. 1973, 219-251

[3] BERGER, M.: Untersuchungen an wartungsfreien trockenlaufenden Verbundgleitlagern. Aachen: Shaker, 2000

[4] DOMININGHAUS, H.: Die Kunststoffe und ihre Eigenschaften, Düsseldorf: VDI, 1992

[5] WENGER, E.:Bestimmung des Lagerspiels von Kunstharzpreßstofflagern unter Berücksichtigung der Verformbarkeit des Werkstoffes. Stahl und Eisen 74 (1964), S. 1202-1212

[6] DUBBEL, H.:Taschenbuch für den Maschinenbau. Berlin/Heidelberg: Springer, 1990

[7] BAEHR, H.D.; STEPHAN, K.: Wärme- und Stoffübertragung. Berlin-Heidelberg: Springer, 1994

[8] FLEISCHER, G.; GRÖGER, H.; THUM, H.: Verschleiß und Zuverlässigkeit. Berlin: Technik, 1980

SESSION VI
TRIBOLOGICAL STUDIES OF DRY AND LUBRICATED SURFACES (2) - DRY AND LUBRICATED

Chairman : Professor G Dalmaz

Paper VI (i) **Tribological Design Criteria and Standards for Spacecraft Mechanisms**

Paper VI (ii) **A Simple Numerical Procedure to Calculate the Input Data of Greenwood-Williamson Model of Asperity Contact for Actual Engineering Surfaces**

Paper VI (iii) **The Study of Self-lubricating Sintered Nickel Matrix Material and its Applications**

Paper VI (iv) **Behaviours of Ultra-thin Lubricant Films Between Solid Surfaces Having Atomic Scale Roughness**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribological Design Criteria and Standards for Spacecraft Mechanisms

R A Rowntree

Engineering Solutions, AEA Technology
410 The Quadrant, Birchwood, Warrington, Ches., WA3 6AT, United Kingdom

The tribologist is called upon to solve friction and wear problems in many diverse environments. One of the more unusual is vacuum and, more specifically, the vacuum of space, where modern spacecraft bristle with moving components. This invited paper describes a recently developed Space Engineering Mechanism Standard (ECSS-E-30, Part 3), that includes a significant number of mandatory tribological criteria to ensure the continued reliability of newly developed hardware.

1. INTRODUCTION

The design of mechanisms used for the basic functions of a spacecraft continues to evolve. Given the low numbers of spacecraft that have the same or similar design, or which use the same mechanisms, the competitive nature of the space mechanism procurement process and ever changing technical specifications, then there is a continual process of new mechanism design and supply. Achieving reliable mechanism design, under these conditions, is therefore not a trivial task.

To this end, the European Space Agency, through its ECSS programme (European Co-operation for Space Standardisation) have developed Space Engineering Standards (Figure 1). One of these relates to Mechanisms (ECSS-E-30, Part 3) and this includes a significant number of mandatory tribological criteria [1]. This specific standard has support documentation ('Informative References') of a Space Tribology Handbook [2] and a Tribometer User's Guide for Space Mechanism Applications [3] both produced by the European Space Tribology Laboratory (ESTL), also part of AEA Technology. It is intended, ultimately, to incorporate both documents in the ECSS system.

The purpose of this invited paper is to give an overview of the tribological design criteria and principles contained within this design standard and supportive documents, and to look at some of the general issues that arise, not just spacecraft-specific, when put into practice.

2. SPACE MECHANISM TYPES

There are a surprisingly large number of mechanisms with moving surfaces on modern spacecraft. These will be briefly reviewed here.

Following launch the first set of mechanisms operated are those involved in the deployment of antennae, booms, solar arrays etc. Deployment normally occurs very soon after launch to ensure that communication and power are established as rapidly as possible. These are invariably 'one-shot' devices, which may have been tested in excess of 50 times before launch. Key features are the separation of latches, launch locks and clamped interfaces without adhesion. Mechanisms range from simple springs to pin-pullers, memory alloy elements, cable cutters, thermal knives, powered hinges, DC motor gearheads, dampers of various types etc.

The next major mechanism grouping is that involved in the attitude control of the spacecraft. To achieve 3-axis stabilisation, a combination of inertial and thruster control is used. Inertial devices are momentum wheels (fixed speed, typically up to 5000 rpm) or reaction wheels (variable speed, typically $\pm$3000 rpm), both are usually linked to gyroscopes (typically speed 10,000 rpm, but used intermittently). Inertial devices are required to continuously operate up to 15 years ($>10^{10}$ revs) with near constant torque behaviour. Bearings, cage stability and lubrication are the key tribological issues. Thrust is also achieved by mono- or bi-propellants controlled by valves. The sealing of

valves to both vacuum and propellant, with very low leak rates, is mandatory. For certain science applications spin stabilisation is preferred. Under these conditions, for communication back to Earth or to point instruments, a despin mechanism is used. These devices rotate typically between 5 to 60 rpm for up 15 years. Bearings and rotating electrical transfer devices are the main tribological elements.

The third major grouping of mechanisms is associated with power generation. Solar array drives are used to rotate the arrays to be normal with respect to the sun, and transmit electrical power. Rotating star sensors and sun trackers are also used to help orientate the solar arrays. Whilst lifetimes are again 7 to 15 years, solar array drives rotate at 1rev/day, hence total bearing revs is usually <5000 (strictly for geostationary orbits). Star & sun sensors rotate typical at 200 to 400 rpm hence lifetimes are $>10^9$ revs. Gears and sliding electrical contacts are also employed in solar array drives.

The final set of major mechanisms is those associated with the payload. For communications this involves antenna pointing devices and RF switching units. For remote sensing and science applications this includes radiometers, scanners, telescopes, tape recorders, filter wheels, beam splitters, shutters and instrument pointing devices of all forms (gimbals, rotating wedges, linear actuators etc) are used. The types of tribological element used are only limited by the imagination of the design engineer; the more common include bearings (rotational & linear; ball, roller & bush); gears (external & internal, spur, planetary, worm & harmonic); pivots; threaded, ball or roller screws; friction drives & brakes.

3. SPACE ENVIRONMENT

Any discussion on spacecraft tribology usually focuses on the thermal vacuum environment of space. In practice all mechanisms have to be able to operate both on ground (ie in humid air and during all ground testing) and in vacuum, whilst clearly surviving the launch environment. Many reviews [4 -7] define the vacuum environment in detail, however few seem to fully recognise the tribological difficulties. Operation in air or in vacuum is relatively straightforward, but operation in both with >99% reliability, after periods of sometimes long storage, followed by extreme vibration and within specified margins is the challenge.

The vacuum and residual species experienced by a spacecraft mechanism depends on where it is on the spacecraft (internal or external), where the spacecraft is (low earth orbit, geostationary orbit, or deep space) and the pollution of the local environment by the spacecraft (from outgassing of the materials used and thruster firings). The minimum stable orbit for spacecraft is around 200 km, with a background pressure of around 10^{-6} torr, whilst in geostationary orbit, at 36000 km, the background pressure is about 10^{-13} torr. However, most spacecraft mechanisms are internal, hence shielded and partially sealed.

In orbit, therefore, the partial pressure of oxygen is typically $<10^{-10}$ torr. Thus ruptured surface oxide films, arising from relative motion of contacting surfaces, are not repaired by either physical adsorption or chemical reaction. Progressive oxide film attrition can rapidly produce high friction and severe wear, for materials that would exhibit mild wear and low friction on ground. The function of a space lubricant is clearly to prevent high wear, and to give low and predictable friction.

In the authors view the practical problems of space tribology include:

- It has to work in all three environments
- Lack of appreciation of lubricant characteristics and properties, and of their limitations.
- Assumption of constant friction
- Limited range of lubricants
- Long process of qualification or acceptance of new lubricant solutions or other new surface treatment.
- Inadequate or unrepresentative testing
- Relubrication is difficult, but not impossible

4. SPACE ENGINEERING STANDARD ECSS-E-30 Part 3A

This Standard is one of a series of ECSS Standards intended to be applied together for the management, engineering and product assurance of space projects. It is a co-operative effort of the European Space

Agency, national agencies and European industrial associations.

ECSS-E-30 defines the mechanical engineering requirements, with Part 3A for mechanisms. The latter contains a mandatory list of requirements applicable to concept definition, design, analysis, development, production, test verification and in-orbit operation of mechanisms to meet mission requirements. The main tribological criteria are contained within sections in the Standard on Design, and on Verification. These will now be considered in more detail.

4.1 Design

These are further sub-divided into the following subject areas:

- *General tribology requirements*, including the definition of satisfactory lubrication - 'adequate lubrication function between surfaces to meet all performance requirements throughout the required lifetime and to minimise wear' (note some wear is deemed acceptable). The need for lubricants to be space approved (mainly with respect to outgassing and particulate contamination properties); the requirement that any lubricant degradation on ground or in-orbit shall not lead to any degradation in component or mechanism performance; the need for cleaning of surfaces prior to lubrication for optimum adhesion or wetting purposes; and the avoidance, where possible, of sliding surfaces. If not possible, the steps to be taken to prevent the risk of adhesion (surface hardening, use of self-lubricating or dissimilar materials.
- *Dry lubrication*, covering their selection process; deposition process validation including co-deposited witness coupons and their evaluation; application verifications to achieve stated performance, life and repeatability; and definition and documents of all aspects of handling storage, and use of dry lubricated components.
- *Fluid lubrication*, requiring justification on lubricant quantities given leakage paths, acceptable outgassing rates and, if not, means of sealing; use anti-creep barriers and means of ensuring their integrity.
- *Materials*, listing minimum surface roughness for end stops (<0.4 microns Ra); minimum hardness of metal surfaces in contact (500 HV) unless one is self-lubricating or additional lubrication is present; and their life demonstration
- *Bearing preloading*, specifically for ball bearings (the most widely used bearing system in space mechanisms), defining the magnitude and type of preload (hard or by spring); the use of dedicated sleeves if spring preloading systems are used; the checking of bearing preload after installation, and if appropriate after run-in; and checks on bearing adequacy under vibration if internal gapping occurs.

Under the same heading of Design, Mechanical Design and Sizing issues and requirements are as follows:

- *Factors of safety*, computational margins for metallic materials of 1.25 against yield, 1.5 against ultimate stress and a minimum fatigue factor of 4.
- *Motorization factors*, designed to ensure adequate driving margins are always present for a mechanism to operate throughout its operational lifetime and over its full range of travel, output torques or forces which exceed at least 2 times the combined worst case resistive torques or forces, in addition to any required deliverable output torque or force, T_L. In order to derive the factored worst case quasi-static resistive forces or torques, the components of resistance shall be multiplied by the minimum uncertainty factors given in Table1.

Table 1
Minimum uncertainty factors

Component of Resistance	Symbol	Factor
Inertia	I_T (or I_F)	1.1
Spring	S	1.2
Friction	F_R	3 (1.5 if measured)
Hysteresis	H_Y	3 (1.5 if measured)
Others (eg harness)	H_A	3 (1.5 if measured)
Adhesion	H_D	3 (1.5 if measured)

The minimum required actuation torque T_{min} is defined as:

$$T_{min} = 2\ (1.1\ I_T + 1.2\ S + 3\ F_R + 3\ H_Y + 3\ H_A + 3\ H_D) + T_L \quad (1)$$

The minimum required actuation force F_{min} can also be defined in the same way with T_L replaced with FL. When the motorization equation is applied to mechanisms which have to fulfil a specified acceleration requirement or for which an indirect acceleration requirement can be deduced, then the external torque or force is increased by a factor of 1.25.

- *Separable surfaces*, designed to minimise the adhesion forces and where the contact between mating surfaces shall be 'well characterised and reproducible' for the area of contact, alignment and load conditions (it is noted both Hertzian and non-Hertzian contacts are apparently acceptable). Avoidance of sliding and high hardness (>500 HV) of metallic surfaces (if used) are mandatory. As a check for the avoidance of yielding in a Herztian contact (with margin factor 2), the standard states that the peak Hertzian contact pressure shall be demonstrated to be <93% of the (uniaxial) yield limit of the weakest material.
- *Bearing and gear sizing*, to be in accordance with ISO76 for ball bearings and ISO 6886 for gears.

4.2 Verification

Verification is mandatory and shall comprise of both analytical and testing elements.

- *Analysis* of tolerance budgets on preload; Hertzian contact analysis for separable and sliding contacts for worst case conditions to verify selection of base materials, lubricant and any coatings employed; bearing and gear sizing, to be in accordance with ISO76 for ball bearings and ISO 6886 for gears; an assessment of the choice of lubrication system and its adequate dimensioning and quantity for the proposed application and lifetime; identification of any life limited components.
- *Test,* particularly related to the qualification or life test. In the qualification testing the adequacy of the lubrication system shall be demonstrated under a more extreme range of simulated conditions (usually thermal). In the life test the adequacy of the mechanism design, lubricant lifetime and performance shall be demonstrated under representative conditions (eg thermal, motion profile, durations, loads etc). If accelerated life testing is performed (which is normal); the conditions must be chosen to be representative of the worst case environmental conditions with respect to potential degradation modes. Thus, if liquid lubricated, there are severe constraints to acceleration factors possible, as otherwise the specific film thickness will be different between normal use and accelerated life test – a condition unacceptable to tribologists. Finally a margin is also applied to the life test durations, where the predicted number of cycles in each part of the test phase is factored cumulatively according to Table 2.

Table 2
Life Test Duration Factors

Type	Number of Predicted cycles	Factor
Ground Testing	Number of test cycles (>10)	4
In-Orbit	1 to 10	10
	11 to 1000	4
	1001 to 100,000	2
	>100,000	1.25

Lifetime acceptance criteria are also defined, where at the end of the test, following a strip examination:

- No metal to metal contact identified in the interface of solid lubricated contact surfaces.
- Surface properties of contact surfaces unmodified in its essential performance properties.
- No detrimental chemical deterioration of liquid lubricants is found.
- The amount and size of wear products conforms to contamination requirements and overall mechanism performance requirements.
- Worst case variation or degradation in the peak torque or force overall throughout the life test is compatible with the torque force ratio requirements of equation 1 and less than or equal to 50%.

4.3 In practice

It would be wrong to suggest this space mechanism design standard has cured all ills, has fully covered all mechanism design permutations or that has it been universally welcomed. But it is fundamental step in improving the design process

and at flushing out at an early stage potential tribological issues in the development of a new mechanism. The various contributors to the standard found that trying to distil years of background experience into a design standard more difficult than at first sight. For example, simply defining when a test is tribologically successful or not, is not trivial.

5. SPACE TRIBOLOGY HANDBOOK

The ECSS standard described in the preceding sections defines the requirements only. It was recognised that a source of validated and approved tribological data was also essential if the standard was to be rigorously applied. A particular concern was that this data was spread in many different documents in different locations and not readily available to the mechanism's engineer. In many cases data contained in different documents was somewhat contradictory and its validity difficult to verify because of inadequate specification of the test conditions. With support from the European Space Agency, ESTL has written and produced in 1998 a Space Tribology Handbook to meet this need, with the following aims:

- To provide a definitive reference book/manual on space tribology
- To specify design criteria for selection of tribo-materials
- To specify rules and guidelines for the selection, use and limitations of tribo-materials
- To provide useful data on typical systems where tribology is important either to performance or reliability
- To aid the designer in the selection of bearing configurations

This collects and distils in a single volume, some 30 years experience and substantiated test results in European space tribology. Regularly updated it also includes known friction coefficients for different material combinations and lubricants as functions of numerous environmental and stress conditions.



6. A TRIBOMETER USER'S GUIDE FOR SPACE MECHANISM APPLICATIONS

A space tribometer user's guideline has also been produced, since tribometers are frequently used to simulate in a simplified manner the tribo-system on the intended space application at an early stage of the mechanism design. However their use often leads to incoherent data when different test facilities involved for the 'same' test. This can reduce the confidence of the data acquired at a given test site, often requiring duplications of tests to resolve. By standardising the test parameters involved, an improvement in the quality and reproducibility of the tribological data and measurements can be achieved.

A specification of parameters influencing the friction and wear applicable to the basic classes of tribometers has been established. Concepts and approaches for using tribometers to simulate

tribological processes occurring in real space engineering applications, and the use and interpretation of the results have been outlined. The User Guide also aims to help engineers unfamiliar with tribometers to identify the device most able to solve a specific tribological problem. Finally the User Guide gives an overview of existing space tribometer test facilities and their capabilities for potential consultation by users.

7. CONCLUDING REMARKS

An approach to ever improving tribological design in spacecraft mechanisms has been presented in this paper. Based on a standards approach with mandatory requirements, with supporting validated data, ESA and European space industries seek to improve space product performance and cost-efficiency of space missions. However the application of this standard does mean more work and justifications for the design engineer.

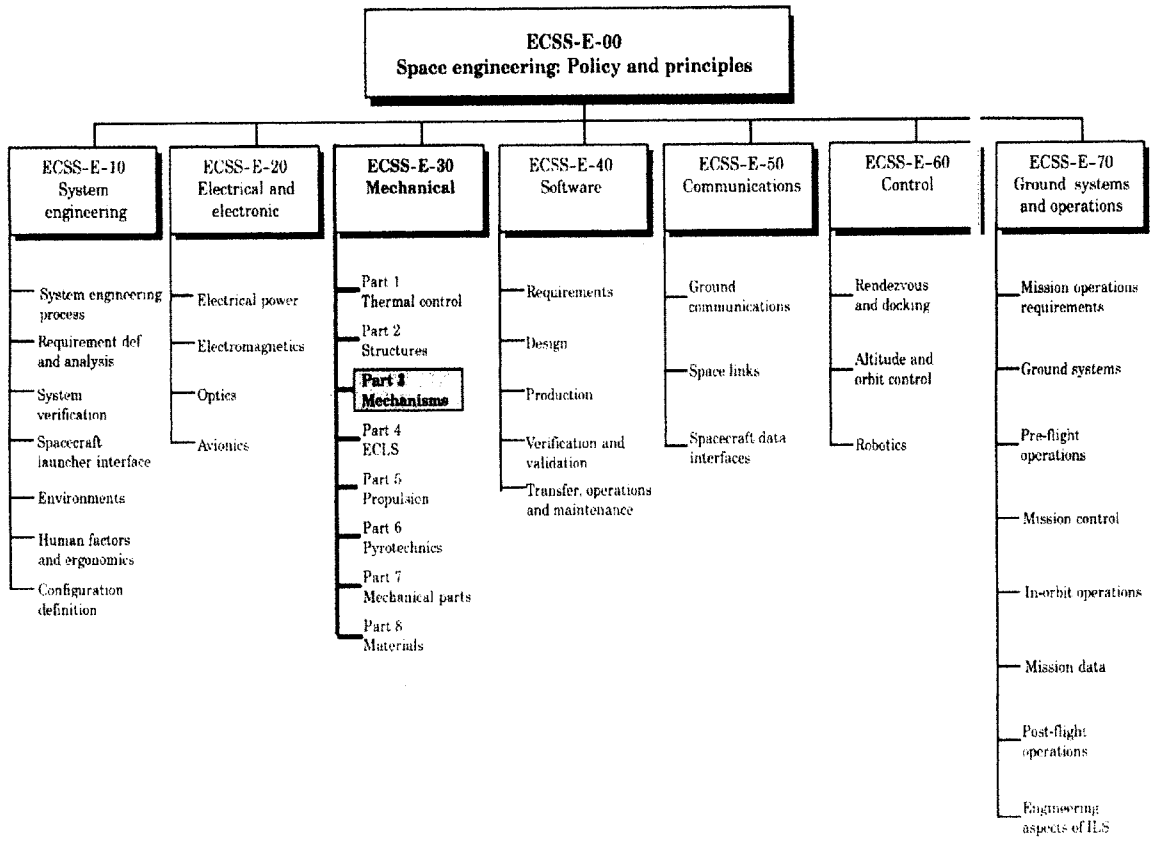
A fundamental question remains: can space tribology be reduced to a set of requirements and validated test results. In the author's view, that answer is not quite, otherwise not only is our role significantly reduced, there is the greater risk of assuming all is well because the design meets the criteria specified in the standard. Tribological design is about the getting the details correct and not assuming. Given the diversity of mechanism designs and their operational conditions, a single all embracing design standard and existing data set is unlikely.

But there is not doubt either that trying to define process and issues, had resulted in a set of common criteria which forces the mechanism designer to consider much more the tribology details and provide justifying analyses. Provided this standard is updated with current industrial successful practice and the generation of the required validated data, continued improvements in mechanism reliability, particularly for new mechanism designs, will be achieved.

REFERENCES

1. Space Engineering, Mechanical – Part 3: Mechanisms, ECSS-E-30 Part 3a, April 2000.
2. Space Tribology Handbook, ESTL, AEA Technology, 3rd edition, 2002.
3. A Tribometer User's Guide for Space Mechanism Applications, Cunningham M J, ESTL/TM/139, 1994.
4. Heaton J W. The environment of space. Inst. Phys. Conf. Serial No 34, 1977.
5. Carpenter L G. A survey of vacuum technology for satellites. Proc. Inst. Mech. Engrs. Vol 179, Part 3G, 1964-1965, pp 164-171.
6. Godwin R C & Woods W J. Properties and behaviour of materials in outer space. Proc. Inst. Mech. Engrs. Vol 179, Part 3G, 1964 – 1965, pp 172-180.
7. Patrick T J. Space environment and vacuum properties of spacecraft materials. Vacuum, Vol. 3 No. 8, 1981, pp 351-357.

Figure 1
Structure of ECSS Engineering Standard System



This Page Intentionally Left Blank

D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

A simple numerical procedure to calculate the input data of Greenwood-Williamson model of asperity contact for actual engineering surfaces

Eduardo Tomanik*, Haroldo Chacon, Giovanni Teixeira

Mahle S.A. – Brazil Tech Center

Greenwood-Williamson model for asperity contact has been extensively used for prediction of tribological behaviour. However, characterization of the necessary input data, summit asperity height deviation (σ_S), mean radius (β) and density (η) are not available from commercial roughness measuring equipment. A relatively simple numerical method is proposed to calculate the asperities data for actual surfaces. Summits were considered to be the local maximum points above the surface mean line and calculation of the standard deviation, mean radius and density of the summits are directly numerical performed. The method was applied for new and worn ICE cylinder bores. The product $\sigma_S.\beta.\eta$ was found to be between 0.01 to 0.05, lower than the usually accepted range 0.03 to 0.05. The lowest values were found on plateau honing finish, after 15 h test. For all cases, the calculated summit mean radius was found to be lower than previous published values.

Some preliminary asperity pressures calculation with the found parameters, suggest that the usual film parameter on the Stribeck curve can lead to wrong conclusions about the tribological severity. Some discussion about the expected asperity pressures of new and worn ICE cylinder bores are also presented.

1. INTRODUCTION

The Greenwood-Williamson (G-W) model [1] or the later, more realistic, Greenwood-Tripp (G-T) model [2] is extensively used to calculate micro-contact and pressures that arise when two rough surfaces approach each other. For simplification, the present work will focus on the G-W model, and it will be applied specially for piston ring/cylinder interactions, but the proposed method can be directly applied for other rough surfaces. It can also be easily extended for the G-T model. In the G-W model the rough surface is presumed to be covered with asperities spherical in shape, with constant radius β. The summits are uniformly distributed with density η of summits per area unit. One surface is considered to have the combined roughness of the two original surfaces and the other is considered to be smooth. See figure 1.

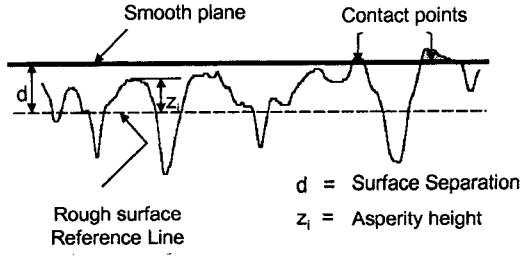


Fig. 1. Schematic contact of two surfaces according the G-W model.

The mean height of summits is located above the mean height of the surface as a whole, as indicated in fig.2. The summit heights, z_S, are assumed to follow a gaussian or normal distribution with standard deviation σ_S. If two rough surfaces are brought in contact, their asperities will start to interact, and a pressure, P_{ASP}, caused by asperities deformation will appear. P_{ASP} can be calculated by:

$$P_{ASP} = \frac{4}{3} \cdot \eta \cdot E' \cdot \sqrt{\beta} \cdot \sigma_c^{3/2} F_{3/2}(t) \quad (1)$$

* E-mail address: atomanik@mahle.com.br

$$F_{3/2}(t) = \int_{d}^{\infty} \phi(z) \cdot (z-t)^{3/2} \cdot \delta z \qquad (2)$$

Where:

σ_C: composite roughness of the 2 surfaces

β: average summit radius

η: summit density

t: $= \sigma_C/d$, d: distance between the two surfaces

E': Combined Modulus of Elasticity.

The integral in the equation (2) basically accounts for the probability of a given summit height, $z_{S(i)}$, be higher than the distance d between the surfaces. $\phi(z)$ is the probability density function, PDF, of the summit heights.

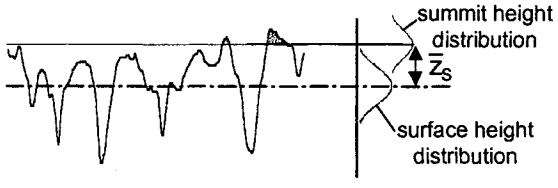


Fig. 2. Surface and summit mean planes and distributions.

Although the G-W model is being extensively used in computer programs that deal with tribological behaviour of piston rings, except for considering σ_S = Rq, there is little experimental data about the values of the two other surface parameters: β, η. The product $\sigma_S.\beta.\eta$ is usually taken to be in the range published by G-T: $0.03 < \sigma_S.\beta.\eta < 0.05$, although some recent works obtained different values. Table 1 reproduces some values of the G-W parameters found in the literature.

Arcoumanis [3] used the composite equivalent spectral moments of the surface to obtain the summit parameters. Calculation of the spectral moments from experimental data was found difficult by the authors of the present work. See section 2.

Table 1
G-W surface parameters

σ_S (μm)	β (μm)	η (1/m^2)	$\sigma_S.\beta.\eta$	Ref.
0.12			0.03 to 0.05	[1]
			0.05	[4]
0.24	66	3.9 E9	0.06	[5]
			0.04	[6]
0.2 to 1.0	5000	1.0 E9	1 to 5	[7]
0.185	275	3.1 E9	0.16	[3]
			0.63 to 1.92	[8]

Other complication to practical utilization of the G-W model is that most of the engineering sliding surfaces are non-gaussian. E.g., ICE cylinder bores are finished to have much more valleys than peaks. Fig. 3 shows two cylinder roughness profiles from large volume, SI European engines. Notice that, for the plateau honed cylinder, the highest peaks are of the same magnitude of Rq. With the usual assumption that σ_S = Rq, the G-W model would estimate that a smooth surface will start to contact the cylinder at a distance of 3Rq, much higher than the actual highest surface peak. To address the fact that the peaks are much smaller than the valleys, some correction is usually applied. E.g., Priest [9] used for σ_S a value of Rq reduced by a factor from 0.32 to 0.50; Tomanik [10] assumed that σ_S = Rpk. Those corrections were chosen to adjust specific models to experimental wear data and no general rule was given in how to estimate the input parameters.

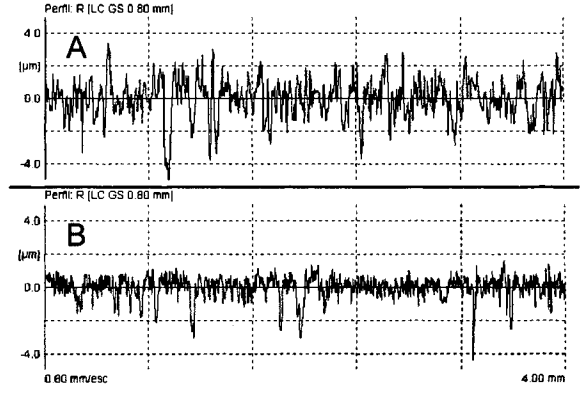


Fig. 3. Roughness profile of gasoline ICE cylinders, as produced.

A- Std Finish. Rq: 1.2 μm; Rpk: 1.0 μm.

B- Plateau honing. Rq: 0.7 μm; Rpk: 0.3 μm.

2. SURFACE SPECTRAL MOMENTS

Nayak [11] proposed the parameters m_0, m_2 and m_4, spectral moments of the profile Power Spectral Density, PSD, function, to determine a large number of useful statistics for isotropic, gaussian surfaces. The PSD function is the Fourier transform of the profile autocorrelation function. According to McCool [5], these spectral moments may be calculated, for isotropic surfaces, by the expressions:

$$m_0 = E\{(z(x))^2\} \tag{3}$$

$$m_2 = E\{\left(\frac{dz}{dx}\right)^2\} \tag{4}$$

$$m_4 = E\{\left(\frac{d^2z}{dx^2}\right)^2\} \tag{5}$$

where z(x) is a profile in an arbitrary direction x and $E\{\ \}$ denotes a statistical estimator. m_0 is simply the mean square height, i.e. $(Rq)^2$. m_2 is equivalent to $(\Delta q)^2$. Both Rq and Δq are common roughness parameters. But m_4 is approximately the mean square curvature, not yet available from commercial roughness measurement equipment. With the knowledge of m_0, m_2 and m_4, it would be possible to calculate the asperity parameters according to random field theory, [5]:

$$\sigma_S = \sqrt{\left(1 - \frac{0.8968}{\alpha}\right) \cdot m_0} \tag{6}$$

$$\beta = \frac{3 \cdot \sqrt{\pi}}{8 \cdot \sqrt{m4}} \tag{7}$$

$$\eta = \frac{m_4 / m_2}{6 \cdot \pi \cdot \sqrt{3}} \tag{8}$$

where:

$$\alpha = \frac{m_0 \cdot m_4}{m_2^{\ 2}} \tag{9}$$

However, calculation of m_2 and m_4 involves practical difficulties, as sample spacing and signal processing. For this work, the authors made trials to determine the second and fourth spectral moments, calculating the variation (dz/dx) and (d^2z/dx^2) of the measured points by different methods, as described below. The profile file measured by the MAHR Concept Perthometer PGK equipment was used. The file contains 8064 points, 0.695 μm spaced. Because, in this equipment, the data contained in the file is supplied as measured, an additional filtering software has to be developed and it is described in Annex I. This user filtering process can be omitted if the equipment supplies the profile already filtered.

Three different methods for calculating the first and second derivatives were made:

Simple derivation (SD):

$$(dz/dx)_i = (z_{i+1} - z_i)/h \tag{10}$$

$$(d^2z/dx^2)_i = \frac{z_{i+1} - 2z_i + z_{i-1}}{h^2} \tag{11}$$

Three points derivation (3PD):

$$\left(\frac{dz}{dx}\right)_i = \frac{z_{i+1} - z_{i-2}}{3h} \tag{12}$$

$$\left(\frac{d^2z}{dx^2}\right)_i = \frac{z_{i+2} - 2z_{i-1} + z_{i-4}}{9h^2} \tag{13}$$

Finite central differences (FCD):

$$\left(\frac{dz}{dx}\right)_i = \frac{-z_{i+2} + 8z_{i+1} - 8z_{i-1} + z_{i-2}}{12h} \tag{14}$$

$$\left(\frac{d^2z}{dx^2}\right)_i = \frac{-z_{i+2} + 16z_{i+1} - 30z_i + 16z_{i-1} - z_{i-2}}{12h^2} \tag{15}$$

Where: $h = x_{i+1} - x_i = 0.695$ μm

In order to test these methods, they were applied in measurements of standard roughness profiles Mahr Perthen. These standards are used to calibrate

the equipment. They may be considered gaussian, although not isotropic. Table 2 reproduces some roughness parameters, measured by the equipment. All the measurements cited in this paper were obtained using a MAHR Perthomether equipment, with a 2 μm, 90° stylus. 0.8 mm Cut-off length.

Table 2
Standard profiles, measured values

Profile	Skewness	Rq	Δq
PGN01	-0.09	0.59	0.040
PGN03	0.00	1.12	0.061
PGN10	0.00	2.67	0.085

As expected, the calculated values of m_2 agreed for the different methods as well as showed agreement (±5%) with the measured Δq. However the calculated values of m_4 showed great variation between one method and another. See table 3.

Table 3
Calculated spectral moments, standard profiles.

Profile	m_2 (10^{-3})			m_4 (10^{-4})		
	SD	3PD	FCD	SD	3PD	FCD
PGN01	1.6	1.5	1.6	4.1	0.7	6.2
PGN03	3.7	3.6	3.7	5.3	1.0	7.9
PGN10	7.2	7.1	7.2	5.8	1.1	8.6

It was considered that it would be unreliable to estimate the summits radius and density with the spectral moments. A different, more direct, approach was developed to calculate the asperity parameters. It will be described in the next sections, specially in section 4.

3. STATISTICAL DISTRIBUTION OF THE HEIGHT IN ACTUAL SURFACES.

The G-W model assumes that the summit heights have a gaussian distribution. Most of the works on piston ring modelling accept this assumption, in part because there are evidences that the summit heights have gaussian distribution, even if the surface, as a whole, does not. See fig. 4. This was verified, by the present authors, for the surface finish of various ICE cylinders. E.g. fig. 5 shows the Normal plot of the cumulative height distribution of the plateau honing cylinder, fig 3B. As typical for a two-stage plateau honing, the surface is not gaussian, the valleys are much deeper than the peaks. Even with peaks, z>0, having an almost gaussian distribution, the usual simplification that σ_S = Rq greatly overestimates the micro-contact occurrence and pressures. E.g., the profile of fig. 3B has Rq = 0.7 μm and maximum height 1.67 μm. At the height of the highest peak, the usual, gaussian, estimation would be that 0.8% of the profile is in contact*. Actually, this contact amount is only reached at a much lower distance, less than 1 μm. At height of 1.0 μm, the usual prediction would be that 7.6% of the profile will be in contact**, actually only 1.3% is. See fig. 5. This difference would impact in the calculated tribological parameters, as asperity pressures, minimum oil film thickness, etc.

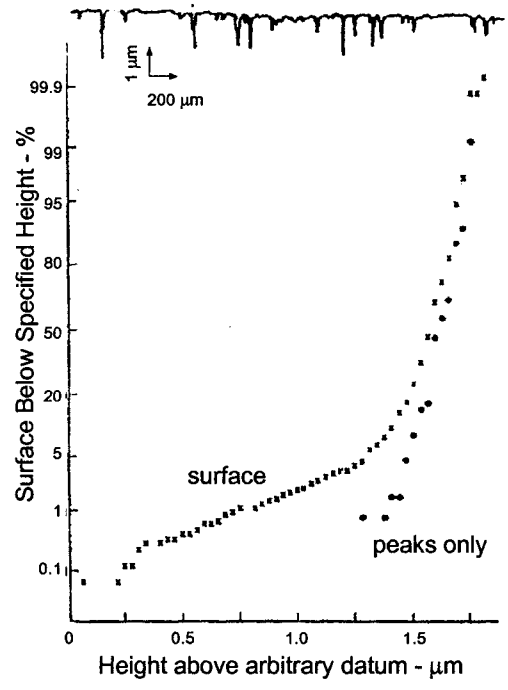


Fig. 4. Cumulative height distribution plotted on normal probability paper. [2].

* 0.00820 (0.8%) is the tabulated area under the standardized normal curve for z = 2.39, i.e. 1.67μm divided by Rq (0.7μm).

** 0.076 (7.6%) is the tabulated area under the standardized normal curve for z = 1.43, i.e. 1.0μm divided by Rq (0.7μm).

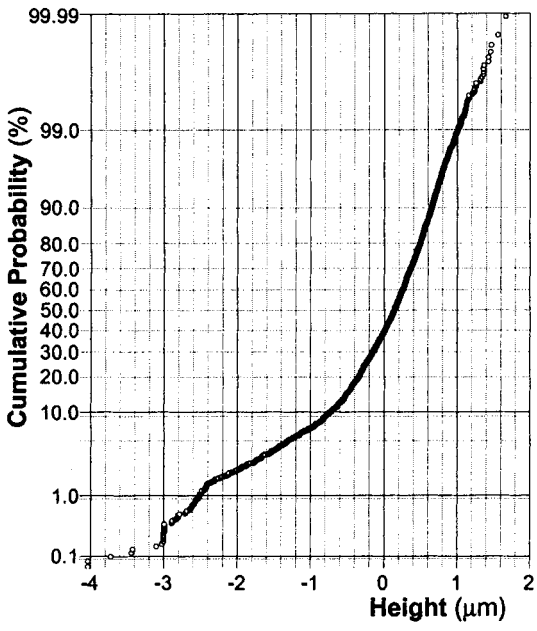


Fig. 5. Cumulative height distribution. Cylinder bore, plateau honing finish. Profile fig.3B.

To comply with the fact that actual surfaces are usually non-Gaussian, the proposed procedure is described in the next section:

4. METHOD FOR CALCULATION OF THE ASPERITIES PARAMETERS

First step is to verify if the peaks follow a gaussian distribution. This was verified to be true for most of the analyzed, new and worn, cylinder bores. Some form and waviness filtering was necessary in the profile file supplied by the equipment used for this work. See Annex I.

4.1. Calculation of η

The summit density is calculated by computing the number of occurrences of positive local maxima. Only the portion of the filtered profile that is above or in the mean line is considered, i.e. the number of occurrences in the profile that:

$$z_i > (z_{i-1}, z_{i+1}) \quad and \quad z_i > 0$$

The equipment stylus, hence the measured profile, does not in general pass exactly at the summit of a peak. The measured profile does not exactly correspond to the surface summits. Nayak [11] proposes that, for random surfaces, the density of the summits, η, is equal to 1.2 times the square of the density of peaks, D_P. $\eta = 1.2\ D_P^2$. Because the studied surfaces are not random, we chose to consider that $\eta = D_P^2$. It was assumed that the surface is isotropic and the density was calculated by:

$$\eta = \left(\frac{\text{number of local maxima}}{\text{profile length}} \right)^2 \qquad (16)$$

There is, also, risk of finding some false local maxima. See discussion in [14], as well as not counting possible local maxima that exists between the acquired measurement points. These errors appear to be inherent of one dimensional, discrete, measurement. They were neglected in this work.

4.2. Calculation of σ_S

Following the definition given in 4.1, σ_S is assumed to be equal to the standard deviation of the profile peak heights. The mean height of the peaks is also calculated. $\overline{Z}_s$ is assumed to be equal to the average height of the peaks.

4.3. Calculation of β

For each positive local maximum, it is calculated the radius of the circle that passes through the maximum, z_i, and its nearest neighbours: (z_{i-1}, z_{i+1}). The mean radius of the summits, β, is considered to be the average of these peak radii.

It was also calculated the radius that passes through the maximum and its second nearest neighbours: (z_{i-2}, z_{i+2}). In average the mean radius increased three times, while the density decreased 46% for the standard bores and 38% for the plateau ones. These results are shown as β_2, η_2 in table 4.

4.4. Application of the method on actual cylinder bores

The proposed methodology was applied for 2 engine blocks of large production, European SI 4 cylinder engine. One block had standard bore finish. The other, plateau honing. The blocks were measured, as produced (new) and after 15 hour dynamometer test. Topography parameters were calculated, according to the proposed method, for each measurement. See table 4, where SN: Std new, SU: Std after 15h test, PN: Plateau New, PU: Plateau after 15 h test.

As expected, σ_S decreased after test, and it was lower on the plateau honed cylinders. This may be understood as a preliminary indication of the lower tribological severity which is achieved with the plateau and with worn surfaces. In average, the product $\sigma_S.\beta.\eta$ also decreased after test and is lower on the plateau ones. In comparison with the literature, see table 1, the following comments can be done about the calculated values:

- some cases, including all PU, had the $\sigma_S.\beta.\eta$ lower than the usual limit 0.03. This may be due to the fact that ICE cylinders are produced to have smaller peaks, and may be expected for worn surfaces, like the ones after 15 h engine test. Or it can be caused by the calculated low mean radius, see discussion ahead.

- the calculated summit radius, β, is much lower than the values previously published. β increased if the second, not the first, nearest neighbours were used. But still to values lower than literature. Preliminary, it was decided to use the value calculated with the nearest neighbours, but deeper studies are necessary. A visual examination of the profile local maxima, where the radius were calculated, supports the choice of β_1. Fig. 6 shows a measured profile detail, with the same magnification in both axis. The calculated radius for this peak is 5.0 μm.

Table 4
ICE cylinders – Measured and Calculated Values

		Measured								Calculated					
	#	Rq	Rz	RSk	Rpk	Rk	Rvk	MR1	MR2	$\bar{Z}_s$	σ_S	β_1	η_1	β_2	η_2
		μm	μm		μm	μm	μm	%	%	μm	μm	μm	10^9 1/m²	μm	10^9 1/m²
SN	1	1.2	6.4	-0.7	1.0	2.4	1.8	9.8	85.0	0.90	0.65	7.5	9.7	21.2	4.6
	2	1.4	8.9	-1.6	1.0	2.8	2.9	6.0	85.0	1.05	0.66	6.7	10	15.4	5.2
	3	1.3	8.1	0.1	1.4	2.8	1.7	11.4	88.0	1.11	0.90	5.5	7.2	19.4	3.9
	4	1.1	6.8	-0.2	1.0	2.8	1.4	8.9	89.8	1.00	0.66	4.1	7.7	15.2	4.1
SU	1	0.7	4.0	-3.6	0.1	1.0	1.3	1.9	75.3	0.43	0.22	16.1	8.5	28.6	3.7
	2	0.6	3.3	-2.5	0.1	0.9	1.2	3.5	81.3	0.34	0.17	20.4	9.5	22.8	5.0
	3	0.6	3.9	-2.6	0.1	1.0	1.3	6.3	82.4	0.39	0.22	20.4	8.5	40.8	4.9
	4	0.6	3.2	-2.4	0.1	1.1	1.1	4.9	84.8	0.39	0.19	9.5	7.2	33.5	4.3
PN	1	0.7	3.9	-1.7	0.3	1.1	1.4	7.4	82.4	0.50	0.26	4.6	16.4	20.0	9.8
	2	0.9	6.4	-4.1	0.4	1.0	2.3	8.3	84.8	0.54	0.31	9.3	17.2	27.9	10.0
	3	0.6	3.9	-1.8	0.3	1.2	1.3	7.2	84.0	0.51	0.26	8.5	14.4	22.3	9.1
	4	0.9	4.8	-2.0	0.3	1.3	1.8	7.6	79.3	0.61	0.32	4.1	16.0	18.0	9.8
PU	1	0.5	3.0	-2.2	0.1	0.6	1.2	4.7	75.4	0.30	0.14	15.0	10.3	35.2	5.8
	2	0.7	4.1	-7.0	0.2	0.5	1.0	6.2	75.7	0.28	0.17	19.0	7.1	55.2	4.2
	3	0.6	4.0	-4.0	0.1	0.4	1.1	6.0	64.8	0.31	0.13	8.8	8.2	42.2	5.5
	4	0.6	4.1	-3.3	0.1	0.4	1.2	3.9	69.5	0.32	0.13	6.6	10.2	28.7	6.4

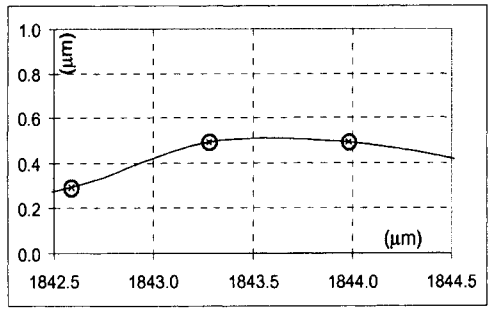


Fig. 6. Profile detail showing a typical peak. Same magnification on both axes.

The calculated σ_S was compared with some common measured roughness parameters. One of them was obviously Rq. See fig.7. The correlation is good, R^2 = 0.88, in a relatively large range, covering all the 16 studied bores. But not for smaller values of Rq. In the spotted region (Rq < 1.0) that considers all cases except the SN, R^2 decreases to 0.64. Much better correlation of σ_S, was found with Rpk + Rk/2. See fig. 8. For all 16 cases, R^2 = 0.98. For the spotted region, R^2 = 0.91. This very good correlation can be explained by the fact that both parameters, σ_S and (Rpk + Rk/2), are basically derived from the portion above the surface mean line.

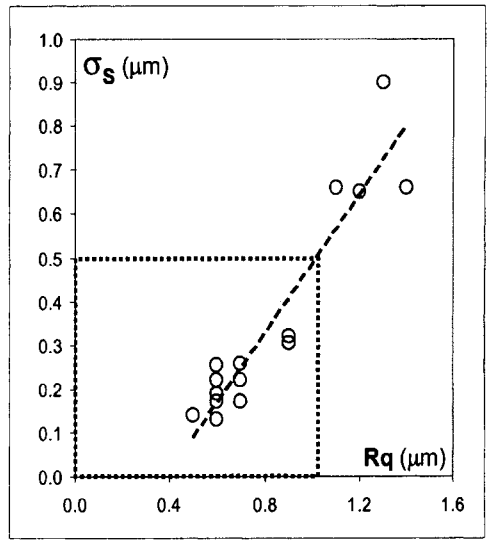


Fig. 7. Correlation between the calculated σ_S and the measured Rq.

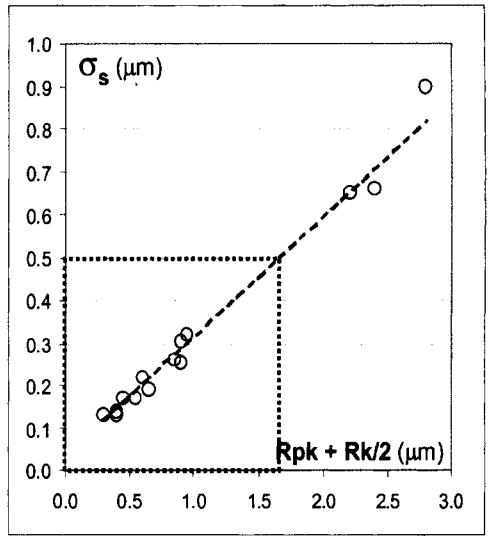


Fig. 8. Correlation between the calculated σ_S and the measured Rpk + Rk/2.

5. CALCULATION OF MICRO-CONTACT PRESSURES WITH THE CALCULATED PARAMETERS.

Some numerical data treatment was necessary to address the usual simplifying hypothesis that the summit mean line coincides with the surface mean line. This hypothesis is present in several current commercial, as well proprietary, computer codes for piston ring simulation. In these codes, it is suggested to use not the calculated σ_S, but a value double of the originally calculated. This is done to accommodate all the summit distribution, even neglecting its mean height, $\overline{Z}_s$. See fig. 9. It should be emphasized that this is only a numerical artifice to be used in computer codes that do not consider the summit mean height.

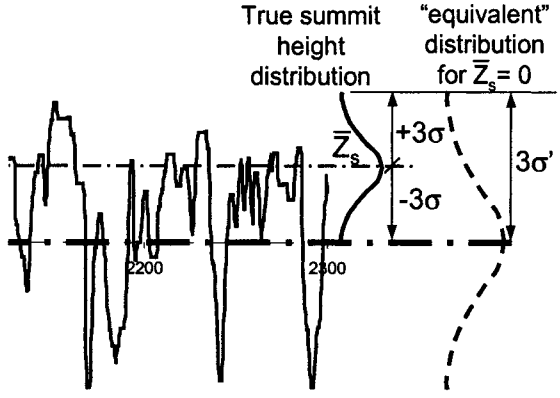


Fig. 9. Schematic equivalent summit height distribution (dotted line) for computer codes that do not consider the Summit mean Height ($\overline{Z}_s$).

The micro-contact pressures, that arise when a lubricated piston ring slides against the cylinder bore, were calculated with the computer program described in [15]. The asperity pressures were calculated according the G-W model. The hydrodynamic pressures are calculated by a simplified Reynolds equation where the calculation of the squeeze effect was turned off . Increase of oil viscosity by pressure was not considered, neither shear thinning effects. The following condition was investigated:

- 1.2 mm wide ring, with 10 μm barrel drop, symmetrical. The ring face barrel profile was approximated by a parabola.
- E' = 100 GPa
- oil SAE 40 at 100 °C (viscosity= 0.013 Pa.s)
- loading pressure = 3.5 MPa (35 bar)
- piston speed = 1 m/s

To study only the influence of the bore roughness, it was assumed that the combined roughness is equal to $2\sigma_S$ (due to the cited correction because $\overline{Z}_s = 0$), the ring was assumed to be smooth. The pressures were calculated with the average values found for each bore condition of table 4. For comparison, the same ring profile and operation condition, but with the σ_S, β, η values given by [3], was also calculated. Table 5 summarises the results showing the calculated values of minimum lube film (Hm) between ring and bore; the film parameter ($t = Hm/\sigma_C$) and the percentage of the load supported by hydrodynamic pressures, P_{HYD}. A fully hydrodynamic regime, t > 3, would have P_{HYD} = 100%.

Table 5
Calculated Hydrodynamic indicators.

	σ_C μm	β μm	η 10^9	Hm μm	t	P_{HYD} %
SN	1.44	5.9	8.3	3.13	2.2	4
SU	0.40	16.6	8.1	0.64	1.6	35
PN	0.58	6.6	15.8	1.08	1.9	18
PU	0.28	12.4	8.7	0.41	1.5	58
[3]	0.185	275	3.0	0.30	1.6	81

It is interesting to notice that, at least for the calculated cases, two parameters that can be seen as indicators of the tribological severity didn't follow the same trend: the percentage of the load supported by the hydrodynamic pressures, P_{HYD}, was lowest for the Standard New (SN) case that at the same time presented the highest parameter film, t. The latter would be an indicative of a more hydrodynamic, less severe regime. This didn't happen because the relatively large lubricant film (3.13 μm) was unable to produce the same load capacity as in the other cases, with lower lubricant film thickness.

So, according to the percentage of the load supported by the Hydrodynamic pressures P_{HYD}, the tribological severity increases from:

$$PU < SU < PN < SN$$

while just the opposite was found for the film parameter, t:

$$SN < PN < SU < PU$$

The lowest lube film thickness was obtained with the data published in [3], mainly due to the very low σ_C (0.185 μm). However this low oil film thickness produced the highest P_{HYD} (81%), but with the second lowest film parameter (t= 1.6).

6. CONCLUSIONS AND DISCUSSION

A numerical method, relatively easy to implement, was developed to calculate the input parameters, σ_S, β, η, of the G-W and G-T micro-contact models. The calculation was made from the usually available measurement profile from commercial equipment.

The method was applied for new and worn ICE cylinder bores. The product $\sigma_S.\beta.\eta$ was found to be between 0.01 to 0.05, lower than the usually accepted range 0.03 to 0.05. The lowest values were obtained on plateau honing finish, after 15 h test. For all cases, the calculated summit mean radius was found to be lower than previous published values.

The method allows a more realistic modelling of actual, non-gaussian surfaces with engineering use, as e.g. ICE cylinders with different bore finish. Some preliminary asperity pressure calculations with the found parameters, suggest that the usual film parameter, t= Hm/σ_C, on the Stribeck curve can lead to wrong conclusions about the tribological severity.

More investigations are necessary to expand the method to consider the non-isotropy of the surfaces. The basic geometrical calculations, that were made in the profile, can be expanded to a 3D surface. More investigations of the correction applied on the calculated σ used in simulation codes, $\sigma' = 2\sigma_S$, is also recommended.

NOMENCLATURE

β	Summit mean radius
η	Summit density
σ_S	Summit height standard deviation
σ'	Corrected σ for use in simulation codes that assume that $\bar{Z}_s = 0$, $\sigma' = 2\sigma_S$
Hm	Minimum lube oil film thickness
m_0	0^{th} Spectral moment
m_2	2^{nd} Spectral moment
m_4	4^{th} Spectral moment
MR1	Material ratio of peaks
MR2	Material ratio of valleys
P_{HYD}	Percentage of the Load that is supported by Hidrodynamic pressures
Δq	Root mean slope of the profile
Rpk	Reduced peak height
Rk	Core roughness depth
Rvk	Reduced valley depth
Rq	Root mean square of profile heights
Rsk	Skewness
t	Oil film parameter, $t = Hm / \sigma_C$
z	Asperities height
$\bar{Z}_s$	Average Asperity Summit height

REFERENCES

1. Greenwood, J.A.; Williamson, J. "Contact of nominally flat surfaces". Proc. R. Soc. London, Series A, v.295, p.300-319, 1966

2. Greenwood, J.A.; Tripp, J.H. "The Contact of Two Nominally Flat Rough Surfaces". Proc Inst Mech Engrs 1970-71 v.185, n.48, p.625-633 , 1970

3. Arcoumanis, C. et al. "Mixed lubrication Modelling of Newtonian and Shear Thinning Liquids in a Piston-Ring Configuration". paper SAE 972924, 1997

4. Johnson, K. L. et al "A simple theory of asperity contact in Elastohydrodynamic lubrication" Wear, 19 (1972) p.91-108

5. McCool, J.I. "Comparison of models for the contact of rough surfaces". Wear, v.107, 1986

6. Hu, Y. et al. "Numerical Simulation of Piston Ring in Mixed Lubrication- A Nonaxisymetrical

Analysis". Trans. of the ASME v.116, July 1994, p.470-478, 1994

7. Gulwadi, S.D. "Analysis of Lubrication, Friction, Blow-by and Oil Consumption in a Piston Ring Pack". ASME-ICE Division, Spring Conference. Fort Collins, Colorado, USA, 1997.

8. Thomas, T.; Rosén, B.-G "Determination of the optimum sampling interval for rough contact mechanics" Tribology International 33 (2000) p.601-610, 2000

9. Priest, M. et al. "Predictive wear modelling of lubricated piston rings in a diesel engine" Wear v.231, p.89-101, 1999

10. Tomanik, E.; Nigro, F. "Piston Ring Pack and Cylinder Wear Modelling" SAE paper 2001-01-0572

11. Nayak, P.R.; "Random Process Model of Rough Surfaces" Journal of Lubrication Technology, ASME transactions p.398-407, July 1971

12. Standard DIN 4777 – Phase correct filters for use in electrical stylus instruments.

13. Form Talysurf series – Appendix: Instrument theory, parameters and definitions.

14. Greenwood, J.A. "A unified theory of surface roughness" Proc. R. Soc. London A 393, p.133-157, 1984.

15. Tomanik, E; Ferrarese, A. "Use of a micro-contact model to optimize SI engine's 3-piece oil ring profiles regarding wear and lubrication" – ASME ICE-Vol. 39, p.427-435, 2002.

ANNEX I- PROFILE FILTERING

In order to calculate the surface parameters, it was necessary first to filter the profile file available by the equipment. This was done by an user made program, using MatLab® routines. The purpose was to replicate the filter usually applied by the equipment for roughness measurement. This program would not be necessary in an equipment that supplies a file, with the data already filtered. The filter process has two steps:

Form Removal

Elimination of the nominal shape of the component from the assessment of texture. The nominal shape may be defined by using various mathematical functions such as circle, conical and aspheric functions. The form removal of the signal was preliminary implemented by a least-square curve fitting of 2nd order polynomial.

Waviness Filtering

Corresponds to a high-pass filter which attenuates long-wave signals, depending on its characteristic, and transmits higher-frequency or short wave signals. This was implemented according to DIN4777 [12]. The filter characteristic of the waviness profile (i.e. of the mean line), is determined from the weighting function by means of the Fourier transformation. It is expressed by the following equation:

$$\frac{a_1}{a_0} = e^{-\pi\left(\frac{0.4697 \cdot \lambda_c}{\lambda}\right)^2}$$

Where,

a_0 amplitude of the profile before filtering.
a_1 amplitude of the profile in the mean line.
λ_c cut-off wavelength of the profile filter.

Fig. I to III illustrate the filtering process. Fig. I reproduces the profile, as supplied by the equipment. 0.8 mm Cut-off, 5.6 mm measurement length, 8064 points. Fig. II shows the user calculated waviness profile, first and last cut-off were considered in the calculation of the waviness, but not plotted. Fig. III shows the profile after form removal and waviness filtering. First and last cut-off length are discarded, equivalent of using a 2CR PC filter [13]. The result is a filtered text file, that for the studied cases, a file containing 5760 points, 0.695 μm spaced, that was then used for calculating the G-W surface parameters.

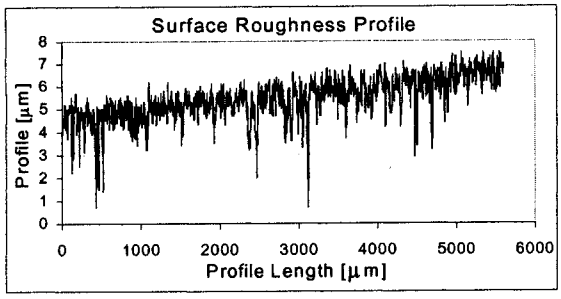


Fig. I. Surface Roughness Profile, as supplied by the equipment, before form removal and filtering.

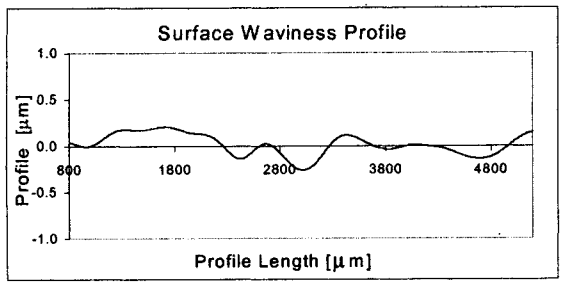


Fig. II. User calculated Surface waviness.

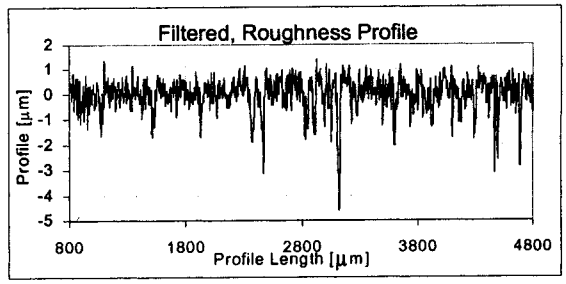


Fig. III. User roughness profile, after form removal and waviness filtering.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

The Study of Self-lubricating Sintered Nickel Matrix Material and its Applications

Liu Zuomin[a], Zhou Yan[b], Zhang Yibing[a], Gao Zhiping[b], and Wang Erhu[a]

[a] P.O.Box 117,Institute of Tribology, Wuhan University of Technology , Wuhan 430070, China,
[b] Beijing Steel-Iron Ltd, China

Abstract: In the paper, a self-lubricating sintering nickel matrix material has been developed and inserted into ball bearing as new solid lubricant at 600℃. For obtaining good lubricating property, some additives are added into the nickel matrix and sintered at elevated temperature. Research shown that CaF_2 has great contribution to lubrication, but amount of CaF_2 must be controlled, and that the lubricating behaviors can be further improved by adding a mixture of CaF_2 and BaF_2, but the proportion of CaF_2 / BaF_2 should be also controlled so that a eutectic compound of Ni and CaF_2 / BaF_2 can form during the sintering. In the paper, the high temperature self-lubricating mechanism of the material has been investigated. Topography and compositions of surface film are analyzed by using SEM and X-ray. Based on good self-lubricating behaviors of the sintering nickel matrix material, an oscillating ball face bearing has be manufactured. Its service test result show that the bearing has a service life over 1.2×10^5 cycle numbers and lower friction coefficient μ=0.26 - 0.12 from room temperature to 600℃at different loads.
Key words: Dry bearing, Sintering self-lubricating materials, Friction and lubrication at elevated temperature.

1. INTRODUCTION

In metallurgical industry, operating conditions of bearing generally is at high temperature and heavy load. A typical example is oscillating bearing in burning valves. In the past, because of that self-lubricating material of the bearing contained some organic lubricants, therefore, its operating temperature is about 450℃. Presently, the operating temperature of the bearing is over 600℃ due to the application of natural gas burning system for controlling environment pollution. For satisfying the requirement of extreme condition, it is necessary to develop a new self-lubricating material so that it has good lubricating property at 600℃ or over.

Sintering, as a powder metallurgical processing, has successfully sintered some solid lubricants into the metal matrix so that friction and wear performances of the materials can be improved [1]. However, some additives couldn't bear high sintering temperature, because of that a part of lubricating phase constructions could be destroyed during sintering process. On the other hand, some additives are easily oxygenated within certain temperature so that friction coefficient increases with operating time [2]. Therefore, for obtaining good sintering quality and self-lubricating property, design of composites and amount of the solid lubricants is very important [3]. A lots of researches shown that the most of fluorides have high melting point and good chemical thermal stability at elevated temperature [4], especially, the CaF_2 and BaF_2 present a better lubricating property at 500°C to 950°C[5]. Some research shown that when on certain conditions, some solid lubricants can be sintered with nickel matrix [6]. However, some

prior research shown that optimum of solid lubricants is very important in controlling sintering quality and lubricating property of material.

The purpose of the paper is to develop a new self-lubricating material by PM technology so that it can be inserted into metal substrate as a new solid lubricant for manufacturing an oscillating bearing which operating temperature is up to 600℃or above.

2. MATERIALS AND EXPERIMENT

2.1 Materials

The fluorides ($CaF_2 + BaF_2$) and small amount Cu are added into the nickel matrix and sintered at 950 ~ 1050℃. The sintering nickel matrix material is manufactured to pin specimens, which are 18 mm long and 10mm diameter with plate tips. The matching material is W6Mo5Cr4V2 which is manufactured to disc specimens of 60mm diameter and 5 mm thickness, hardness HRC =62.

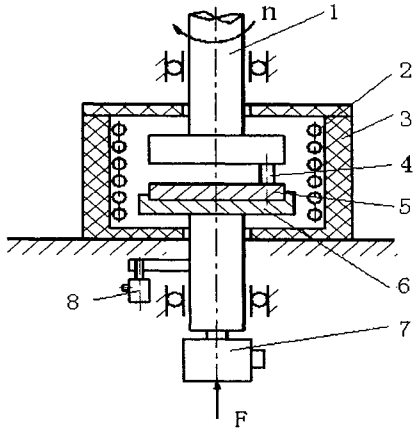


Fig.1 Schematic diagram of pin and disc machine

2.2 Experiment

A test rig (as shown in Fig.1) was developed for purpose of hot wear up to 900°C. The pin (4) and disc (5) specimens are placed at the center of a wire-wound tube furnace (3) which temperature range can be changed from room to 900°C by an instrument. The pin and disc are mounted in holders at the end of resistant hot steel tubes filled with insulating ceramic fiber.

A loading arm carries out the loading between the two specimens by transducer (7). Load transducer (7) and (8) respectively measures the pressure and friction force between the pin and disc. The measuring result is online printed out. A PC controls all test parameters.

The experiment conditions: Sliding velocity V = 0.12m/sec, loading P=150N, temperature is changed from ambient to 750°C.

3. TEST RESULT AND ANALYSIS

3.1. Effect of the CaF_2 on the material.

As high temperature alloy, the sintering nickel matrix material has some lubricity, but strong stick phenomenon also can be observed when it slips on the surface of metal at elevated temperature. For obtaining good self-lubricating properties, CaF_2 and BaF_2 additives are added into the sintering nickel matrix based on CaF_2 and BaF_2 having higher melting point and chemical stability at elevated temperature. However, the fluorides over amount could weaken sintering quality of material. Therefore, it is important to study the amount and proportion of the additives.

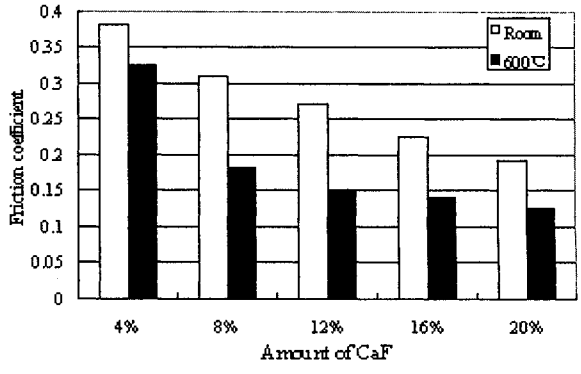


Fig.2 Effect of amount of CaF_2 on lubrication

Figure 2 shows the test results of friction coefficient vs. amount of CaF_2 at room temperature and 600°C. It can be observed that the friction

coefficients decrease from 0.38 to 0.19 at ambient and from 0.33 to 0.12 at 600°C with the increase of CaF_2, and that a rapidly change occurs between at 4% and 8% at 600°C. It can be also observed that the change of friction coefficient at 600°C is very small after 8% CaF_2. This shown that amount of CaF_2 is an important factor in improving self-lubricating behaviors of the sintering nickel matrix material.

On the other hand, amount of CaF_2 also affects the wear rate of the sintering nickel material. Generally, density of the sintering nickel matrix material decrease with increase of CaF_2, and when CaF_2 is beyond 20%, wear rate of the sintering material dramatically increases. Fig.3 is the test results of wear rate vs. amount of CaF_2 at room temperature and 600°C. It is obvious the higher wear rate occurs at 20% CaF_2. In the experiment, it is also observed that serious stick occurs at 4% CaF_2. The stick phenomenon gradually disappeared and surface film presents on the worn surface after CaF_2 beyond 8%.

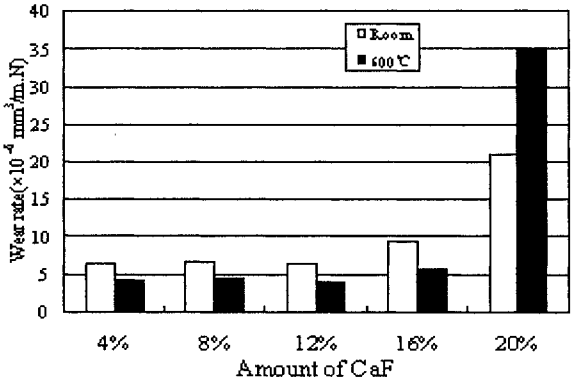


Fig. 3 Effect of amount of CaF_2 on wear rate

The result above shown that for making the sintering nickel matrix material having good tribological properties, the amount of CaF_2 should be controlled within certain range.

3.2 Effect of CaF_2 / BaF_2 on material

A early research shown that composite matter of CaF_2 / BaF_2 has good high temperature self-lubricating behaviors while they were used as coating [4]. However, when CaF_2 and BaF_2 are simultaneously added into nickel matrix and sintered at high temperature, the situation has little change. First of all, the fluorides could be dissolved at high temperature so that a part of lubricating properties could be lost. Secondly, it is difficult how to optimize the proportion of CaF_2 / BaF_2 so that good self-lubricating film can be formed on frictional surface.

In the experiment, the sintering temperature is considered as an important factor for controlling the loss of the solid lubricants. The range of temperature is controlled within 950-1050℃ (details will be discussed in another paper).

For optimizing of the proportion of CaF_2 and BaF_2, a ratio i ($i = BaF_2 / CaF_2$) of CaF_2 and BaF_2 is considered. When i =0, BaF_2 = 0. Only CaF_2 is added into the nickel matrix. When $i \neq 0$, BaF_2 and CaF_2 are simultaneously added into the nickel matrix.

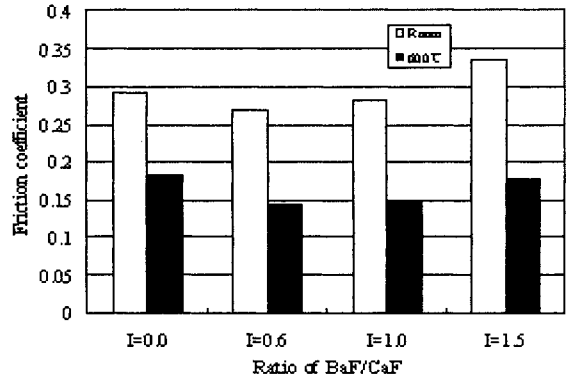


Fig.4 Effect of BaF_2 / CaF_2 on lubrication

Figure 4 is test results for i =0, 0.6,1.0 and 1.5. The results shown that at ambient, of all ratios, the friction coefficient has the highest value at i =1.5 and the lowest value at i =0.6, and that at 600°C, of all ratios, the friction coefficient has the highest value at i =0 and the lowest value at i =0.6. Furthermore, it is also observed that at

both ambient and 600°C, the friction coefficient increases with i from 0.6 to 1.5, and at i =0.6 has the lowest friction coefficient.

Fig.5 is test results of wear rate of the sintering material vs. ration of CaF_2/BaF_2 at 600°C. Comparing the Fir.5 with Fig.4, a corresponding relation between the wear rate and friction coefficient at 600°C can be observed. It is evident that mixture of CaF_2/BaF_2 can effectively improve tribological properties of the sintering nickel matrix material if a reasonable ratio of CaF_2 / BaF_2 could be designed.

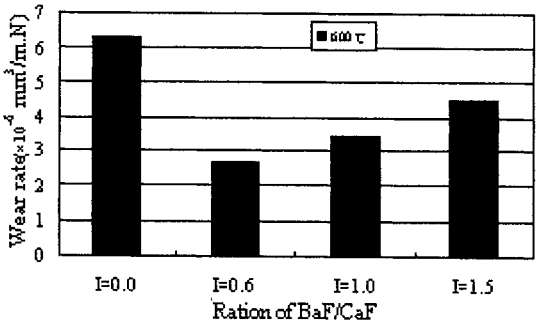


Fig.5 Effect of BaF_2/CaF_2 on wear rate

3.3. Self-lubricating mechanisms of the material

Fig.6 is test curves of friction coefficient vs. time for i =0 and i =0.6 at different operating temperature. Two main characteristics are presented in the figure 6. One of the characteristics is that friction coefficient containing mixture of CaF_2/BaF_2 is lower than that of only containing CaF_2 and the characteristic become very obvious at elevated temperature. Another characteristic is that the friction coefficients decrease with the increase of temperature and have a change point at 200℃.

The change of friction coefficient is mainly affected by two factors. One of the factors is phase construction of fluorides. The fluoride of metal is product of ionic bound of fluoride and metal. Therefore, the fluorides of metal have some special phase construction. For example, fluorides calcium (CaF_2), it is a non-layer micro-structure, but its phase construction has a slip plane (Ca atomic compact plane, for example "abc" face shown in figure 6), and atomic force on the slip plane is lower than other planes. Whiling friction, the slip plane is easily sheared off so that CaF_2 presents better lubricating behaviors, especially, when high temperature weakens the atomic force on the slip plane.

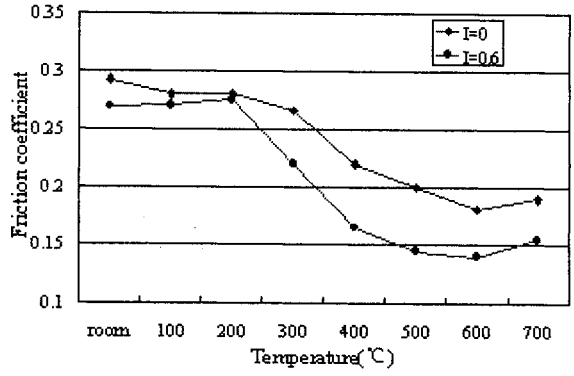


Fig.6 Effect of temperature on lubrication

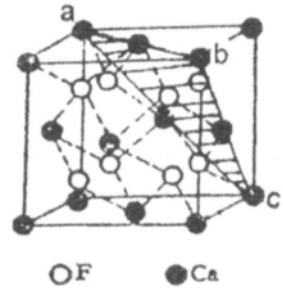


Fig.7 Phase construction of CaF_2

Another factors is eutectic compound of Ni and CaF_2 / BaF_2. Subject to higher melting point and good chemical satiability at elevated temperature, a eutectic compound of Ni and CaF_2 / BaF_2 could be formed during sintering process. The eutectic compound has a softening characteristic at certain temperature so that soften or molten film be easily produced and covered on the frictional surface. Fig 8 is topography of worn scar of the sintered nickel matrix material for i =0.6 at 600°C. It is obvious that a thin layer film present on worn surface.

Figure 9 is a measuring result of x-ray on the film

zone. It indicated that on the film surface, only Ca, Ba , Ni and Cu elements ,which come from the sintering nickel matrix, presents on the film zone without other elements of matching pair elements such as Fe.

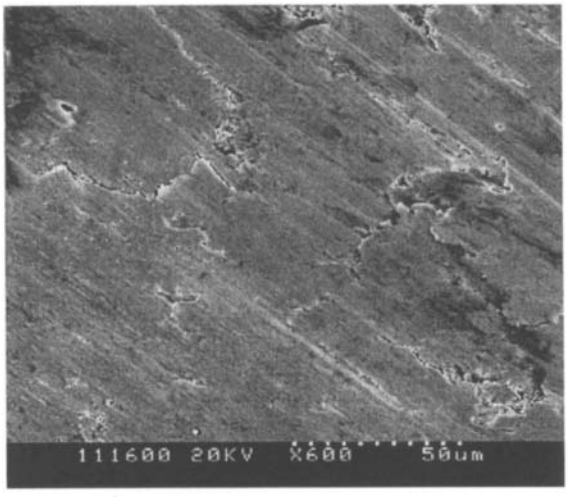


Fig.8 Topography of worn scar

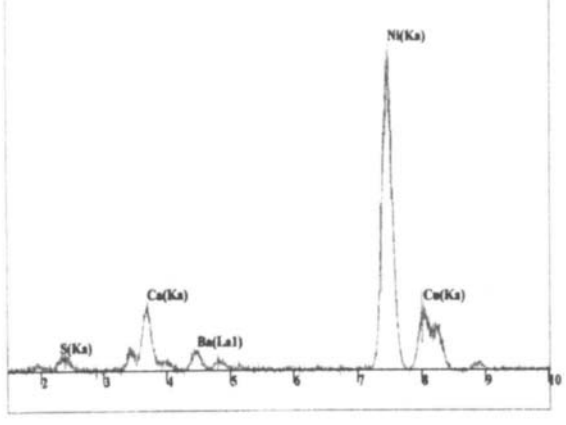


Fig. 9 Measuring result of x-ray of the film zone.

It is evident that as new solid lubricant, the self-lubricating property of the sintering nickel matrix material mainly depends on surface film containing eutectic compound of Ni and CaF_2 / BaF_2.

4. APPLICATION OF THE MATERIAL

Subject to good self-lubricating property of the sintering material at elevated temperature, it was used in an oscillating ball face bearing.

4.1 Design of bearing.

The bearing consists of two parts (as shown in Fig 10): metal spheroid (diameter 120mm) and spherical seat. The sintering nickel matrix material is inserted into the spherical seat substrate with given regular arrangement. For creating enough lubricating film on the frictional surface, the inserting area of sintering nickel matrix material is about 25-32% of surface area of the bearing seat.

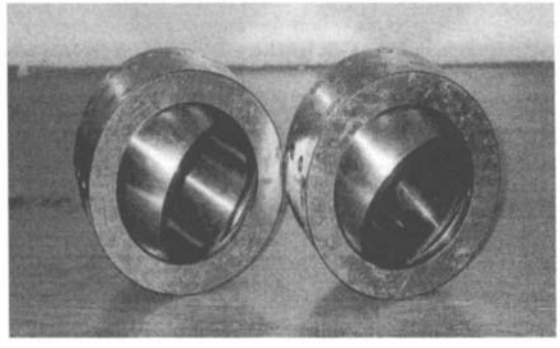


Fig.10 New self-lubricating ball bearing

4.2 Lubricating behaviors of bearing.

The service life test of bearing is carried out in a simulating tester of oscillating bearing (as shown in Fig.11). The bearing is installed in a hot-furnace of

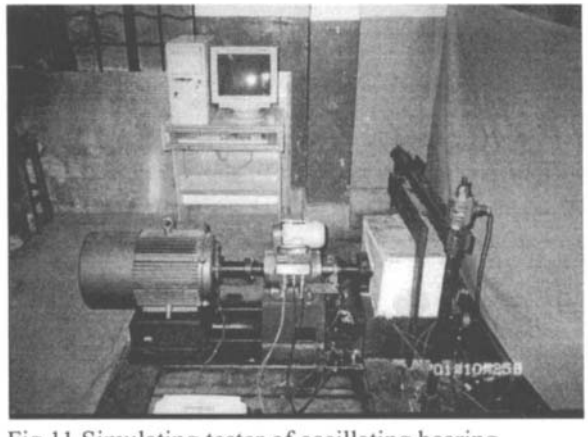


Fig.11 Simulating tester of oscillating bearing

the tester. The furnace is sealed and insulated with resistant hot material. Operating parameters (such as temperature, velocity and load etc) of the tester are controlled by PC. Oscillating frequency of the

test is 30-cycle/each minute. Friction coefficient of the bearing is online measured and printed out. Test results are respectively plotted in figure 12 and 14.

Figure 12 is curve of average friction coefficient vs. oscillating cycle numbers. The test is carried out at the axial load alone (zone A in Fig.12) and at both axial and radial synthetic loads (zone B in Fig.12). At the zone A, The change range of the friction coefficient is from 0.08~0.10 within 60000 cycles at 600℃ at radial load Pr =5300N. The change range of the friction coefficient is from 0.14~0.15 within 65000 cycles at 600℃ at both radial load Pr =5300N and axial load Pa = 900N.

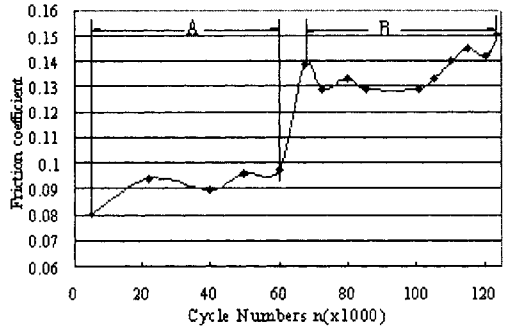


Fig.12 Self-lubricating ball bearing after oscillating 120000-cycles.

The result shown that the synthetic force of radial and axial loads increases the contact stress of the bearing so that surface film bears a higher pressure than only radial load and resulting in higher friction coefficient.

Figure 13 is a ball bearing responding to test conditions of Fig.12 above. It can be observed that only some very small worn traces present on the spheroid surface of the bearing without any sticky phenomenon after operating 120000-cycles. This shown that the sintering nickel matrix can forms a surface film and separates the contact surfaces between the spheroid and spherical seat.

For verifying the loading capacity of the lubricating film, a serial of tests has been carried out at different loads and temperature (as shown in Fig. 14). The test results shown that friction coefficient of the bearing decreases with the increase of test temperature. but increases with load, and that the friction coefficient has a change point at 200℃ and keeps within 0.2 from 400℃-600℃.



Fig.13 Bearing after operating 1.2×10^5 -cycles

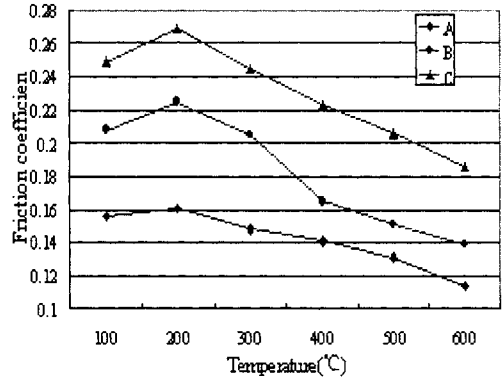


Fig 14 Curves of Friction coefficient vs.temperatures and loads.

(A—Pr= 5300N Pa= 1500N, B—Pr= 5300N Pa= 900N, C—Pr=4550N Pa= 900N. Test temperature: 600℃)

The change of friction coefficient of the bearing shown that lubricating behaviors of the sintering nickel matrix depend on its operating parameters such as temperature and loads, and that an ideal operating temperature of the bearing is over 400℃, and the operating temperature of the oscillating bearing is recommended between 400℃-600℃.

5. CONCLUSION

(1) CaF_2/BaF_2 can be added into nickel matrix and sintered at elevated temperature, but the amount and ratio of the fluorides should be controlled for obtaining both good sintering quality and tribological properties. In the experiment, amount of 8-16% CaF_2 and 0.6 and 1.0 ratio of CaF_2/BaF_2 are recommended.

(2) For the sintering nickel matrix material, some lubricity can be observed after 200℃, but good self-lubricating behavior is over 400℃. Main lubricating mechanism of the material at elevated temperature is its softening film containing eutectic compound of Ni and CaF_2/BaF_2.

(3) When the sintering nickel matrix material is used as solid lubricant and inserted into metal spherical seat substrate, the ball bearing has a good lubricating behavior and higher loading capability.

(4) In the experiment, the oscillating ball bearing has a service life over 1.2×10^5 cycle numbers and lower friction coefficient within the cycle numbers.

REFERENCES

1. Liu Zuomin and T.H.C.Childs, The Influence of TiC, CaF_2 and MnS Additives on Friction and Lubrication of Sintered High Speed Steels at Elevated Temperature , Wear (1996) 193.

2. D-P.Feng etc., Development of Inorganic Solid Lubricants with high Temperature Resistance, Lubrication Engineering (In Chinese) No.6 (1997)

3. S-Q.Niu, J-P. Zhu and J-L.Quyang , Study on Development and Properties of Several New Type High Temperature Self-lubricating Composites, Tribology (In Chinese), No 4 (1995)15.

4. H.E.Sliney, Solid lubricant materials for high temperature ---a Review, Tribology International. No.4 (1982) 15.

5. H.E. Sliney , T.N.Stom and G.P.Allen Fluoride solid lubricants for extreme temperature and corrosive conditions, ASLE Trans. No.8 (1965) 4.

6. R-T.Liu and X-B.Li etc. Nickel–based High-temperature Self- lubricating Material, Hunan Nonferrous metals (In Chinese) No.3 (1998)14.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Behaviours of ultra-thin lubricant films between solid surfaces having atomic scale roughness

T. Okumura[a], J. Sugimura[a] and Y. Yamamoto[a]

[a] Department of Mechanical Engineering, Kyushu University, Fukuoka 812-8581, Japan

Molecular dynamics simulations were conducted in order to study effects of mono-atomic scale steps on sliding surfaces on the dynamic behaviour of lubricant molecules under high pressure and shear. Hydrocarbons, including n-hexane, cyclohexane and n-hexadecane, were assumed as lubricants. Simulations were made such that two layers of lubricant molecules were formed. It was found that steps on the surfaces had dramatic effects on interactions between surface atoms and lubricant molecules. Movements of lubricant molecules, and thus traction between the surfaces, were affected by interactions between the molecules and those between the molecules and the roughness structure, both of which depended on the molecular structure of the lubricant.

1. INTRODUCTION

Recent advances in highly accurate machinery have inevitably imposed very severe conditions on tribological elements. This is due partly to advances in manufacturing processes and also to the need for smaller and more efficient systems for energy conservation. The thickness of lubricating films has been getting smaller and smaller, while there is the necessity of attaining low friction and minimum wear. It is important, therefore, to establish a methodology for designing tribological systems in the so-called boundary/mixed lubrication, or ultra-thin film regime. In this regime, any phenomena are of atomic and molecular scales.

Atomic-level lubrication has been investigated both experimentally and theoretically. Experimental techniques, such as the surface force apparatus (SFA), have been used to study thin molecular films. It has been shown that behaviours of thin films, of a few molecular diameters thick, are generally quite different from their bulk behaviours [1-4]. For example, it is well known that the equilibrium density of the fluid in such a system is not uniform across the film, but shows a periodic profile. While it is difficult at the moment to experimentally observe each of the atoms and molecules involved in actual sliding systems, numerical simulation is capable of providing detailed information of how they behave, though still under simplified conditions. Recently, molecular dynamics (MD) simulations have been used to obtain molecular and atomic-level information relevant to tribological phenomena. In MD simulations, it has been observed that for thin molecular film under shear, layers of molecules between the surfaces are aligned parallel to the surfaces [5-7].

Most experimental or theoretical studies of nano-order lubrication have considered atomically flat solid surfaces although actual solid surfaces usually have roughness. Therefore, an investigation connecting lubrication with solid surface roughness is important. Jabbarzadeh et al. have studied the effect of roughness on lubrication by

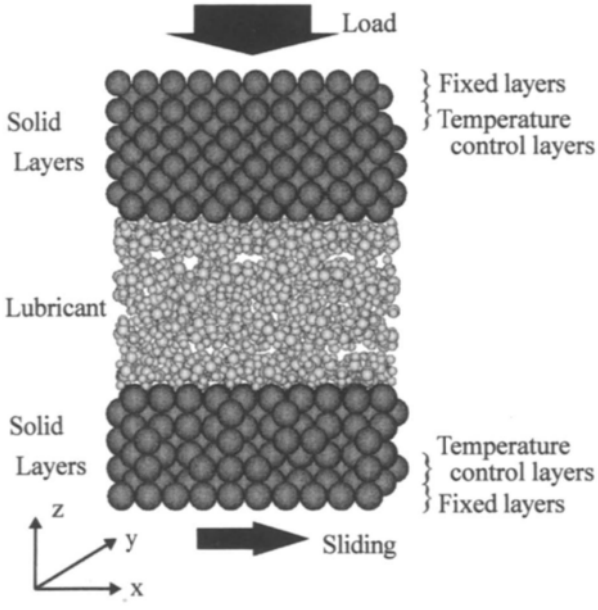


Figure 1 Simulation model.

MD simulation, and have shown roughness effects on properties of the lubrication film [8,9].

The present study focuses on mono-atomic steps on solid surfaces. Behaviours of lubricant molecules subject to high pressure and shear between two solid surfaces sliding very close to each other are examined by using MD simulations. On the basis of the results of the simulations, effects of atomic scale roughness on dynamic behaviour of lubricant molecules and traction between the surfaces in ultra-thin film lubrication are discussed.

2. SIMULATION DETAILS

The details of MD simulation are described here. Figure 1 shows the model system of the simulation. The system is composed of two solid surfaces and lubricant molecules. The lubricant molecules are confined between two surfaces. Periodic boundary conditions are imposed in the x and y directions. Before the two surfaces are sheared, equilibrium structures of all the atoms contained in the system are obtained at a given temperature under a given normal force, or a given pressure, loaded on the upper surface. Motion of the upper surface in the z direction is governed by the net force between the fluid molecules and the surface atoms, and the given load. In sliding simulations, the lower surface is moved at a velocity in the x direction with the upper surface being stationary.

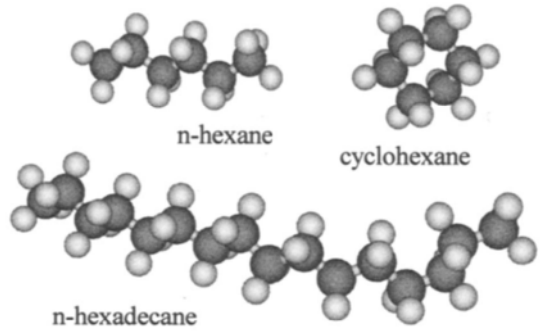


Figure 2 3-D models of lubricants.

2.1 Lubricants

In this study, n-hexane, n-hexadecane and cyclohexane are assumed as lubricants. The 3-D models of lubricants are shown in Figure 2. The intramolecular and intermolecular interactions are calculated using the OPLS-AA, optimized potentials for liquid simulations - all atom, force field potential [10,11]. In this force field, the potential energy function consists of harmonic bond-stretching and angle-bending terms, a Fourier series for torsional energetics, and Coulomb and Lennard-Jones terms for the nonbonded interactions, as defined in eqs. (1) to (4).

$$E_{bond} = \sum_{bond} k_r (r - r_{eq})^2 \qquad (1)$$

$$E_{angle} = \sum_{angle} k_\theta (\theta - \theta_{eq})^2 \qquad (2)$$

$$E_{torsion} = \sum_i \left[\frac{V_1^i}{2}(1 + \cos\varphi_i) + \frac{V_2^i}{2}(1 - \cos 2\varphi_i) + \frac{V_3^i}{2}(1 + \cos 3\varphi_i) \right] \qquad (3)$$

$$E_{non-bond} = \sum_i \sum_{j>i} \left\{ \frac{q_i q_j e^2}{r_{ij}} + 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \right\} \quad (4)$$

where k_r and k_θ are the force constants, r and r_{eq} are bond length and bond length at minimum energy, and θ and θ_{eq} thickness. The parameters are the force constants k_r and k_θ, the bond length r and the bond length at minimum energy r_{eq}, the bending angle θ and the bending angle at minimum energy θ_{eq}, the Fourier coefficients V_1-V_3, the torsion angle φ, the partial atomic charges q and the Lennard-Jones radii and well-depths, σ_{ij} and ε_{ij}. Values of these parameters are given in Table 1. The values of these parameters are from the THINKER molecular modelling software, Force Field Parameter Sets. The combining rules for L-J interactions are used such that $\sigma_{ij}=(\sigma_{ii}\sigma_{jj})^{1/2}$ and $\varepsilon_{ij}=(\varepsilon_{ii}\varepsilon_{jj})^{1/2}$, where i and j denote atom type. The non-bonded interactions are evaluated for intermolecular atom pairs, and also for intramolecular atom pairs separated by three or more bonds. Interactions between intramolecular atom pairs separated by three bonds are reduced by a factor of 2 in order to use the same parameters for both intra- and intermolecular interactions.

Table 1 List of the potential parameters

Atom type	Charge, q
C/CH_2	-0.12
C/CH_3	-0.18
H	0.06

Atom type	σ, Å	ε, kcal mol^{-1}
C	3.500	0.066
H	2.500	0.030
Fe	2.321	12.13

Bond	k_r, kcal mol^{-1} $Å^{-2}$	r_{eq}, Å
C-C	268.0	1.529
C-H	340.0	1.090

Angle	k_θ, kcal mol^{-1} rad^{-2}	θ, deg
C-C-C	58.35	112.70
C-C-H	37.50	110.70
H-C-H	33.00	107.80

Torsion	V_1, kcal mol^{-1}	V_2	V_3
C-C-C-C	1.740	-0.157	0.279
C-C-C-H	0.000	0.000	0.366
H-C-C-H	0.000	0.000	0.318

2.2 Solid surfaces

Solid surfaces are made of iron atoms and form a bcc lattice with the surfaces being the (001) surface. The mono-atomic steps for roughness are set at bcc lattice position by superimposing Fe atoms on the flat surface. Interactions between Fe atoms are expressed by the Johnson potential [12]. Atoms in two layers from the outermost layer are fixed at lattice positions. Two layers inside the fixed layers are for temperature control. Temperature in these layers is kept constant by the velocity scaling method. Thereby it is assumed that the average temperature of the bulk solid was kept constant and heat generated by friction is transferred to the solids.

The Lennard-Jones type potential is used for the interaction between Fe atoms and lubricant molecules. The parameter ε in the Lennard-Jones potential was changed in order to investigate the effect of the interaction between lubricants and the surfaces. There are oxide or any other contaminants present on actual solid surfaces, which may make interaction between surface atoms and lubricant molecules weaker. The calculations are conducted by using the original value ε, and the half value of ε, denoted by $\varepsilon_{0.5}$.

2.3 Calculation conditions

All the calculations are made at a temperature of 300 K, a pressure of 0.5 GPa and a sliding velocity of 100m/s. Solid surfaces have a unit cell size of 28.6Å by 28.6Å.

The RESPA (Reversible Reference System Propagator Algorithms) [13] is used to integrate Newton's equations of motion. Three timesteps are used in the RESPA MD. A timestep of 0.5 fs was used for all the intramolecular

interactions. The L-J interactions and Coulomb potentials are integrated using 1.0 fs and 4.0 fs timestep, respectively. Coulomb potential is calculated by using the Ewald method, and the L-J potential is cut off at 3.0σ.

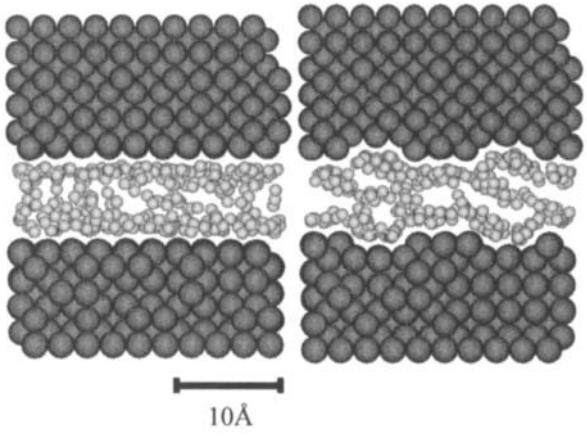


Figure 3 Snapshots of n-hexane molecules under shear.

The number of carbon atoms in the calculated system is 288, which corresponds to 48 molecules of hexane and cyclohexane, and 18 of hexadecane. The mono-atomic steps lie in the y direction, forming transverse roughness. The height and width of steps are 1.43 and 5.72Å, respectively. The simulations are conducted in the cases where upper/lower surface combinations are flat/flat, step/flat and step/step.

3. RESULTS AND DUSCUSSION

Figure 3 shows snapshots of n-hexane molecules under shear between two solid surfaces without/with steps. The surface atoms and carbon atoms of lubricant molecules are shown in the figure. The layered structure parallel to the surface is formed in both conditions. A similar structure is found for other molecular types. This result

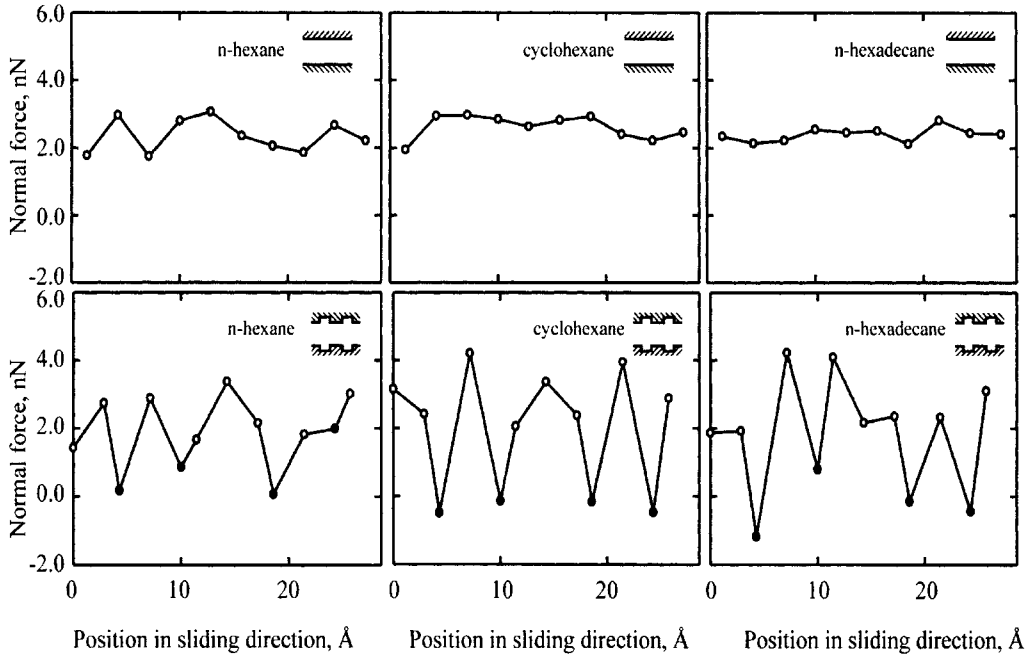


Figure 4 Normal forces acting on the upper surface atoms against the position in sliding direction. (L-J parameter : ε)

agrees with experiments [2, 14] and other simulations [5-7] of confined films. In n-hexane and n-hexadecane simulations, a few molecules that connect the upper and lower molecular layers are observed. It has been reported that such bridging molecules increase shear stress [7].

3.1 Forces acting on the step atoms

Forces acting on the upper surface atoms from lubricant molecules are calculated. Figure 4 shows variations of the force in the z direction with the position in sliding direction for n-hexane, cyclohexane and n-hexadecane molecules. The original values of the L-J parameter ε are used here. Values of forces averaged over 28.6 ps are plotted, which corresponds to the time that the lower surface moves a unit cell width. Positive and negative value of normal force means repulsive and attractive force, respectively. Filled points in the figures indicate the atoms at step position.

While the forces are approximately constant in sliding direction in the case of flat surface, steps give rise to force variation. Comparison of the results of flat and stepped surfaces suggests that only forces acting on the step atoms deviate greatly from the forces on the other positions. Forces for $\varepsilon_{0.5}$ have the similar results.

Interaction between step or kink atoms and lubricant molecules is greater than that between terrace atoms and the molecules. This is because the former atoms have higher activity due to more dangling bonds than the latter, and also because the number of lubricant molecules around the steps is larger than that around flat part of the surfaces.

3.2 Effects on lubricant movement

Density and velocity profiles of carbon atoms of the lubricant molecules across the gap are shown in Figures 5 to 7. The vertical axis is the distance from the lower surface. Solid lines and filled points indicate the distribution across the gap of local density and the velocity of carbon atoms, respectively. The density is the number of carbon atoms per unit length. The velocity profiles

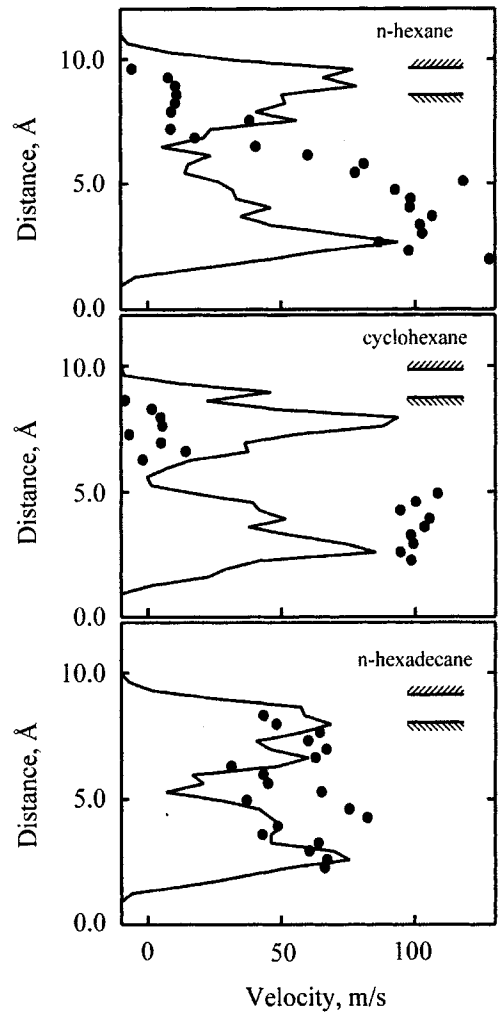


Figure 5 Density profile and the velocity profile. L-J parameter : ε

are obtained by averaging the x-directional velocities of carbon atoms in each of thin layers of 0.3 Å thick.

The profiles for n-hexane, cyclohexane and n-hexadecane between flat surfaces are shown in Figure 5. There are two peaks in each density profile. This means that lubricant films are bilayer. The two peaks of the profile for cyclohexane are more pronounced than those of the other fluids. The velocity profile of cyclohexane exhibits complete separation. This implies that cyclohexane molecules adsorb on the surfaces and have few bridging molecules in this condition. In the case of

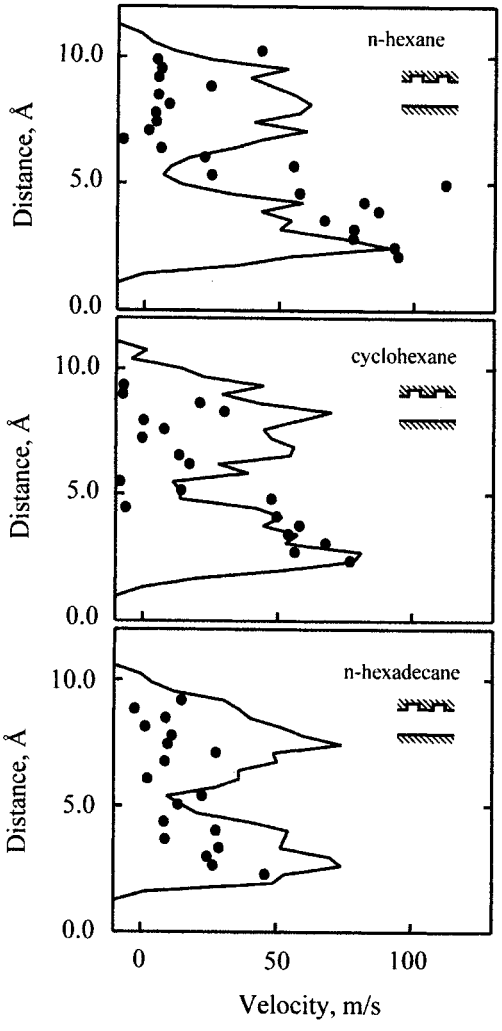


Figure 6 Density profile and the velocity profile.
L-J parameter : ε

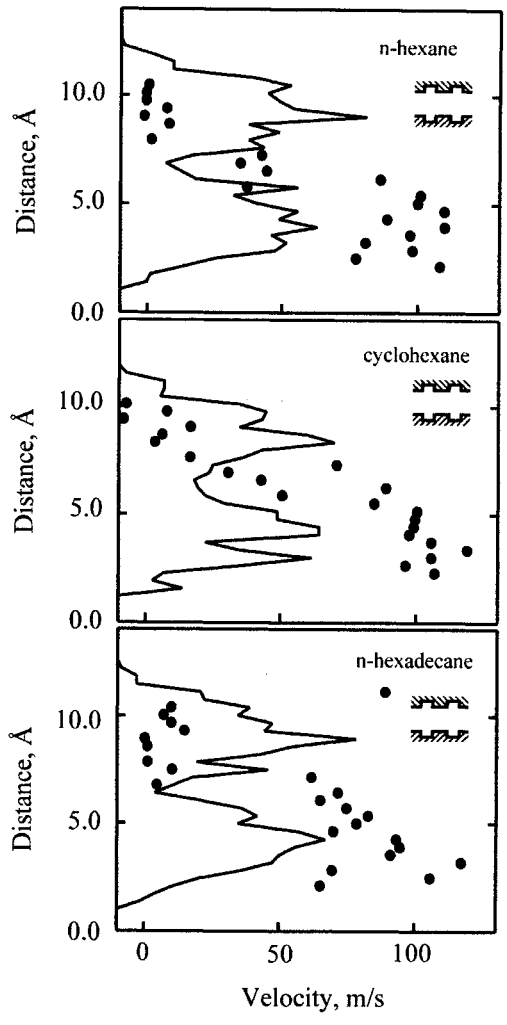


Figure 7 Density profile and the velocity profile.
L-J parameter : ε

n-hexane, the velocity profile is continuous from 0 to 100 m/s. Slip occurs at solid surfaces only in the case of n-hexadecane, where the velocity of lubricant is almost constant across the gap at the average velocity of both surfaces. The length of n-hexadecane molecule is much longer than that of the other molecules, which must have enhanced bridging between the two layers.

Figure 6 shows the profiles in the cases when stationary surface has steps. The velocity of lubricant in the vicinity of the upper surface is about 0 m/s in all types of molecules. No slip is observed for n-hexadecane molecules unlike the case of flat surfaces. These results suggest that the existence of steps on solid surface increase interaction between the lubricants and the surfaces. Steps affect not only the local velocity in the vicinity of the surface but also whole velocity profile. Under such thin film conditions as forming bilayer, one of the surfaces affect the lubricant molecules adsorbed the other surface directly and via lubricant molecules.

Figure 7 shows profiles in the case when both upper and lower surfaces have steps. The profiles demonstrate little or no slip at the surfaces, which is quite different

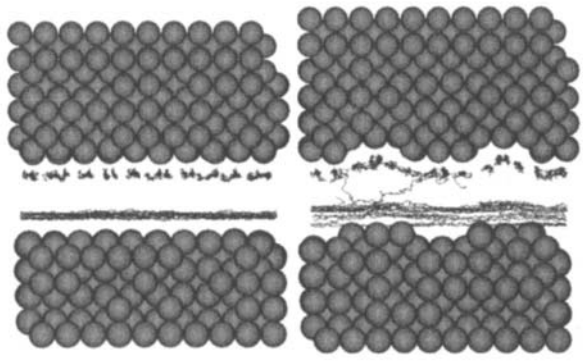


Figure 8 Side views of trajectories of the barycenter of cyclohexane molecules.
(L-J parameter : ε, for 20 ps)

from the results shown in Fig. 6. In particular, cyclohexane exhibits entirely different behaviour from the flat/flat case.

In order to observe behaviours of the cyclohexane molecules, trajectories of the molecules are examined. Figure 8 shows side views of trajectories of the barycenter of cyclohexane molecules over 20 ps in the case of flat/flat and step/step. The movements of molecules form lower and upper layers in the case of step/step, while there is no movement in flat/flat case.

To investigate the effect of interaction between surface atoms and lubricants on the velocity profiles, simulation was made with the half L-J parameter $\varepsilon_{0.5}$. Figure 9 shows the results. Weak interaction between lubricant molecules and surface atoms allows slip at both the upper and lower surfaces except for the case of the lower surface with cyclohexane. The weak interaction also allows the molecules to move more freely, resulting in rather straight density profiles. Figure 10 shows the profiles in the case when both upper and lower surface have steps and the L-J parameter $\varepsilon_{0.5}$ is employed. No slip is seen at the surfaces, and the velocity profiles are continuous in all cases. Motions of admolecules on the stepped surfaces are examined. Figure 11 shows top views of trajectories of n-hexane molecules adsorbed on the upper surface with/without steps for the L-J parameter $\varepsilon_{0.5}$.

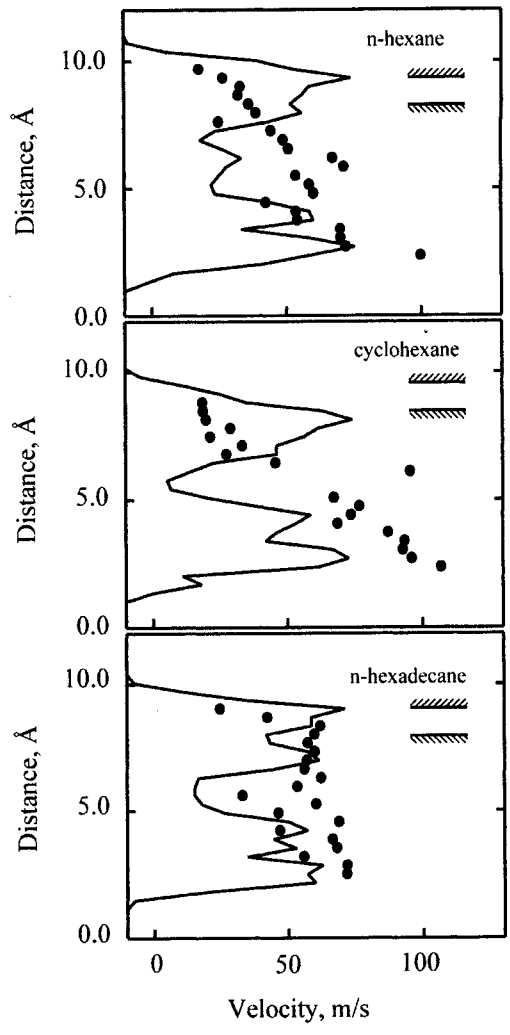


Figure 9 Density profile and the velocity profile.
L-J parameter : $\varepsilon_{0.5}$

Admolecules on the flat surface slip in sliding direction In contrast, most of admolecules stay where they are or the stepped surface. Interaction between lubrican molecules and the surface atoms is thus enhanced by steps on the surfaces.

3.3 Effects on traction coefficient

Traction coefficient is calculated for each set o simulations. Traction coefficient μ is defined here as the ratio F/N where N is the normal load and F is the time average of sum of all the x-direction force acting on the

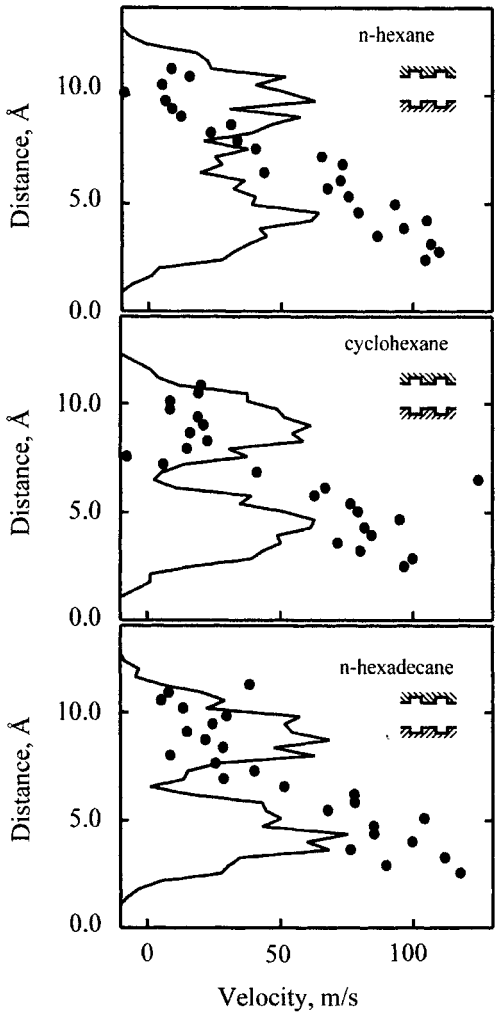


Figure 10 Density profile and the velocity profile. L-J parameter : $\varepsilon_{0.5}$

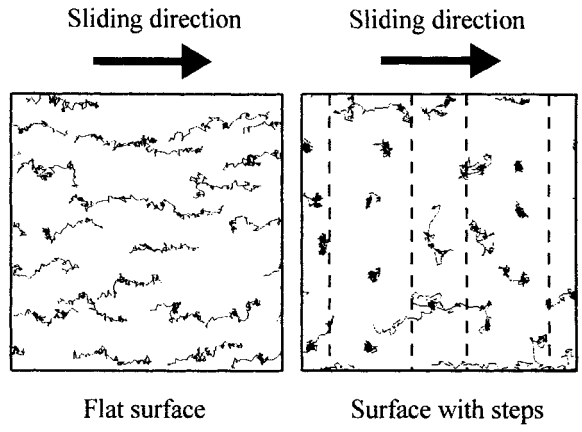


Figure 11 Top views of the trajectories of n-hexane molecules adsorbed on the upper surface with/without steps. Dashed lines indicate step atoms position. (L-J parameter : $\varepsilon_{0.5}$, for 20 ps)

Table 2 Traction coefficient, L-J parameter : ε

lubricant	The number of molecules	flat/flat	step/step
hexane	48	0.23	0.22
cyclohexane	48	0.14	0.22
hexadecane	18	0.42	0.32

Table 3 Traction coefficient, L-J parameter : $\varepsilon_{0.5}$

lubricant	The number of molecules	flat/flat	step/step
hexane	48	0.18	0.23
cyclohexane	48	0.16	0.19
hexadecane	18	0.20	0.24

top fixed atoms of the upper surface. Tables 2 and 3 show the results for the L-J parameters ε and $\varepsilon_{0.5}$, respectively. Results for both the flat/flat and step/step cases are shown in the tables.

The traction coefficients for n-hexane with/without steps are about the same. This is expected from the results shown in Figs. 5 and 7, in which no clear difference in the density and velocity profiles is seen. For cyclohexane, the traction with the stepped surfaces is larger than that with the flat surface. This may be explained by the difference in the density profiles as described 3.2, the bridging molecules in the case of the stepped surfaces cause increase in the traction coefficient. The traction coefficient for n-hexadecane with the stepped surfaces is smaller than that with flat surfaces. As shown in Fig. 7, slip occurs not at the surfaces but between lubricant layers in the case with the stepped surfaces, probably because the interaction between lubricants and surfaces is larger

than the interaction between lubricants to maintain the continuous velocity profile.

The traction coefficients calculated in the present simulations are different from those by Tamura et al.[7]. The traction coefficients of bilayers in their calculation are 0.18 and 0.10 for cyclohexane and n-hexane, respectively. A given pressure in the present calculation is 0.5 GPa, while they used 1.0 GPa, which may have provided higher traction for cyclohexane which may have become solid-like under that high pressure. On the other hand, the number of molecules in the simulation system is larger in this study. It is likely that in the case of n-hexane the relative number of bridging molecules is larger than that in their study, resulting in a higher coefficient.

Table 3 shows the traction coefficients for $\varepsilon_{0.5}$. The values are generally smaller than those shown in Table 2 in all the cases except for n-hexane. The traction coefficients with the surfaces with steps are larger than with the flat surfaces for all the molecules. These results are explained the velocity profiles in Figs. 9 and 10.

Thus the differences in the traction coefficient between flat and stepped surfaces are related with the behaviours of lubricant molecules, including adsorption, slip, aligning and bridging, which are described with the velocity and density profiles. Factors affecting these behaviours are the structures of lubricant molecules, the interactions between lubricants and the interactions between lubricants and the surfaces.

4. CONCLUSIONS

Molecular dynamics simulations have been conducted in order to investigate effects of mono-atomic steps on sliding surfaces on dynamic behaviour of lubricant molecules under high pressure and shear. The steps on the surfaces have a dramatic effect of increasing interactions between surface atoms and lubricant molecules. Movement of lubricant molecules on sliding surfaces is strongly governed by these interactions. In addition, behaviours of lubricant molecules within thin film, including aligning and bridging of molecules, depend on molecular structures of lubricants, the interactions between the molecules and those between the molecules and the surface atoms. Combination of all these effects results in the differences in the traction coefficient.

REFERENCES

1. A. M. Homola, J. N. Israelachvili, P. M. McGuiggan and M. L. Gee, Wear, 136 (1990) 65-83.
2. M. L. Gee, P. M. McGuiggan, J. N. Israelachvili and A. M. Homola, J. Chem. Phys., 93 (1990) 1895-1906.
3. H. Yoshizawa, Y. L. Chen and J. N. Israelachvili, Wear, 168 (1993) 161-166.
4. S. Granick, Science, 253 (1991) 1374-1379.
5. J. Gao, W. D. Luedtke and U. Landman, J. Chem. Phys., 106 (1997) 4309-4318.
6. H. Yamano, *et al.*, Thin Solid Films, 281-282 (1996) 598-601.
7. H. Tamura, *et al.*, Langmuir, 15 (1999) 7816-7821.
8. A. Jabbarzadeh, J. D. Atkinson and R. I. Tanner, J. Phys. Rev. E., 61 (2000) 690-699.
9. A. Jabbarzadeh, J. D. Atkinson and R. I. Tanner, Proc. 27th Leeds-Lyon Symposium on Tribology, (2001) 779-785.
10. W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, J. Am. Chem. Soc., 118 (1996) 11225-11236.
11. R. C. Rizzo and W. L. Jorgensen, J. Am. Chem. Soc., 121 (1999) 4827-4836.
12. R. A. Johnson, Phys. Rev., 134 (1964) A1329-A133611. R. A. Johnson, Phys. Rev., 134 (1964) A1329-A1336.
13. M. Tuckerman, B. Berne and G. Martyna, J. Chem. Phys., 97 (1992) 1990-2001.
14. H. Matsuoka and T. Kato, Trans. ASME J. of Tribology, 118 (1996) 832-838.

This Page Intentionally Left Blank

SESSION VII
BIO-TRIBOLOGY (1) – HIP AND KNEE JOINTS

Chairman : Professor N D Spencer

Paper VII (i) **Fluid Film Lubrication in Artificial Hip Joints**

Paper VII (ii) **Wear of Alumina/Zirconia Nano-composite for Ceramic/Ceramic Joint Prostheses**

Paper VII (iii) **Surface Characteristics and Tribological Behaviour of New Zirconia Ceramics for Cementless Knee Joint Prostheses**

Paper VII (iv) **The Relationship Between Steady-state Wear Rate and Theoretical Film Thickness in Metal-on-Metal Total Replacement Hip Joints**

Paper VII (v) **On the Role of the Lambda Parameter in Simulator Wear of Metal-on-Metal Hip Implants**

Paper VII (vi) **Advances in Simulator Testing of Orthopaedic Joint Prostheses**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Fluid film lubrication in artificial hip joints

ZM Jin[a], JB Medley[b] and D Dowson[c]

[a]Department of Mechanical and Medical Engineering, University of Bradford, UK

[b]Department of Mechanical Engineering, University of Waterloo, Canada

[b]School of Mechanical Engineering, University of Leeds, UK

The importance of fluid-film lubrication mechanisms experienced in hip implants is reviewed. For the different types of hip implants currently used in clinical practice, including conventional/cross-linked polyethylene-on-metal/ceramic, metal-on-metal and ceramic-on-ceramic bearing couples, a detailed review is presented of theoretical fluid-film lubrication analyses along with friction measurements and wear studies, both having a strong emphasis on the fluid-film mechanisms involved. Future studies and improvements of hip implants from a fluid-film lubrication point of view are discussed.

NOMENCLATURE

c	Radial clearance, R_C-R_H
d	Thickness of acetabular cup
E	Modulus of elasticity
h_{cen}	Central film thickness
h_{min}	Minimum film thickness
R_{aH}	Average surface roughness of the femoral head
R_{aC}	Average surface roughness of the acetabular cup
R_H	Radius of curvature of the head
R_C	Radius of curvature of the cup
R	Effective radius defined in equation (2)
t	Time
T	Frictional torque
u	Entrainment velocity, $\omega R_H / 2$
w	Applied load
η	Viscosity of synovial fluid
λ	Lambda ratio defined in equation (1)
ν	Poisson's ratio
σ_H	RMS surface roughness of the femoral head
σ_C	RMS surface roughness of the acetabular cup
ω	Angular velocity of femoral head relative to the cup

1. INTRODUCTION

Natural synovial joints such as hips and knees are remarkable bearings. These bearings are expected to function in the human body for a lifetime whilst transmitting large dynamic loads and accommodating a wide range of movements. However, diseases such as osteoarthritis, rheumatoid arthritis and trauma sometimes require these natural bearings to be replaced by artificial ones. Total joint replacement has been the most successful surgical treatment for hip joint diseases for almost fifty years. Currently, more than 800,000 hip joint replacements are carried out world-wide every year (Dowson 2001). The majority of current artificial hip joint replacements utilise a material combination of ultra high molecular weight polyethylene (UHMWPE) as the acetabular cup articulating against either a metallic (stainless steel, cobalt-based alloy) or ceramic (alumina) femoral head (Figure 1). These man made bearings can sometimes last more than 20 years in the body without failure. However, osteolysis and loosening of the prosthesis has recently been identified as the main factor limiting the *in-vivo* performance of implants often leading to revision after only about 10 years (Amstutz et al, 1991; Harris, 1995). This complication has become more significant because of the increasing use of these devices in younger patients with life expectancies after surgery in excess of 25 years. Osteolysis and loosening are usually caused by an adverse tissue reaction to wear particles of UHMWPE, as reviewed recently by Ingham and

Fisher (2000). Therefore, minimising wear and thus wear particle generation is an important strategy in

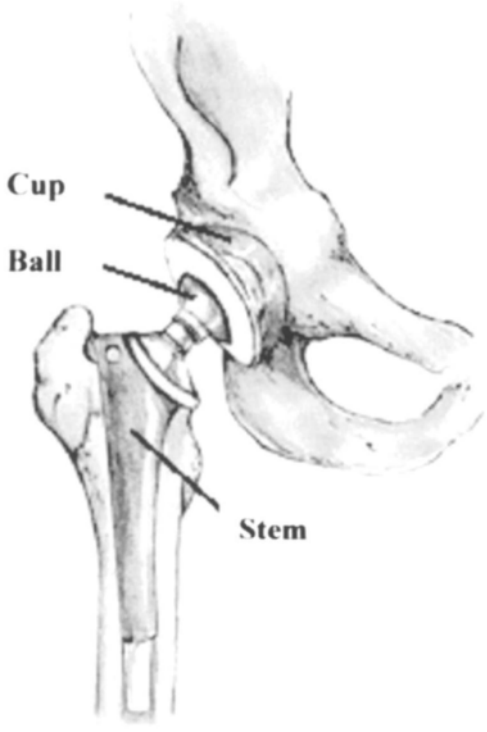


Figure 1. Hip implant consisting of an acetabular cup fixed to the pelvis and a femoral ball fixed to the femur.

prolonging the service life of artificial joint replacements. To reduce wear and wear particle generation, new bearing materials have been introduced such as cross-linked UHMWPE (XPE), both wrought and cast cobalt-based (CoCr) alloys, zirconia ceramic, alumina ceramic composites and polyurethane elastomers (PU). These new materials have been used in various couples and when considered with the older material couples (Table 1) a number of novel bearing combinations are created. Figure 2 shows five typical examples of hip implants.

Some of these couples are still in the research and development stage, including elastomer-on-metal/ceramic (cushion form) bearings (Medley et al 1980; Unsworth et al 1987; Dowson et al 1991) and metal-on-ceramic (Firkins et al 2001a). However, other bearings are increasingly used in clinical practice, such as CoCr-on-CoCr, alumina-on-alumina, and XPE-on-CoCr/alumina.

Tribology plays a major role in the long-term success of artificial hip joints. Wear particles are generated by direct surface contact, causing adhesive and abrasive wear, as well as sub-surface stresses and fatigue. The wear can be reduced by developing or simply taking advantage of favourable surface microstructures. For example, the development of both XPE and alumina composites as hip implant materials can be considered, in their different ways, as attempts to develop a more wear-resistant microstructure.

Since a synovial fluid type of lubricant forms in the joint capsule after total hip arthroplasty (Section 2.1), it may be possible to promote maximum lubricant protection of the implant surfaces. Surface adherent boundary additives can be placed in the synovial fluid to reduce the amount of direct surface contact, but repeated and frequent replacement of such material might be necessary. It is, perhaps, more practical to develop bearings that promote fluid film lubrication (while avoiding fatigue from sub-surface stresses), thus separating the surfaces fully or partially and reducing the proportion of the total load carried by asperity contacts. Therefore, it is important to understand the fluid film lubrication

Table 1
Bearing couples available for hip implants.

FEMORAL HEAD	ACETABULAR CUP					
	UHMWPE	XPE	CoCr	Alumina	PU	Alumina Composite
Stainless Steel	√					
CoCr	√	√	√		√	
Alumina	√	√	√	√	√	
Zirconia	√	√		√	√	√
Alumina Composite				√		√

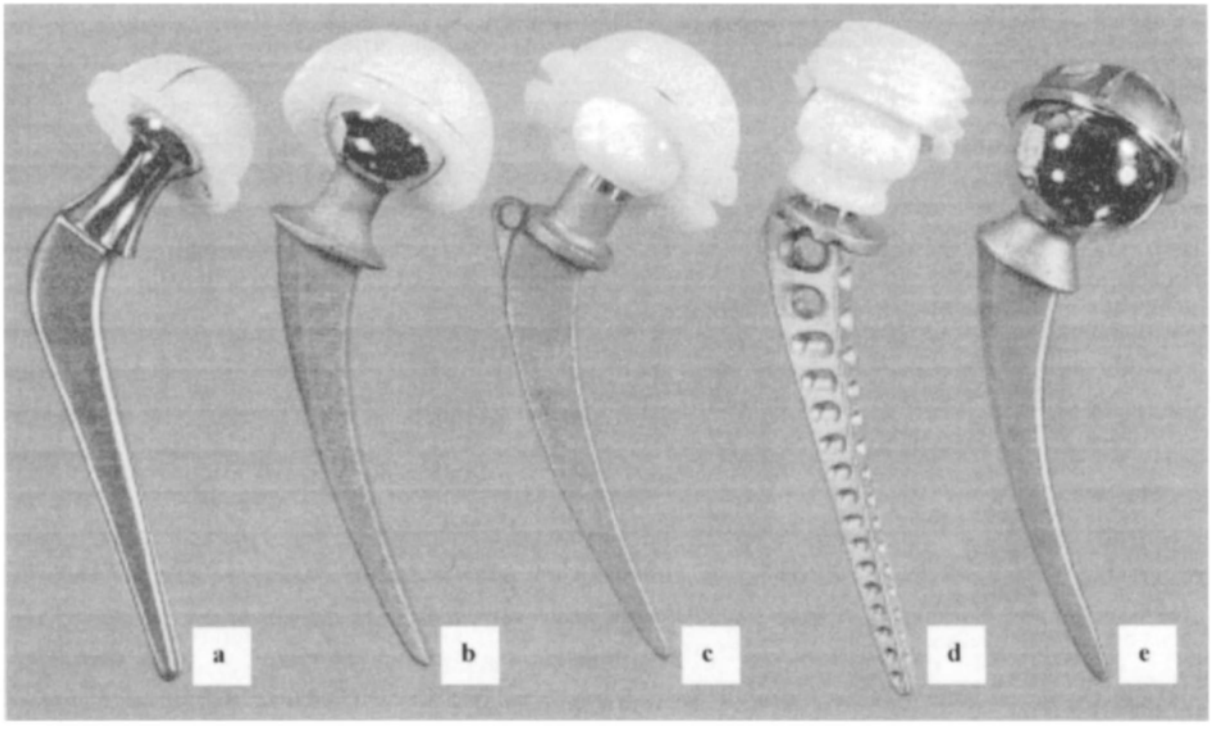


Figure 2 Five examples of hip implants: a) Charnley UHMWPE-on-Metal b) Muller UHMWPE-on-Metal c) UHMWPE-on-Ceramic d) Ceramic-on-Ceramic e) Metal-on-Metal.

mechanism experienced by various forms of hip implants.

Conventional engineering methods of assessing fluid film lubrication regimes can be applied to hip implants. These can be classified into two categories, experimental measurements and theoretical predictions. The experimental methods usually have involved either friction measurements, related to the Stribeck curve, or the detection of separation between the two bearing surfaces using a simple resistivity technique. The theoretical predictions have usually been based on the λ ratio defined as follows (Dowson 2001; Medley et al 2001):

$$\lambda = \frac{h_{\min}}{\left[R_{aH}^{\ 2} + R_{aC}^{\ 2}\right]^{1/2}} \qquad (1a)$$

or

$$\lambda = \frac{h_{cen}}{\left[\sigma_H^{\ 2} + \sigma_C^{\ 2}\right]^{1/2}} \qquad (1b)$$

It should be noted that in most cases the numerical values of both definitions of lambda are almost identical and the interpretations are very similar (Medley et al 2001). The key to the theoretical assessment of the potential for fluid-film lubrication is the accurate prediction of a representative film thickness for the bearing and accurate measurements of surface roughness.

A major objective of the present paper was to review fluid film lubrication studies of hip implants; in particular to focus on an assessment of;

- the efficacy of the fluid film through theoretical analysis (Section 2)
- experimental measurements of friction and film thickness (Section 3)
- the relationship of fluid film lubrication to implant wear in a hip simulator (Section 4).

Other important topics, such as the influence of bearing material microstructure on wear (Dowson 2001) and biological reactions to wear particles (Ingham and Fisher 2000) were beyond the scope of the present review.

2. THEORETICAL ANALYSIS

2.1 Fluid film lubrication model

A simple ball-in-socket configuration, representing the femoral head and the acetabular cup, is generally adopted for the analysis of fluid film lubrication in hip implants (Figure 3a). The materials used to construct the femoral head are metals (stainless steel and cobalt chromium alloy) and ceramics (alumina and zirconia). The acetabular cup is usually made from UHMWPE or more recently XPE, metal (CoCr) and ceramic (alumina). Since the surface deformation influences the film thickness and the mode of lubrication can readily be shown to be isoviscous-elastic, the elastic modulus (E) and Poisson's ratio (ν) of the bearing surfaces plus the underlying bone cement and/or bone (Table 2) are input to the model. The important geometric

Table 2
Elastic modulus (E) and Poisson's ratio (ν) for various bearing materials used in hip implants.

Bearing/Structural materials	***E (GPa)***	ν
PU	0.02	0.4 – 0.5
UHMWPE	0.5 – 1.	0.4
XPE	0.8 – 2.	0.4
Bone cement	2.5	0.25
Bone	0.8 – 17.	0.3
Stainless steel	210	0.3
CoCr	230	0.3
Zirconia	210	0.26
Alumina	380	0.26

parameters required for the lubrication analysis are the radii of both the femoral head (R_H) and the acetabular cup (R_C) from which the radial clearance ($c=R_C - R_H$) can be determined. The geometry of the underlying cement and bone must also be considered because they can sometimes have a significant influence on surface deformation. An equivalent ball-on-plane model can also be used as shown in Figure 3b, with an effective radius determined as follows:

$$\frac{1}{R} = \frac{1}{R_H} - \frac{1}{R_C}$$

or

$$R = \frac{R_H R_C}{R_C - R_H} \qquad (2)$$

Typical values of femoral head radius, radial clearance and equivalent radius are shown in Table 3.

The lubricant in healthy natural joints is synovial fluid and the lubricant in total joint replacements is similar to synovial fluid obtained from patients with osteoarthritis (Saarri et al 1993; Delecrin et al, 1994). Laboratory measurements of viscosity at various shear rates for synovial fluid from osteoarthritic patients have shown powerful non-Newtonian shear thinning characteristics under relatively low shear rates up to 10^3 s^{-1} but this shear thinning declines significantly at higher shear rates (Cooke et al 1978). Thus, according to Jin et al (1997) the non-Newtonian synovial fluid can be approximated with reasonable accuracy as a Newtonian fluid at the very high shear rates which are likely to be experienced in artificial hip replacements under steady walking conditions.

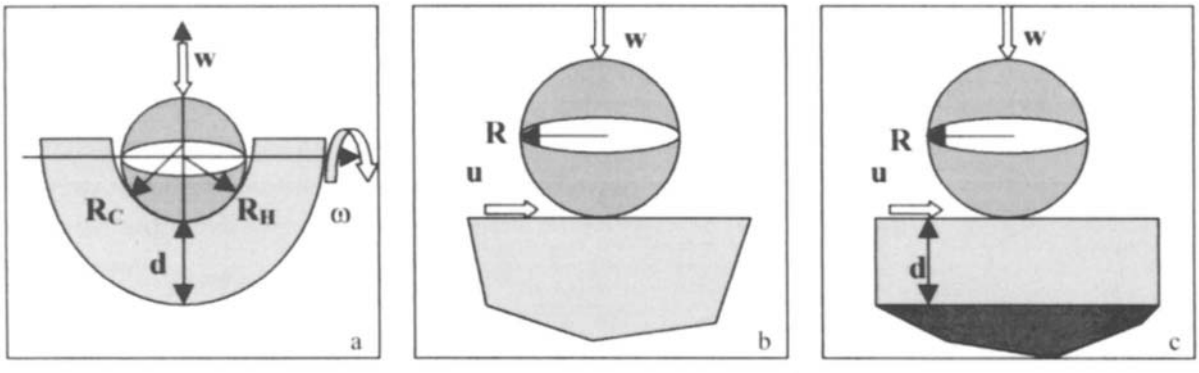


Figure 3. Models for lubrication analysis for hip implants. a) ball-in-socket; b) ball-on-plane semi-infinite solids and c) ball-on-plane layered solid.

Furthermore, the major piezo-viscous effect observed for mineral oils does not appear to occur in synovial fluid from osteoarthritic patients for pressures up to 100 MPa (Cooke et al 1978) and this behaviour is unlikely to change for synovial fluid in arthroplasty patients. This pressure level is also unlikely to be exceeded in fluid films encountered in UHMWPE-on-metal (Jin et al 1999) or even in more rigid metal-on-metal or ceramic-on-ceramic hip implants (Jagatia and Jin 2001; Mak and Jin 2002). Therefore, synovial fluid found in arthroplasty patients is not likely to be influenced by film pressures or by non-Newtonian shear thinning.

The lubricant used for simulator testing is usually bovine serum, diluted to various concentrations. This lubricant also shows significant shear thinning effects (Dowson et al 2003), but a Newtonian, isoviscous lubricant model has still been assumed to represent bovine serum for the majority of lubrication analyses reported in the literature.

A representative viscosity of 0.9 mPas for the 25% bovine serum frequently used in simulator studies has been suggested by Dowson (2002). It is important to distinguish between the effective viscosity of the joint fluid bathing implants in vivo and that of the fluids used in simulator tests. The later is important in the analysis of experimental data on wear, friction and lubrication under carefully controlled conditions, while the former determines the tribological performance of implants in the body. The viscosity of synovial fluid from arthroplasty patients is probably somewhat higher than that of most of the lubricants adopted in simulator tests, as pointed out by Mazzucco et al (2002), and therefore a representative value of between 1.5 and 5 mPas appears to be reasonable for in-vivo evaluations.

Hip implants in vivo are subjected to intermittent motion and in fluid film lubrication modelling, it is important to remember that some direct surface contact frequently occurs no matter how favourable the conditions are with the surfaces in motion (Roter et al, 2002). However, the dynamics and kinematics of cyclic walking is often used to represent in vivo conditions, thus providing transient load and motion patterns for lubrication models and hip simulator testing. During the stance phase, the load can reach up to 5 times bodyweight while the surface velocities are relatively low. During the swing phase, the load is greatly reduced and the entraining velocity increases. In addition, the directions of both the total load and the velocity vectors are transient, although the major load component is in the vertical direction and the major velocity component is in the flexion/extension direction. Figures 4a and 4b show the transient variation of the load in the vertical direction and the resultant angular velocity (mainly in the flexion/extension direction) in hip joints experienced during cyclic walking (Medley et al 2001), which can be used as input for lubrication analysis as shown in Figure 3a. The load and speed employed in hip simulators range from a three-dimensional physiological pattern to simplified versions of specified components, generally consisting of a vertical load and a flexion/extension motion, with an internal-external rotation often added. An average load of between 1200 N (average for a complete cycle) and 2500 N (average in the stance phase), equivalent to about 3 times bodyweight of 750 N and an average resultant angular velocity of about 1.5 rad/s (mainly in the flexion/extension direction) have been suggested for steady-state lubrication analysis under in-vivo conditions (Jin et al, 1997; Medley et al, 2001).

Table 3
Typical values (range) of femoral head radius, radial clearance and effective radius for hip implants with various bearing couples.

Bearing Couples	***Femoral head radius (mm)***	***Radial Clearance (μm)***	***Effective radius (m)***
PU-on-metal[a]	16 (16 – 25[b])	~ 250	1.04 (1 – 2)
UHMWPE-on-metal[c]	14 (11 – 16)	150 (80 – 180)	1.32 (0.5 – 2.5)
XPE-on-metal	14 (14 – 22[d])	150 (100 – 200)	1.32 (1 – 5)
Metal-on-metal	14 (14 – 30[e])	30 (30 – 150)	~ 5
Ceramic-on-ceramic[f]	14 (14 – 16)	40 (10 – 40)	~ 5

[a]Dowson et al (1991); [b]Jennings and Fisher (2002); [c]Jin et al (1999); [d]Muratoglu et al (2001); [e]resurfacing type hip prosthesis; [f]Mak and Jin (2002).

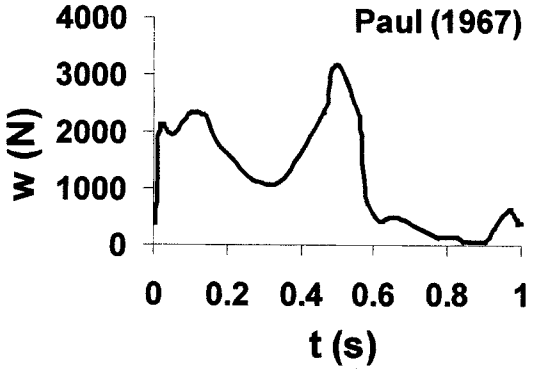


Figure 4a Transient load variation in hip joint during walking for a body weight of 750 N (from Medley et al 2001).

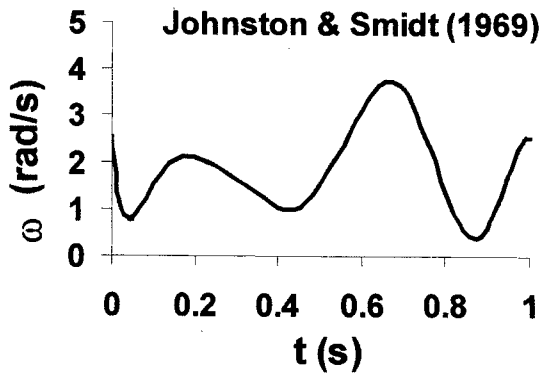


Figure 4b Transient resultant angular velocity in hip joint during walking (from Medley et al 2001).

The governing equations for lubrication analysis of hip implants consist of the Reynolds equation and the elasticity equation, subjected to various boundary conditions. The Reynolds equation may be expressed in Cartesian or spherical coordinates and may or may not include a squeeze velocity term. The elasticity equation depends on the bearing system. For example, a simple column model, which assumes that elastic deformation is directly proportional to local pressure, has been used for UHMWPE-on-metal hip implants (Jalali-Vahid et al, 2000). However, for hard-on-hard bearing couples such as metal-on-metal or ceramic-on-ceramic, finite element methods are often required to obtain the displacement coefficients, which can then be used for calculating elastic deformation and hydrodynamic pressures (Jagatia and Jin 2001).

2.2. Steady-state conditions

A simple estimation of the lubricating film thickness can be made readily under the assumption of average steady-state conditions experienced in hip implants using the equivalent ball-on-plane model shown in Figure 3b. It will be shown in a subsequent section that such a steady-state estimation can give a good approximation to the average lubricating film thickness under more realistic variations of transient loading and motions. The first analysis of these steady-state film thicknesses for various hip implants was carried out by Jin et al (1997) and for metal-on-metal by Medley et al (1996, 1997), using versions of the Hamrock and Dowson (1978) formulae for the isoviscous-elastic lubrication of an equivalent ball-on-plane configuration (Figure 3b). For both metal-on-metal and ceramic-on-ceramic hip implants, the contact area is generally small, and thus the assumption of semi-infinite solids for the bearing surfaces can readily be justified, as can the neglect of the effect of the curvature associated with the ball-in-socket configuration. However, for UHMWPE-on-metal hip implants, the contact dimensions can readily become similar to or greater than the thickness of the UHMWPE cup wall, and the assumption of semi-infinite solids is no longer valid. It is then necessary to consider the elastohydrodynamic lubrication of a layered surface geometry (Figure 3c) as noted by Jin (2000).

If the hydrodynamic pressure distribution extends over a significant proportion of the surface of the acetabular cup, the curvature associated with the ball-in-socket configuration becomes important. The first attempt to address this problem was made by Goenka (1980), who formulated the hydrodynamic lubrication problem in spherical co-ordinates and solved the resultant equations using the finite element method (Goenka and Booker, 1980). The effect of the cup inclination angle was examined recently by Meyer and Tichy (1999). These solutions can only be applied to conforming high elastic modulus bearing surfaces if the elastic deformation of the bearing surfaces is small. Unfortunately, the following condition from Hamrock (1994) indicates that elastic deformation is significant for high elastic modulus surfaces for virtually all realistic in vivo loading patterns (Medley et al 1997);

$$\frac{w^{2.67}}{\left(E' R^5\right)^{0.667} (\eta u)^2} \geq 141.2 \qquad (3)$$

where $E' = \frac{E}{1-\upsilon^2}$.

For UHMWPE-on-metal hip implants, it follows that deformation of the polyethylene bearing surface must be included in an elastohydrodynamic lubrication analysis and it may be significant enough to introduce curvature effects. Jalali-Vahid et al (2000) simultaneously solved the Reynolds equation and an elasticity equation based on the column model in spherical coordinates. Subsequently, a parametric analysis of lubricating film thickness was carried out by Jalali-Vahid et al (2001a, b) and a relatively constant minimum film thickness of between 0.15 to 0.25 μm was predicted, based upon a relatively high angular velocity of 2 rad/s and a viscosity of 0.01 Pas, despite a wide range of design parameters such as the head radius, radial clearance, cup wall thickness and UHMWPE elastic modulus. However, this range of film thickness is considerably less than the initial surface roughness (R_{aC}) of 1 μm that is often quoted for the UHMWPE bearing surface. While the surface roughness can decrease to about 0.05 μm due to wear and creep (see Discussion to Jalali-Vahid et al 2000), the in vivo wear of UHMWPE appears to continue linearly with implantation time. This suggests that mixed rather fluid-film lubrication dominates the tribological performance of such implants. The predicted film thickness range of 0.15 to 0.25 μm was no doubt excessive due to the high value of viscosity adopted in the calculation. Another possible explanation for this continuing wear is the undesirable changes in the inlet zone geometry that also occur because of creep and wear (see Atkinson et al 1985).

A general methodology for elastohydrodynamic lubrication analysis of metal-on-metal hip implants by employing the finite element method for the calculation of the elastic deformation of the bearing surfaces has been presented recently by Jagatia and Jin (2001). For metal-on-metal hip implants with a relatively thick cup wall (9 mm), the predicted minimum lubricating film thickness was found to be similar to that obtained for steady-state conditions using the Hamrock and Dowson formulae. However, for other metal-on-metal bearing systems, the structural deformation of the bearing surfaces can become important. Typical examples include a cobalt chromium alloy acetabular insert located in a titanium fixation shell through a taper (Besong et al 2001) or bonded to an UHMWPE backing (Rieker et al 2001), and the original McKee-Farrar metal-on-metal hip implants (Jin et al 2002). The significant increase in the elastic deformation of the bearing surfaces cannot be neglected in the elastohydrodynamic lubrication analysis of these implant systems (Jagatia and Jin 2002; Jin et al 2002). Figure 5a shows an original McKee-Farrar metal-on-metal hip implant, fixed to the pelvis with bone cement, and Figure 5b shows the corresponding finite element model employed to perform contact mechanics and lubrication analysis (Jin et al 2002).

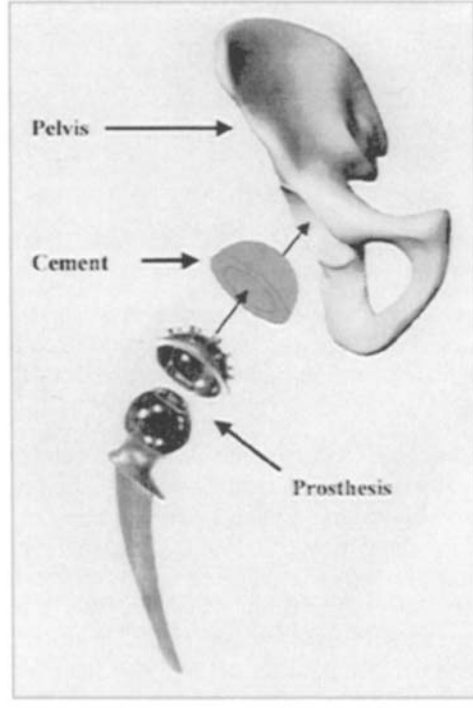


Figure 5a McKee-Farrar metal-on-metal hip implant fixed to pelvis using bone cement.

2.3. Transient lubrication analysis

The transient condition experienced in artificial hip joints during activities such as walking requires the solution of both the Reynolds equation including the squeeze-film motion and the elasticity equation throughout the walking cycle. This usually takes a

significant amount of computing time, and various simplifications are often adopted.

A simplified representation of the transient load and angular speed in steady walking reported by Paul (1967) was employed in the lubrication analysis for rigid spherical and elliptical bearing surfaces

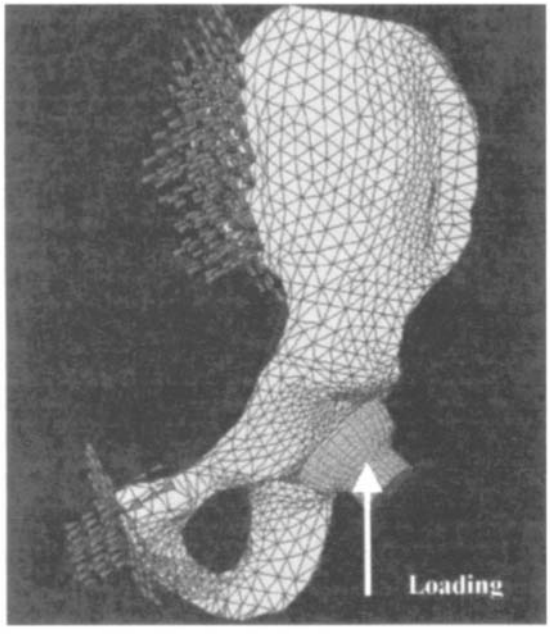


Figure 5b Finite element model used for contact mechanics and lubrication analysis of McKee-Farrar hip implant shown in Figure 5a (Jin et al 2002).

(Goenka, 1980; Goenka and Booker, 1980; Kothari et al 1995). A similar analysis was subsequently reported by Ai and Cheng (1997), where only one loading component and one angular velocity component were considered. A full transient hydrodynamic lubrication analysis was recently presented by Jin and Dowson (1999), in which three-dimensional components of both the load and the angular velocity were considered. The main conclusion from this analysis was that modification of the bearing surfaces in terms of a pocket or recess in the loading direction might significantly increase the lubricating film thickness, whereas the adoption of a symmetrically located recess at the pole had little influence. Furthermore, it was shown that the hydrodynamic pressure was sufficiently high to cause a large elastic deformation of the bearing surfaces, and hence elastohydrodynamic lubrication would be required for hip implants, even those constructed with all metallic or ceramic materials.

Jin et al (1993) developed a simple transient lubrication model for natural hip joints, based upon the linear superposition of loads carried by the entraining and the squeeze-film actions respectively. The film thickness formulae for steady-state entraining motion was combined with the time-dependent squeeze-film equation and then solved for any given transient variations of both the load and the speed. This method was modified and applied to metal-on-metal hip implants tested in hip simulators (Chan et al, 1998) and implanted in patients (Medley et al 2001), based on the assumption of a ball-on-plane semi-infinite solid model. For both simulator and in vivo analyses, it was shown that central film thickness did not vary much over the walking cycle and that using the average load and entrainment velocity gave a central film thickness that was quite close to the lowest value found in the transient analysis (Chan et al, 1998; Roter et al, 2002). Figure 6 shows a typical prediction of the transient central film thickness for a metal-on-metal hip implant throughout one walking cycle under the load and speed given in Figures 4a and 4b, using the methodology outlined by Chan et al (1998).

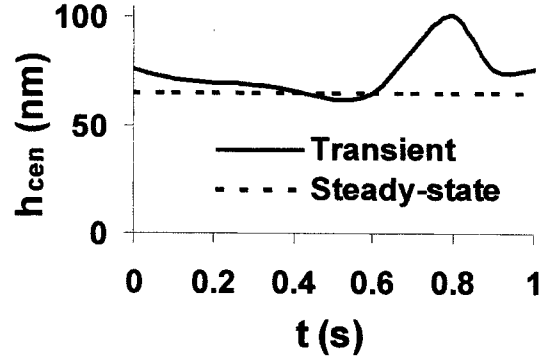


Figure 6 Prediction of transient central film thickness in a metal-on-metal hip implant through a walking cycle (R_H = 14 mm, c = 30 μm, η = 0.005 Pas; steady-state was based on average load and average angular velocity shown in Figures 4a and 4b).

A full transient elastohydrodynamic lubrication analysis for UHMWPE-on-metal hip implants was undertaken by Jalali-Vahid and Jin (2002). A simplified elasticity equation, based on the column model, was employed for calculating the elastic

deformation of the UHMWPE bearing surface in order to save computational time. The main conclusion from this study was that the predicted cyclic minimum and central film thicknesses stayed relatively constant during the cycle due to the combined effect of entraining and squeeze-film actions. Furthermore, it was again pointed out that the predicted film thickness under the quasi-static condition based on the average load and speed was very close to the transient film thickness predicted for a cyclic variation of the load and the speed. However, when load or both load and entrainment velocity increased rapidly, the film thickness in the inlet half of the contact zone was considerably higher than the central film thickness due to spatial differences in local squeeze-film velocities. The simple transient models of Jin et al (1993) and Chan et al (1998) could not predict the true spatial variation in film thickness at an instant in time.

Under certain conditions, the transient lubrication analysis can be de-coupled into a steady-state entraining (Section 2.2) and a time-dependent squeeze-film problem. An elastohydrodynamic squeeze-film lubrication analysis has recently been carried out by Jagatia et al (2001) for UHMWPE-on-metal hip implants, following on the earlier studies by Mabuchi and Sasada (1990) and Wang et al (1990). Such a simplified analysis can be particularly useful for examining the lubrication of hip joints during the stance phase, since the speed encountered is usually very low, and micro-separation between the femoral head and the cup might occur (Dennis et al 2001).

2.4. Discussion

It is clear from the above sections that elastohydrodynamic lubrication analysis is generally necessary for all types of hip implants considered. The use of a ball-on-plane model is reasonably accurate for many types of hip implants. A steady-state lubrication analysis based on the average load and entraining motion often gives a very good approximation to the lowest film thickness during transient analysis for cyclic conditions. Therefore, it is possible to estimate the lowest minimum film thickness likely to be experienced in different artificial hip joints, implanted in patients (Table 4) based on typical parameter values given in Tables 2 and 3.

The mode of fluid-film in implants can be assessed by comparing the estimated theoretical film thickness to typical surface roughnesses of the two bearing surfaces. However, Dowson (2002) has shown that in metal- on-metal total replacement hip joints the correlation between steady-state simulator wear rates and theoretical film thickness is quite good, while that with lambda ratio is often poor. The reason for this is not yet clear, but it may be related to the highly non-Gaussian form of many worn head and cup surfaces. The representative ranges of computed lambda ratios are nevertheless shown in Table 4. The values show that the main mode of lubrication for UHMWPE-on-metal hip implants is between boundary and mixed, for metal-on-metal between mixed and fluid-film and for ceramic-on-ceramic predominantly fluid-film. However, Chan et al (1999) suggested that for two of the twenty-two implants they studied, the surfaces were smooth enough to indicate predominantly fluid-film lubrication. Understanding lubrication mechanisms is important when wear of various hip implants is investigated.

Table 4

Theoretical estimation of in vivo lubrication regimes experienced in various hip implants based on the λ ratio, using the Hamrock and Dowson formulae (R_H = 14 mm; w = 2.5 kN; η= 5 mPa s; ω = 1.5 rad/s).

Bearing Couples	***Minimum film thickness (nm)***	***Composite R_a roughness (nm)***	***λ ratio (lubrication regime)***
UHMWPE-on-metal	83	50 – 1000	0.08 – 1.7 (boundary to mixed)
Metal-on-metal	36	14 – 28	1.3 – 2.6 (mixed to fluid-film)
Ceramic-on-ceramic	24	7	3.4 (fluid-film)

3. EXPERIMENTAL FRICTION AND LUBRICATION STUDIES

3.1. Friction

The measurement of friction experienced in hip implants is usually carried out in the laboratory by using pendulum simulators with the lower stage floating on hydrostatic bearings so that virtually all of the frictional torque comes from the implant. Constant loads were applied in the early pendulum machines (Unsworth et al 1975), but in due course more sophisticated machines subjected to dynamic loads were developed (O'Kelly et al 1977; Roberts et al 1982; Unsworth et al 1987; Auger et al 1993; Scholes and Unsworth 2000). It was often necessary to simplify loading and motion cycles in order to obtain accurate friction measurements, a typical example being a flexion/extension motion under a vertical load. The friction alone often indicates the mode of lubrication, since each mechanism is associated with broad ranges of the coefficient of friction or friction factor defined as $(T/(R_H \mathrm{w}))$ (Scholes and Unsworth 2000). This can be further extended to investigate the variation of the friction factor against the Sommerfeld number, defined as $(\eta u R_H / w)$ in order to assess lubrication regimes. If the measured friction factors remain constant, fall or increase as the Sommerfeld number is increased, the associated modes of lubrication are boundary, mixed or fluid-film respectively. For UHMWPE-on-metal hip implants, the experimental studies by Unsworth et al (1975) clearly indicated boundary or mixed lubrication under steady load, although a suddenly applied load was found to invoke squeeze-film lubrication action. Broadly similar conclusions were reached by O'Kelly et al (1977) employing a hip function simulator under dynamic loading, for both Charnley (diameter 22.225 mm as shown in Figure 2a) and Muller (diameter 32 mm as shown in Figure 2b) hip prostheses. In a more recent study by Scholes and Unsworth (2000), where a simple harmonic oscillatory motion and dynamic loading were applied in a hip function simulator, the friction factors were in the range 0.02 to 0.06 for 28 mm diameter metal heads and UHMWPE cups, broadly representative of mixed lubrication. For the cushion form bearing for hip implants employing a PU-on-metal material combination, a much lower friction factor has been found, representative of fluid-film lubrication (Unsworth et al 1987; Auger et al 1993). For metal-on-metal hip implants, a mixed lubrication regime has been shown to be dominant by Scholes and Unsworth (2000). A similar conclusion was found more recently when a number of large-sized metal-on-metal hip implants (diameter 50 mm) with various radial clearances (20 to 190 μm) were tested (Kampen et al 2002). For ceramic-on-ceramic hip implants, a very small friction factor can be achieved with a synthetic lubricant of similar viscosity to that of synovial fluid, indicative of fluid film lubrication (Scholes and Unsworth 2000). However, the friction factor can be significantly affected by the proteins contained within biological lubricants such as bovine serum.

When the above friction studies are examined, it is surprising that theoretical predictions of film thickness have not been used to predict friction factors over the full transient cycle during dynamic operation and then both theory and experimental results could be compared. In this manner, the accuracy of the experimental measurements along with both the assumptions (such as the negligible influence of protein surface layers) and predictions of the theoretical models could be judged. Such a comparison was initiated by Auger et al (1990) for a compliant surfaced bearing with lubricants that did not contain proteins. Considerable experimental inaccuracies due to mechanical vibrations and slight misalignments of their pendulum simulator were identified in this study. Despite these experimental inaccuracies, there was some indication that the measured friction was behaving in accordance with that predicted by transient elastohydrodynamic modelling. At about the same time, it was shown by Gladstone and Medley (1990) that measured friction forces could be predicted with reasonable accuracy using a simple transient elastohydrodynamic model. In this study, a reciprocating pin-on-plate apparatus was used with a pin of cylindrical surface geometry that was covered by a thin elastomeric layer and conventional mineral oils as lubricants. This apparatus was much less complex than the pendulum simulator used by Auger et al (1990). The coefficient of friction did show a rising characteristic when plotted against a Sommerfeld parameter as film thickness increased, both in the experiments and in the theoretical predictions.

3.2. Lubrication

Separation of the head and cup surfaces by a lubricant in hip implants can be detected directly by applying a small electrical potential (voltage) across the head/cup contact. It is necessary to have

conducting bearing surfaces that are insulated from each other, except when their asperities tips are in contact, thus making the contact itself a resistance element in the circuit. The voltage causes a correspondingly small current flow through asperity contacts. If the surfaces are separated, the film resistance is large enough to prevent any detectable current flow. This approach has proved particularly useful recently for metal-on-metal hip implants (Dowson et al 2000), and even for ceramic-on-ceramic hip implants with appropriate conducting coatings such as titanium nitride (Smith et al 2001a).

The direct measurement of the separation between two bearing surfaces of metal-on-metal hip implants tested in a hip simulator under normal walking conditions was carried out by Dowson et al (2000) using the above electrical resistance technique. The mode of lubrication in their experimental study was mixed, with some periods of very effective surface separation and others of metal-on-metal contact in each simulated walking cycle. Using the same technique, Smith et al (2001a) have demonstrated surface separation of the femoral head and acetabular cup by a lubricating film in ceramic-on-ceramic hip implants. However, it was not clear whether protein films enhanced the surface separation in both of these studies.

This valuable experimental technique does not with certainty indicate whether the lubricating film is hydrodynamic or boundary in nature, although the voltage traces appeared to be consistent with fluid-film lubrication effects. Once again however, transient elastohydrodynamic lubrication modelling was not performed to provide full theoretical predictions over the walking cycle.

The major conclusion from this direct attempt to detect lubricating films in metal-on-metal and ceramic-on-ceramic implants in a hip joint simulator, was that both forms of joint replacement experienced a mixture of fluid-film and boundary lubrication in each cycle.

3.3. Discussion

Experimental assessments of lubrication regimes in hip implants based on friction and separation have proved to be very useful and informative. Typical friction factors and corresponding lubrication regimes for various hip implants are summarised in Table 5. It should be pointed out, however, that the interpretation of the effective lubrication regimes from these experimental measurements needs to be undertaken with caution. As pointed out by Dowson (2001), if only a small proportion of the load is transmitted through asperity contacts, the overall friction can be increased significantly. This is particular important for metal-on-metal hip implants, since large protruding carbides are generally found on cobalt chromium bearing surfaces. It is also difficult to distinguish between boundary and fluid-film layers using the electrical resistance technique and perhaps even using the friction factor measurements.

4. EXPERIMENTAL WEAR STUDIES USING HIP SIMULATORS

A wide range of laboratory equipment, test methods and measuring systems have been used to study wear in total replacement hip joints. The three major forms of equipment used in the last quarter of a century have been;-

(a) Pin-on-Disc machines.
(b) Pin-on-Plate machines.
(c) Joint Simulators.

The pin-on-disc machine has been widely used in tribology and is particularly useful in the evaluation of the nature and wear and friction of material pairs under well controlled, steady-state conditions of load, sliding speed and environment.

The pin-on-plate machine sacrifices the steady sliding speed between specimens, but partially simulates the reciprocating action broadly associated with the hip joint in ambulation.

As interest in the comparative performance of hip joints of different designs and material combinations developed, it became necessary to design and build joint simulators for laboratory studies. These machines simulate to a greater or lesser extent the three dimensional loading and motion patterns experienced by hip joints, while bathing the test joints in a lubricant deemed to be physically and chemically similar to synovial fluid.

Two test stations in an early form of versatile hip joint simulator developed at Leeds in the 1980's, which has formed the basis for a number of subsequent designs, are shown in Figure 7 (Dowson and Jobbins, 1988). The two femoral heads were mounted in the lower framework, through which the flexion-extension and abduction-adduction motions were applied, while the acetabular cups were

Table 5
Measurement of friction factors in a pendulum simulator and corresponding assessment of lubrication regimes (Auger et al 1993; Scholes and Unsworth 2000).

Bearing Couples	*Friction factor*	*Variation of friction factor against increasing Sommerfeld number*	*Indicated Lubrication regimes*
Cushion form bearing	0.002 – 0.01	Increasing	Fluid-film
UHMWPE-on-metal	0.06 – 0.08	Constant/decreasing	Boundary/mixed
Metal-on-metal	0.22 – 0.27	Decreasing	Mixed
Ceramic-on-ceramic	0.002 – 0.07	Increasing	Fluid-film

cemented into the pivoting upper frame which also carried the hydraulic loading cylinders and permitted simulation of pelvic rotation. A third station in the machine allowed the test joint to be loaded dynamically, but not to be subjected to articulation. This permitted the contribution of creep to total penetration to be ascertained.

Hip Joint Simulators have developed enormously, in both sophistication and number during the past quarter of a century. Wear can be evaluated by dimensional or gravimetric means.

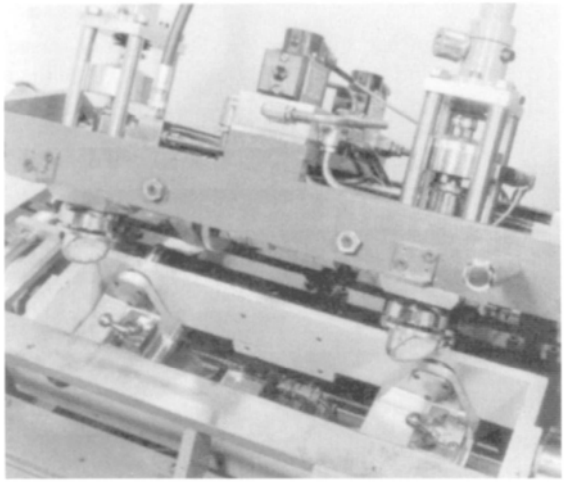


Figure 7 Leeds Hip Joint Simulator (Dowson and Jobbins 1988)

In the 1980's and throughout most of the 1990's few simulators were available, whereas today the laboratories of almost all major orthopaedic companies and several university departments bristle with them. This is a clear indication not only of the value of such devices, but of the international recognition that careful laboratory evaluation of new products should precede clinical trials.

While simulators have been used mainly for wear studies, which themselves reveal much about the prevalent modes of lubrication, they have also been used more recently for direct studies of lubrication (Dowson et al 2000). Some of the essential conclusions about joint lubrication derived from simulator studies on the major forms of implants mentioned earlier are noted below.

4.1. PU-on-Metal

Wear measurements of PU-on-metal hip implants have been carried out at both the University of Leeds (Bigsby et al 1998) and Durham (Smith et al 2000). It has been shown in both studies that the wear rate is significantly reduced, compared with UHMWPE-on-metal material combinations. The volumetric change observed for PU-on-metal bearing couples is attributed mainly to the creep of polyurethane, rather than wear. This is consistent with the beneficial effects of fluid film lubrication in reducing wear experienced in these compliant joints. The advantages of this bearing concept have been further demonstrated in a recent study by Jennings and Fisher (2002) for large-sized components for hip resurfacing procedures.

However, a note of caution must be presented regarding the cushion form concept. The compliant bearing surfaces may be subject to fatigue and delamination caused by their inability to withstand repetitive sub-surface stresses that occur even with full fluid-film lubrication. Such damage was documented in an early study by Auger et al (1995). Although the use of functionally graded compliant layers may alleviate, in theory, the stress discontinuity at the interface between the compliant layer and the substrate (Sunjit and Jin 2002), the experimental verification has not been carried out.

Furthermore, the creep may impair fluid-film formation and promote high start-up friction after rest periods.

4.2. UHMWPE-on-Metal/Ceramic

The majority of simulator studies reported in the literature have been carried out for UHMWPE-on-metal hip implants. Although different simulators with different kinematic and loading input conditions have been adopted, the wear results obtained are relatively consistent, provided that a multi-directional motion path is produced on the UHMWPE bearing surface (Barbour et al 1999; Smith and Unsworth 2000; Kaddick and Wimmer 2001). This finding is consistent with boundary/mixed lubrication regimes experienced in UHMWPE-on-metal hip implants as predicted from the theoretical analysis in Section 2 and deduced from the friction measurement in Section 3. Consequently, the wear in these joints is mainly governed by the adhesive, abrasive and surface fatigue mechanisms associated with boundary/mixed lubrication regimes.

4.2.1. Effect of design parameters

All the simulator studies involving UHMWPE have revealed that an increase in the femoral head radius leads to an increase in the volumetric wear. This is consistent with the boundary/mixed lubrication regime because an increase in the femoral head radius increases the sliding distance and hence wear volume per unit time. The importance of the radial clearance between the femoral head and the cup has been investigated by Wang et al (2001). These authors found that an increase in the radial clearance resulted in a decrease in the initial wear rate, opposite to the general expectation that an increase in the radial clearance would increase the contact stress and hence wear. These simulator results are consistent with those reported previously by Barbour et al (1995) using a simple pin-on-plate wear tester, that an increase in the nominal contact stress could result in a decrease in the wear factor. These experimental observations have been explained on the basis of microscopic elastic asperity contacts between relatively rough UHMWPE and smooth counterface bearing surfaces (Jin and Fisher 2001). A decrease in the nominal contact stress, due to an increase in the nominal contact area, can result in an increase in the real area of contact and hence wear.

4.2.2. Effect of boundary lubricants

Lipids and proteins have been identified as important boundary lubricating constituents from bovine serum and synovial fluid (Hills 2000; Purbach et al 2002). A slight increase in the phospholipid concentration has been shown to lead to a significant reduction in wear rates (Bell et al 2001). The effect of protein concentration has been investigated by Liao et al (1999). The implication of these findings lies in the importance of the concentration of bovine serum used for simulator testing, which can range from 25% to 100%, the lower limit being recommended by ISO/FDIS/14242-1. It is also important to consider the in vivo behaviour of UHMWPE-on-metal/ceramic implants. The biological constituents of synovial fluid and their variation from patient to patient may influence wear significantly. This suggests that there may be possibilities to control the biological composition and tribological behaviour of the synovial fluid of arthroplasty patients.

4.2.3. XPE-on-Metal/Ceramic

There has been significant interest recently in the wear of hip implants employing XPE cups. A remarkable reduction in wear volume has often been observed in simulator studies on XPE. However, there is some controversy regarding the influence of surface roughness on the wear of various XPE cups (McKellop et al, 1999; Endo et al, 2000; Bowsher and Shelton, 2001). In addition, the amount of wear reduction appears to dependent on cross-linking, kinematics, counterface roughness and bovine serum concentration (Galvin et al 2002). More recently, it has been shown that wear is independent of the femoral head sizes up to 46 mm diameter, unlike the conventional UHMWPE cups (Muratoglu et al 2001). Currently, there is still a lack of understanding of wear and lubrication mechanisms experienced with the XPE materials. The early retrievals showed significant creep and plastic deformation of one type of XPE surface but they still seemed to have lower wear than conventional UHMWPE (Muratoglu et al, 2002). Surprisingly, such surfaces had not been found in extensive simulator studies of the same XPE (Muratoglu et al, 1999).

4.3. Metal-on-Metal

The modes of lubrication in metal-on-metal hip implants range from boundary, mixed to fluid-film

depending on both design/manufacturing parameters, and kinematic/loading conditions.

4.3.1. Lambda (λ) ratio

The importance of the lubrication regime on wear volume was first demonstrated by Chan et al (1999). The volumetric wear was shown to correlate fairly well with the λ ratio. An increase in the λ ratio led to an improvement in the lubrication regime and hence a significant reduction in the wear volume. However, even for a given λ ratio, particularly in the mixed lubrication regime, there is usually a large scatter in the wear rates, highlighting the complexity of the lubrication and wear mechanisms involved.

4.3.2. Design parameters

The λ ratio depends on the design parameters such as the radial clearance and the femoral head radius. Regarding the clearance, there appears to be an optimum range, depending on the bearing systems. If the radial clearance is too small, the contact between the two bearing surfaces may extend to the edge of the cup, not only elevating the contact stress but also restricting the lubricant entry and causing lubricant starvation. If the radial clearance is too large, the contact stress may be significantly increased and the mode of lubrication may be shifted towards the boundary lubrication regime. In both cases, wear can increase significantly (Farrar and Schmidt, 1997). The optimum radial clearance was found to be about 30 μm for metal-on-metal hip implants with 28 mm diameter heads (Ultima, made by DePuy Orthopaedics) considered by Farrar and Schmidt (1997). However, Muller (1995) considered the optimal radial clearance to be around 75 μm for Metasul metal-on-metal hip implants of 28 mm diameter heads made by Sulzer. Such conflicting data suggests that the true role of clearance is not well understood. It should be noted that the effective radius shown in equation (2) depends on both the radial clearance and the femoral head radius. An equivalent radius over 4 m was suggested by Medley et al (2001) to achieve effective lubrication in metal-on-metal hip implants. However, the effect of the radial clearance on wear remains illusive. Scholes et al (2001) showed that a reduction in the radial clearance from 40 μm to 22 μm for a fixed femoral head radius of 14 mm resulted in a smaller wear rate, but the difference was not found to be significant.

It is generally expected that the femoral head radius should have a large effect on lubrication and wear, since it is not only related to the effective radius as shown in equation (2), but also the entrainment velocity as well as the sliding distance. This has been confirmed by Smith et al (2001b, c) who showed that an increase in the femoral head diameter from 16 mm to 22.225 mm resulted in an increase in the wear rate, consistent with the boundary lubrication regime. But a further increase in the femoral head radius beyond 22.225 mm led to significant wear reduction, presumably due to improved fluid-film lubrication and reduced asperity contacts (Jin 2002). The benefit of large femoral heads in wear reduction has been further demonstrated in a study of metal-on-metal hip resurfacing prostheses (Bowsher et al 2002).

4.3.3. Effect of kinematic/loading conditions

The λ ratio also depends on the kinematic and loading conditions imposed on hip simulators. Firkins et al (2001b) have shown that in the simulator with two input motions which produced an open elliptical wear path with greater eccentricity, the wear rate was at least ten times higher than that in the simulator with three independent input motions which produced a low level of eccentricity. However, the detailed variation of the motion cycles has been found to have a small effect on wear (Smith et al 2001d).

A new protocol incorporating start-up and stopping has been recently proposed by Chan et al (1999). Other adverse conditions such as micro-separation (Fisher et al, 2002) and stumbling (Bowsher et al 2002) have also been considered. All these adverse conditions generally lead to increases in wear.

4.4. Ceramic-on-Ceramic

Full fluid film lubrication appears to be possible in ceramic-on-ceramic hip implants under most conditions, due to the small clearance and the smoothness of the bearing surfaces. Under normal conditions, extremely low wear rates have been found (Nevelos et al 2001; Cheng et al 2002). However, this is not always reflected in the clinical wear rates and the pattern observed on the retrieved components. Therefore, micro-separation has been introduced into simulator testing and a significant increase in wear has been demonstrated (Nevelos et al 2000). Unfortunately, the presence of a fluid did

not seem to help reduce micro-separation wear and thus improved material microstructures are being sought (Stewart et al, 2002).

4.5. Discussion

Wear rates (volumetric and linear) recorded in the literature are summarised in Table 6 for various implants tested in simulators and retrieved from patients.

5. OVERALL DISCUSSION AND FUTURE STUDIES

5.1. Overall discussion

It is clear from tribological studies of friction, wear and lubrication reported in the literature that different lubrication regimes exist in different forms of hip implants. Furthermore, lubrication mechanisms experienced in these man-made bearings have a strong influence on the wear and the generation of wear particles from the articulating bearing surfaces. Wear debris is now widely recognised as a major factor in clinical problems such as osteolysis and the loosening of implants. Fluid-film lubrication is the most effective method of reducing wear and wear particle generation in most hip bearings. It should be noted, however, that even if a continuous fluid-film is present, wear can still take place due to erosion and fatigue. More importantly, start-up and stopping as well as other adverse conditions such as micro-separation can lead to a significant increase in wear above the values measured in well controlled simulator tests, particularly for hip implants with hard-on-hard bearing couples. The inherent wear resistance of the bearing couples associated with optimal microstructures and surface coatings now plays a more important role in implant performance.

For UHMWPE-on-metal/ceramic hip implants, the mode of lubrication is predominantly boundary, or perhaps mixed under favourable conditions. Changes in the design parameters such as the radial clearance and the femoral head radius are unlikely to change the mode of lubrication and thus the only remaining lubrication mechanism might involve boundary additives. However, a significant wear reduction may be achieved with improved fluid-film lubrication, through modification of the surface topography of the UHMWPE bearing surface.

For metal-on-metal hip implants, the mode of lubrication ranges from boundary to fluid-film including the mixed regime. The lubrication regime experienced in a particular bearing system depends critically on the design parameters such as the radial clearance, the femoral head radius and the bearing structure, as well as the manufacturing of the bearing surfaces in terms of both the sphericity and the smoothness. Design optimization of these parameters in terms of fluid-film lubrication may offer further opportunities to reduce wear and the generation of wear particles.

Ceramic-on-ceramic hip implants appear to have the best chance to benefit from fluid-film lubrication due to their high lambda values. Surface material parameters such as hardness, microstructure and roughness are still important. Further improvement of ceramic-on-ceramic bearings in terms of reducing occasional shock loading within the material or through introducing a soft-mounting system might prove beneficial. It may yet be shown that fluid-film lubrication is important in reducing the wear of ceramic hip components.

5.2. Future studies

The rheology of synovial fluid after total joint replacements, has not been fully characterized, particular under high shear rates (10^5 - 10^7 s^{-1}) likely to be experienced in hip implants under physiological walking conditions.

Table 6
Typical volumetric and linear wear rates for various hip implants.

Bearing Couples	*Volumetric wear rate (mm^3/year)*	*Linear wear rate (μm/year)*
UHMWPE-on-metal	30 – 100	100 – 300
UHMWPE-on-ceramic	15 – 50	50 – 150
Metal-on-metal	0.1 – 1	2 – 20
Ceramic-on-ceramic	0.05 – 1	1 – 20

(One year was assumed to be equal to 1 million cycles)

The roughness of the bearing surfaces plays an important role in lubrication and wear for most types of hip implant. Lubrication analysis that directly includes surface roughness will be important for both UHMWPE/XPE-on-metal/ceramic and metal-on-metal hip implants. In particular for metal-on-metal hip implants, consideration of the surface topography (particularly the valleys) may suggest possibilities for nano-sculptured surfaces to improve lubrication and reduce wear. Further structural optimisation of the bearing surface mounting, in particular for acetabular cups, to achieve continuous fluid film lubrication may offer opportunities to reduce wear to even lower levels. A full transient elastohydrodynamic lubrication analysis will be required for this purpose.

It should be pointed out that the ultimate goal of fluid-film lubrication analysis in hip implants is to provide design guidelines in order to reduce wear and wear particle generation. Further studies should aim at integrating both the analysis of fluid-film and boundary lubrication with wear studies to ultimately develop theoretical models for wear prediction. Such models would offer, in conjunction with simulator testing and other experimental approaches, a screening tool for future development of new bearing couples, thus reducing the time for pre-clinical testing along with costs and improving patient safety. New forms of bearing couples, such as biological surface replacements may become feasible, that would allow even professional athletes in high impact sports to return to their vocations after total hip arthroplasty with some hope of their implants lasting for some 30 - 40 years.

Considerable progress has been made in the development of long lasting total replacement hip joints and the understanding of the underlying tribological factors which influence further progress.

REFERENCES

Ai XL and Cheng HS (1996) Hydrodynamic lubrication analysis of metallic hip joint, Tribology Trans, 39, 103-111.

Amstutz HC, Campell P, Kossovsky N and Clark IC (1991) Mechanism and clinical significance of wear debris-induced osteolysis, Clinical Orthopaedics and Related Research, 276, 7-17.

Atkinson JR, Dowson D, Isaac GH and Wroblewski BM (1985) Laboratory wear tests and clinical observations of the penetration of femoral heads into acetabular cups in total replacement hip joints: iii - the measurement of internal volume changes in explanted Charnley sockets after 2-16 years in vivo and the determination of wear factors, Wear, 104(3), 25-244.

Auger DD, Medley JB, Fisher J and Dowson D (1990) A preliminary investigation of the `cushion bearing' concept for joint replacement implants, Tribology Series 17 (edited by D. Dowson, C.M. Taylor, M. Godet), Elsevier, 251-260.

Auger DD, Dowson D, Fisher J and Jin ZM (1993) Friction and lubrication in cushion form bearings for artificial hip joints. Engineering in Medicine, Proc. Instn. Mech. Engrs., J. Eng. in Med., 207, 25-33.

Auger DD, Dowson D and Fisher J (1995) Cushion form bearings for total knee joint replacement Part 2: wear and durability. Proc. Instn. Mech. Engrs., J. Eng. in Med., 209: 83-91.

Barbour PSM, Barton DC and Fisher J (1995) The influence of contact stress on the wear of UHMWPE for total replacement hip prostheses, Wear, 181 (Part 1), 250-257.

Barbour PSM, Stone MH and Fisher J (1999) A hip joint simulator study using simplified loading and motion cycles generating physiological wear paths and rates, Proc. Instn. Mech. Engrs., J. Eng. in Med., 213 (H6): 455-467.

Bell J, Tipper JL, Ingham E, Stone MH and Fisher J (2001) The influence of phospholipid concentration in protein-containing lubricants on the wear of ultra-high molecular weight polyethylene in artificial hip joints, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H2): 259-263.

Besong AA, Lee R, Farrar R and Jin ZM (2001) Contact mechanics of a novel metal-on-metal total hip replacement, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215, 543-548.

Bigsby RJA, Auger D, Jin ZM and Fisher J (1998) A comparative tribological study of the wear of composite cushion cups in a physiological hip joint simulator, Journal of Biomechanics, 31, 363-369.

Bowsher JG and Shelton JC (2001) Hip simulator study of the influence of patient activity level on the wear of crosslinked polyethylene under smooth and roughened femoral conditions, Wear, 250, 167-179.

Bowsher JG, Nevelos J, Pickard J and Shelton JC (2002) Hip simulator testing - the next generation?, Proc. Int. Conf. Engineers and

Surgeons Joined at the Hip, IMechE, C601/021/2002.

Chan FW, Bobyn JD, Medley JB, Krygier JJ and Tanzer M (1999). The Otto Aufranc Award - Wear and lubrication of metal-on-metal hip implants, Clinical Orthopaedics and Related Research, 369, 10-24.

Chan FW, Medley JB, Bobyn JD and Krygier JJ (1998) Numerical analysis of time-varying fluid film thickness in metal-metal hip implants in simulator tests, In Alternative Bearing Surfaces in Total Joint Replacement, ed. Jacobs JJ, Craig TL, ASTM STP 1346, 111-125, ASTM, West Conshohocken, PA, USA.

Cheng NY, Medley JB, Bobyn JD and Krygier JJ (2001) Lubrication and wear of alumina-alumina hip bearings, Tribology Series 40, Elsevier, 377-387.

Cooke AV, Dowson D and Wright V (1978) The rheology of synovial fluid and some potential synthetic lubricants for degenerate synovial joints, Engng in Medicine, 7, 66-72.

Delecrin J, Oka M, Takahashi S, Yamamuro T and Nakamura T. (1994) Changes in joint fluid after total arthroplasty, Clinical Orthopaedics and Related Research, 307, 240-249.

Dennis DA, Komistek RD, Northcut EJ, Ochoa JA and Ritchie A (2001) "In vivo" determination of hip joint separation and the forces generated due to impact loading conditions, Journal of Biomechanics, 34, 623-629.

Dowson D (2001) New joints for the Millennium: wear control in total replacement hip joints, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H4): 335-358.

Dowson D (2002) Lubrication regimes in total hip replacements, Proc. Int. Conf. Engineers and Surgeons Joined at the Hip, IMechE, C601/051/2002.

Dowson D, Fisher J, Jin ZM, Auger DD and Jobbins B (1991) Design consideration of cushion form bearings in artificial hip joints, Proc. Instn. Mech. Engrs., J. Eng. in Med., 205, 59-68.

Dowson D and Jobbins BJ (1988) Design and development of a versatile hip joint simulator and a preliminary assessment of wear and creep in Charnley total replacement hip joints, Engineering in Medicine, 17(3), 111-117.

Dowson D, Kapur N, McNie CM and Smith SL (2002) Rheology of joint simulator test fluids (to be submitted).

Dowson D, McNie CM and Goldsmith AAJ (2000) Direct experimental evidence of lubrication in a metal-on-metal total hip replacement tested in a joint simulator, Proc. Instn. Mech. Engrs., J. Mech. Eng. Sci., 214, 75-86.

Endo MM, Barbour PSM, Barton DC, Tipper JL, Ingham E and Stone MH (2000) Wear of crosslinked and non crosslinked ultra high molecular weight polyethylene in a hip joint simulator with smooth and damaged femoral heads, Proc 6th World Biomaterials Congress, Vol. 2, Hawaii, 872.

Farrar R and Schmidt MB (1997) The effect of diametral clearance on wear between head and cup for metal-on-metal articulations, 43rd Orthopaedic Research Society, San Francisco, p.71.

Firkins PJ, Tipper JL, Ingham E, Stone MH, Farrar R and Fisher J (2001a) A novel low wearing differential hardness, ceramic-on-metal hip joint prosthesis, Journal of Biomechanics, 34 (10), 1291-1298

Firkins PJ, Tipper JL, Ingham E, Stone MH, Farrar R and Fisher J (2001b) Influence of simulator kinematics on the wear of metal-on-metal hip prostheses, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H1), 119-121.

Fisher J, Ingham E, Matthews B, Stewart T, Stone M, Tipper J and Williams S (2002) Functional biological activity and osteolytic potential of wear debris generated in artificial hip joints, Proc. Int. Conf. Engineers and Surgeons Joined at the Hip, IMechE, C601/011/2002.

Galvin A, Tipper J, Stone M, Ingham E and Fisher J (2002) Reduction in wear of crosslinked polyethylene under different tribological conditions, Proc. Int. Conf. Engineers and Surgeons Joined at the Hip, IMechE, C601/005/2002.

Gladstone JR and Medley JB (1990) Comparison of theoretical and experimental values for friction of lubricated elastomeric surface layers under transient conditions, in Tribology Series 17 (edited by D. Dowson, C.M. Taylor, M. Godet), Elsevier, 241-250.

Goenka PK (1980) Effect of surface ellipticity on dynamically loaded spherical and cylindrical joints and bearings, PhD Thesis, Cornell University, Ithaca, NY.

Goenka PK and Booker JF (1980) Spherical bearings : static and dynamic analysis via the finite

element method, Trans ASME, J Lub Technol, 102, 308-319.

Hamrock BJ (1994) Fundamentals of fluid film lubrication, McGraw-Hill, New York.

Hamrock BJ and Dowson D (1978) Elastohydrodynamic lubrication of elliptical contacts for materials of low elastic modulus. I: fully flooded conjunction, Trans. ASME, J. Lubric. Technol., 100(2), 236-245.

Harris WH (1995) The problem is osteolysis, Clinical Orthopaedics and Related Research, 311, 46-53.

Hills BA (2000) Boundary lubrication in vivo, Proc. Instn. Mech. Engrs., J. Eng. in Med., 214 (H1), 83-94.

Ingham E and Fisher J (2000) Biological reactions to wear debris in total joint replacement, Proc. Instn. Mech. Engrs., J. Eng. in Med., 214(H1), 21-37

Jagatia M, Jalali-Vahid D and Jin ZM (2001) Elastohydrodynamic lubrication analysis of UHMWPE hip joint replacements under squeeze-film motion, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215, 141-152.

Jagatia M and Jin ZM (2001) Elastohydrodynamic lubrication of metal-on-metal hip prosthesis under steady-state Entraining Motion, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215, 531-541.

Jagatia M and Jin ZM (2002) Analysis of elastohydrodynamic lubrication in a novel metal-on-metal hip joint replacement, Proc. Instn. Mech. Engrs., J. Eng. in Med., 216, 185-193.

Jalali-Vahid D, Jagatia M, Jin ZM and Dowson D (2000) Elastohydrodynamic lubrication analysis of UHMWPE hip joint replacements, Proc of 26th Leeds-Lyon Symp on Tribology, 329-339.

Jalali-Vahid D, Jagatia M, Jin ZM and Dowson D (2001a) Prediction of lubrication film thickness in UHMWPE hip joint replacements, Journal of Biomechanics, 34, 261-266.

Jalali-Vahid D, Jagatia M, Jin ZM and Dowson D (2001b) Prediction of lubricating film thickness in a ball-in-socket model with a soft lining representing human natural and artificial hip joints, Proc IMechE, Journal of Engineering Tribology, 215, 363-372.

Jalali-Vahid D and Jin ZM (2002) Transient elastohydrodynamic lubrication analysis of UHMWPE hip joint replacements, Proc. Instn. Mech. Engrs., J. Mech. Eng. Sci., 216 (4), 409-420.

Jennings LM and Fisher J (2002) A biomechanical and tribological investigation of a novel compliant all polyurethane acetabular resurfacing system, Proc. Int. Conf. Engineers and Surgeons Joined at the Hip, IMechE, C601/032/2002.

Jin ZM (2000) Elastohydrodynamic lubrication of circular point contact for a compliant layered surface bonded to a rigid substrate – Part 2 numerical results Proc. Instn. Mech. Engrs., J. Eng. in Med., 14, 281-289.

Jin ZM (2002) Analysis of mixed lubrication in metal-on-metal hip joint replacements, Proc. Instn. Mech. Engrs., J. Eng. in Med., 216, 85-89

Jin ZM and Dowson D (1999) A full numerical analysis of hydrodynamic lubrication in artificial hip joint replacements constructed from hard materials, Proc. Instn. Mech. Engrs., J. Mech. Eng. Sci., 213, 355-370.

Jin ZM and Fisher J (2001) The influence of nominal contact stress on wear of UHMWPE for artificial joint replacements, Proc. of European Society for Biomaterials, 12th-14th Sept., London, T38.

Jin ZM, Dowson D and Fisher J (1993) Fluid film lubrication in natural hip joints, In Thin Films in Tribology, Tribology Series (eds D Dowson, C M Taylor and M Godet) (Amsterdam: Elsevier), pp.545-555.

Jin ZM, Dowson D and Fisher J (1997) Analysis of fluid film lubrication in artificial hip joint replacements with surfaces of high elastic modulus, Proc. Instn. Mech. Engrs., J. Eng. in Med., 211, 247-256.

Jin ZM, Heng SM, Ng H W and Auger DD (1999) An axisymmetric contact model of ultra high molecular weight polyethylene cups against metallic femoral heads for artificial hip joint replacements, Proc. Instn. Mech. Engrs., J. Eng. in Med., 213, 317-327.

Jin ZM, Jagatia M, Yew A and Udofia I (2002) Contact Mechanics and Lubrication Analysis of McKee-Farrar Metal-on-Metal Hip Implants, 48th Annual Meeting of the Orthopaedic Research Society, Paper No. 1015.

Johnston RC and Smidt GL (1969) Measurement of hip-joint motion during walking: evaluation of an electrogoniometric method, Journal of Bone and Joint Surgery, 51(A), 1083-1094.

Kaddick C and Wimmer MA (2001) Hip simulator wear testing according to the newly introduced

standard ISO 14242, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H5), 429-442

Kampen M, Scholes SC and Unsworth A (2002) The lubrication regime in a metal-on-metal total hip replacement, Proc. Int. Conf. Engineers and Surgeons Joined at the Hip, IMechE, C601/035/2002.

Kothari M, Booker JF and Bartel DL (1995) Analysis of artificial hip joints as spherical bearings, Proc of 21st Leeds-Lyon Symp on Tribology, Leeds, 93-98.

Liao YS, Benya PD and McKellop HA (1999) Effect of protein lubrication on the wear properties of materials for prosthetic joints, J. Biomed. Mater. Res. (Appl. Biomater.) 48, 465-473.

Mabuchi K. and Sasada T (1990) Numerical analysis of elastohydrodynamic squeeze film lubrication of total hip prostheses, Wear, 140, 1-16.

Mak MM and Jin ZM (2002) Analysis of contact mechanics in ceramic-on-ceramic hip joint replacements, Proc. Instn. Mech. Engrs., J. Eng. in Med., 216, 231-236.

Mazzucco D, McKinley G, Scott RD and Spector M (2002) Rheology of joint fluid in total knee arthroplasty patients, Journal of Orthopaedic Research, in press.

McKellop H, Shen F, DiMaio, W and Lancaster JG (1999) Wear of gamma-crosslinked polyethylene acetabular cups against roughened femoral balls, Clin. Orth., 369, 73-82.

Medley JB, Bobyn JD, Krygier JJ, Chan FW, Tanzer M and Roter GE (2001) Elastohydrodynamic lubrication and wear of metal-on-metal hip implants, In World Tribology Forum in Arthroplasty, ed. by Claude Rieker et al, Hans Huber, 125-136.

Medley JB, Chan FW, Krygier JJ and Bobyn JD (1996) Comparison of alloys and designs in hip simulator study of metal on metal implants, Clinical Orthopaedics and Related Research, 329(Suppl.), S148-S159.

Medley JB, Dowling JM, Poggie RA, Krygier JJ and Bobyn JD (1998) Simulator wear of some commercially available metal-on-metal hip implants, In Alternative Bearing Surfaces in Total Joint Replacement, ed. Jacobs JJ, Craig TL, ASTM STP 1346, 92-110, ASTM, West Conshohocken, PA, USA.

Medley JB, Krygier JJ, Bobyn JD, Chan FW, Lippincott A and Tanzer, M (1997) Kinematics of the MATCO hip simulator and issues related to wear testing of metal-metal implants, Proc. Instn. Mech. Engrs., J. Eng. in Med., 211, 89-100.

Medley JB, Pilliar RM, Wong EW and Strong AB (1980) Hydrophyllic polyurethane elastomers for hemiarthroplasty, Engng in Medicine, 9, 59-68.

Meyer DM and Tichy JA (1999) Lubrication model of an artificial hip joint: Pressure profile versus inclination angle of the acetabular cup, Journal of Tribology, 492-498.

Muller ME (1995) The benefits of metal-on-metal total hip replacements, Clinical Orthopaedics and Related Research, 311, 54-59.

Muratoglu OK, Bragdon CR, O'Connor D, Jasty M, Harris WH, Gul R and McGarry F (1999) Unified wear model for highly crosslinked ultra-high molecular weight polyethylene (UHMWPE), J. Biomater., 20, 1463-1470.

Muratoglu OK, Bragdon CR, O'Connor D, Perinchief RS, Estok DM, Jasty M and Harris WH (2001) Larger diameter femoral heads used in conjunction with a highly cross-linked ultra-high molecular weight polyethylene - A new concept, Journal of Arthroplasty, 16(8), 24-30 Suppl.

Muratoglu OK, Greenbaum E, Larson S, Jasty M, Freiberg AA, Burk D and Harris WH (2002), Surface analysis of early retrieved acetabular polyethylene liners: A comparison of standard and highly crossslinked polyethylenes, Trans 48th Ann Meeting of the ORS, Vol 27 (2), Poster 1029.

Nevelos JE, Ingham E, Doyle C, Fisher J and Nevelos AB (1999) Analysis of retrieved alumina ceramic components from Mittelmeier total hip prostheses, Biomaterials, 20(19), 1833-40

Nevelos JE, Ingham E, Doyle C, Nevelos AB, and Fisher J (2001) Wear of HIPed and non-HIPed alumina-alumina hip joints under standard and severe simulator testing conditions, Biomaterials, **22**(16), 2191-2197.

Nevelos JE, Ingham E, Doyle C, Streicher R, Nevelos AB, Walter W and Fisher J (2000) Micro-separation of the centres of alumina-alumina artificial hip joints during simulator testing produces clinically relevant wear rates and patterns, The Journal of Arthroplasty, **15**(6), 793-795.

O'Kelly J, Unsworth A, Dowson D, Jobbins B and Wright V (1977) Pendulum and simulator for studies of friction in hip joints, in Evaluation of Artificial Joints, Ed. D Dowson and V Wright, The Biological Engineering Society, 19-20.

Paul JP (1967) Forces transmitted by joints in the human body, in Lubrication and Wear in Living and Artificial Human Joints, Proc. Instn. Mech. Engrs., 181(3J), 8-15.

Purbach B, Hills BA and Wroblewski BM (2002) Surface-active phospholipid in total hip arthroplasty, Clinical Orthopaedics And Related Research, 396, 115-118

Rieker C, Konrad R and Schon R (2001) In vitro comparison of the two hard-hard articulations for total hip replacements, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H2), 153-160

Roberts BJ, Unsworth A and Mian N (1982) Models of lubrication in human hip joints, Ann. Rheum. Dis., 41(3), 217-224.

Roter GE, Medley JB, Bobyn JD, Krygier JJ and Chan FW (2002) Intermittent motion: a novel simulator protocol for the wear of metal-metal hip implants. Tribology Series 40, Elsevier, 367-376.

Saarri H, Santavirta S, Nordstrom D, Paavolainen P and Konttinen YT (1993) Hyaluronate in total hip replacement, J. Rheumatol., 20(1), 87-90.

Scholes SC and Unsworth A (2000) Comparison of friction and lubrication of different hip prostheses, Proc. Instn. Mech. Engrs., J. Eng. in Med., 214(1), 49-57.

Scholes SC, Green SM, Unsworth A (2001) The wear of metal-on-metal total hip prostheses measured in a hip simulator, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H6), 523-530.

Smith SL, Ash HE and Unsworth A (2000) A tribological study of UHMWPE acetabular cups and polyurethane compliant layer acetabular cups, Journal of Biomedical Materials Research, 53 (6): 710-716.

Smith SL, Dowson D, Goldsmith AAJ, Valizadeh R and Colligon JS (2001a) Direct evidence of lubrication in ceramic-on-ceramic total hip replacements, Proc. Instn. Mech. Engrs., J. Mech. Eng. Sci., 215 (3), 265-268

Smith SL, Dowson D and Goldsmith AAJ (2001b) The lubrication of metal-on-metal total hip joints: a slide down the Stribeck curve, Proc. Instn. Mech. Engrs., J. Eng. Tribology, 215 (J5): 483-493.

Smith SL, Dowson D and Goldsmith AAJ (2001c) The effect of femoral head diameter upon lubrication and wear of metal-on-metal total hip replacements, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H2), 161-170.

Smith SL, Dowson D and Goldsmith AAJ (2001d) The effect of diametral clearance, motion and loading cycles upon lubrication of metal-on-metal total hip replacements, Proc. Instn. Mech. Engrs., J. Mech. Eng. Sci., in Med, 215 (1), 1-5.

Smith SL and Unsworth A (2000) Simplified motion and loading compared to physiological motion and loading in a hip joint simulator, Proc. Instn. Mech. Engrs., J. Eng. in Med., 214, 233-238.

Stewart T, Tipper JL, Insley G, Streicher RM, Ingham E and Fisher J (2002) The performance of new ceramic articulations in hip simulator studies with microseparation, Proc. Int. Conf. Engineers and Surgeons Joined at the Hip, IMechE, C601/038/2002.

Virdee SS and Jin ZM (2002) Contact mechanics of a graded layered bearing surface for knee prosthesis, International Journal of Applied Mechanics and Engineering, 7(3), 1025-1032.

Unsworth A, Dowson D, Wright V and Koshal D (1975) The frictional behaviour of human synovial joints – Part II. Artificial joints, Trans ASME, Journal of Lubrication Technology, 97(3), 377-382.

Unsworth A, Pearcy MJ, White EFT and White G (1987) Soft layer lubrication of artificial hip joints, Proc. Int. Conf. 'Fifty Years On', Mechanical Engineering Publications, London, 715-724.

Wang CT, Wang YL, Chen QL and Yang MR (1990) Calculation of elastohydrodynamic lubrication film thickness for hip prostheses during normal walking, Tribology Transactions, 33, 239-245.

Wang A, Essner A and Klein R (2001) Effect of contact stress on friction and wear of ultra-high molecular weight polyethylene in total hip replacement, Proc. Instn. Mech. Engrs., J. Eng. in Med., 215 (H2), 133-139.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Wear of alumina/zirconia nano-composite for ceramic/ceramic joint prostheses

K. Ikeuchi[a], Y. Morita[a], K. Nakata[a], Y.H. Kim[b], T. Sekino[b] and K. Niihara[b]

[a]Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan
[b]The Institute of Scientific and Industrial Research, Osaka University, Osaka 567-0047, Japan

Wear properties of alumina, zirconia and alumina/zirconia nano-composite ceramics were investigated with two types of wear apparatus. An end face apparatus was used to measure wear between flat surfaces in distilled water. Results show that wear of alumina is lower than that of zirconia under poor lubrication conditions. According to the observed wear surfaces, adhesion and micro-ploughing are the major wear mechanism for zirconia. The wear rate under concentrated contact pressure was measured with a sphere on flat apparatus in bovine serum solution. In contrast to the former test, alumina shows higher wear rate than zirconia due to micro-cracking. Wear of alumina/zirconia nano-composite ceramic is lower than alumina and zirconia monolithic ceramics during both tests. In other words, the nano-composite ceramics shows higher wear resistance irrespective of lubrication conditions and contact pressure. It is concluded that the alumina/zirconia nano-composite ceramics have excellent properties for ceramic/ceramic joint prostheses.

1. INTRODUCTION

Wear test with alumina/alumina combination shows low friction [1,2] and high wear resistance [3] in physiological environment if geometric conformity is high. On the other hand, stronger materials are desired for a ceramic/ceramic total joint prosthesis to prevent accidental fracture. Though improved zirconia is stronger than alumina, the wear rate of zirconia is excessively high [3] because severe adhesion arises if the effect of lubrication is not enough.

Alumina/zirconia nano-composite ceramics [4] seem to be a potential material for joint prostheses of the next generation because they are stronger than alumina and zirconia monoliths. Alumina crystals and zirconia crystals coexist in a component of composite ceramic because the mutual solubility is low. The crystals are smaller than the ones in a monolithic material. Bending strength and fatigue strength are improved due to the decrease in crystal size. Alumina/zirconia nano-composite ceramics have good mechanical properties for joint prosthesis while their wear resistance in a physiological environment has not been investigated yet.

The purpose of this study was to investigate the wear properties of alumina/zirconia nano-composite ceramics for ceramic/ceramic total joint prostheses.

2. MATERIAL AND METHOD

Wear rate was measured for square discs of alumina, zirconia and two kinds of alumina/zirconia nano-composite ceramics. The mating face was made of alumina (BIOCERAM NPX: Kyocera Corporation) with the same quality as the alumina femoral components. Mechanical properties of the test materials and the reference material are listed in Table 1. Typical grain size of the alumina was 3 μm and that of the zirconia was 0.8 μm. The grain size of composite ceramics was similar to zirconia.

The centerline average roughness of the sliding surface was less than 0.02 μm before the test. Weight of the specimen was measured before and after the test and the wear factor W was calculated by the following equation [5].

$$W = \frac{V}{L \times D} \qquad (1)$$

Where V is wear volume in mm^3, L is applied load in Newtons and D is sliding distance in mm.

An end face apparatus (Fig. 1) was used to measure wear rate between nominally flat surfaces with minimum lubrication effects. The upper disc is made of alumina, zirconia or alumina/zirconia nano-composite ceramics. The mating disc of medical

grade alumina is regarded as the reference specimen. Its outside diameter is 10 mm and the inside diameter is 4 mm. The lower specimen rotates in one direction with constant velocity. Vertical load of 660 N is applied to the upper specimen. Mean contact pressure is 10 MPa, mean sliding velocity is 40 mm/s and total sliding distance is 10 km. The wear test was performed in distilled water kept at 37 °C while wear volume was calculated by the weight loss and the density.

A pin on flat apparatus (Fig. 2) was used to measure wear rate under concentrated contact and reciprocating motion. The square plate is made of alumina, zirconia or alumina/zirconia nano-composite ceramics. The pin of medical grade alumina had an end face with 10 mm spherical radius. The lower plate oscillates while the vertical load is applied to the pin by dead weights. Friction force is measured with strain gauges attached to the double leaf springs. Applied load is 30 N, oscillation frequency is 2 c/s, stroke is 10 mm and mean sliding velocity is hence 40 mm/s. Wear test was performed in 30 vol.% bovine serum solution kept at 37 °C. Wear volume of the pin was calculated from the area of the wear surface while the wear volume of the disc was calculated by the weight loss.

According to Herzian contact theory, the maximum contact pressure is given by the following equation [6].

$$P_{max} = 0.364 W^{1/3} E'^{2/3} R^{-2/3} \qquad (2)$$

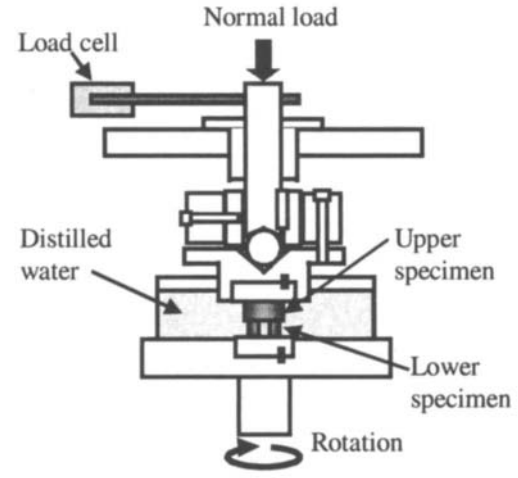


Figure 1. Cross sectional view of end face apparatus.

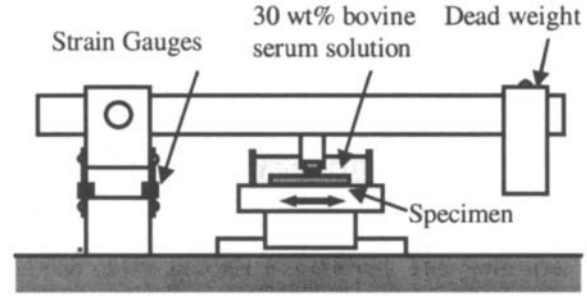


Figure 2. Cross sectional view of pin on flat apparatus.

Table 1
Properties of ceramics

	Density (g/cm^3)	Elastic modulus (GPa)	Poisson's ratio	Hardness (GPa)	Fracture toughness ($MN/m^{1.5}$)	Bending strength (GPa)
3Y-TZP	6.07	214	0.32	12.6	5.9	1.57
Al_2O_3 (TM-DAR)	3.95	365	0.29	16.0	3.3	0.62
TZP/30vol% Al_2O_3 (CZA)	5.37	264	0.30	14.3	5.4	1.86
Al_2O_3/30vol% TZP (CAZ)	4.59	346	0.27	16.6	4.3	1.35
Reference Al_2O_3	3.97	400	0.30	19.0	4.0	0.64

Where W is the applied load, E' is the effective contact modulus and R is effective spherical radius. E' is expressed with elastic moduli and Poisson's ratios of the mating materials as follows.

$$\frac{1}{E'}=\frac{1}{2}\left(\frac{1-\nu_1^2}{E_1}+\frac{1-\nu_2^2}{E_2}\right) \qquad (3)$$

In the test with the pin on disc apparatus, applied load W is 30 N and effective spherical radius R is10 mm. For the alumina/alumina combination, effective contact modulus E' calculated by Eq. (3) is 418GPa and maximum contact pressure P_{max} calculated by Eqs. (2) and (3) is 1.36 GPa at the beginning of the wear test. In the same manner, P_{max}=1.11 GPa for alumina/zirconia, P_{max}=1.33 GPa for alumina/CAZ, P_{max}=1.21 GPa for alumina/CZA, where CAZ indicates the composite of 70 wt% alumina and 30 wt% zirconia, and CZA is composed of 70 wt% zirconia and 30 wt% alumina. The initial maximum contact pressure in the test with the pin-on disc apparatus is higher than 100 times of the mean contact pressure for the test between flat surfaces. However, the contact pressure reduces with an increase of contact area by wear.

3. RESULT

Fig. 3 shows the wear factor measured with the end-face apparatus. The left bars indicate the wear factor of the specimens while the right bars indicate the wear factor of the mating face of alumina. The wear factors are very low for alumina/alumina, CAZ/alumina and CZA/alumina combinations. The wear factors of the reference alumina are lower than the ones of test ceramics. For zirconia/alumina combinations, catastrophic wear was measured both on the zirconia and the alumina surfaces.

Fig. 4 shows the wear factors of the ceramic plates for sliding with the alumina pins. In contrast to the previous case, excessive wear arises for alumina plate while wear of zirconia is lower than the case of alumina/alumina combination. Wear rates of alumina/zirconia nano-composite ceramics are much lower than those of alumina and zirconia monolithic ceramics. Fig. 5 shows that wear rate of the alumina pin is highest when sliding with zirconia disc. This figure also shows that the wear of alumina pin is very low when sliding with alumina/zirconia nano-composite ceramics.

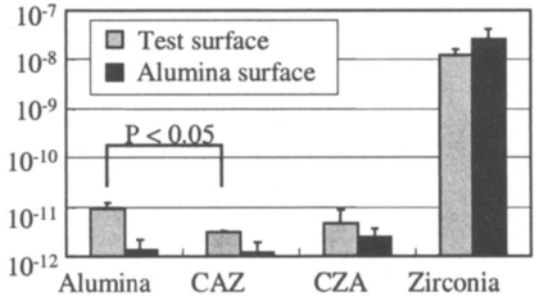


Figure 3. Wear factor of ceramics sliding with alumina disc in distilled water with end face apparatus.

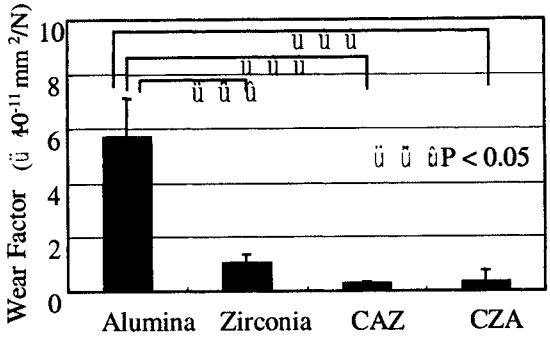


Figure 4. Wear factor of ceramic discs sliding with alumina spherical pins in bovine serum solution.

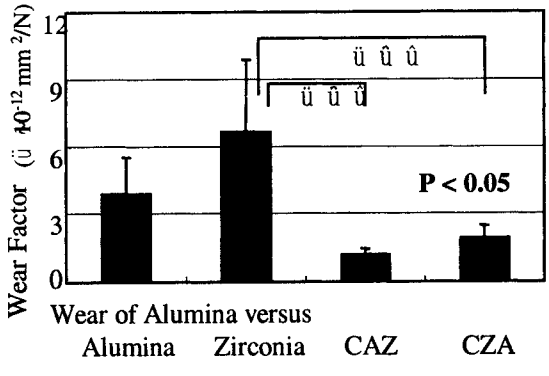


Figure 5. Wear factor of alumina spherical pins sliding with ceramic discs in bovine serum solution.

4. DISCUSSION

Rate of wear for alumina/zirconia combination is excessively high when sliding with the end-face apparatus in distilled water. According to the observation, zirconia begins to adhere to the mating surface of alumina with a start of sliding if lubrication condition is poor. Then the zirconia wear particles cause third body abrasion on both surfaces. As zirconia is softer than alumina, micro-ploughing arises on the zirconia surface (Fig. 6) while micro-cracking develops on the alumina surface (Fig. 7) because alumina is a brittle material. The increase of wear particles causes further grooving wear on the zirconia surface and crack initiation and propagation on the alumina surface respectively. The cooperation of these processes leads to catastrophic wear on the both materials.

High wear rate of zirconia under poor lubrication condition attributes partly to the instability of its crystal structure [7]. During sliding with high friction, a part of the zirconia crystals transform from tetragonal to monoclinic structure due to high friction force. The crystal transformation is followed by volume expansion, loss of strength and increase of wear.

Wear of zirconia decreases drastically by lubricating with bovine serum solution even if concentrated contact pressure is applied because severe adhesion is prevented by proteins or phospholipids. Therefore, it may be concluded that the boundary lubrication is inevitable for a joint prosthesis with zirconia surfaces to keep the wear rate low.

Wear of alumina is kept low even if lubrication effect does not appear. Perhaps, this is a reason why only alumina is applied to ceramic/ceramic total hip prostheses at present. On the other hand, the result with the spherical pin on flat apparatus reveals that stress concentration causes micro cracks and detachment of wear particles from the surface (Fig. 9) because alumina is a brittle material. It is possible to reduce stress concentration in alumina at normal condition by improving geometric conformity of the articulating surface of a joint prosthesis by precise finish. However, the possibility of joint fracture still remains in the case of impingement.

Bending strength and hardness of alumina/zirconia nano-composite ceramics are improved because the crystal size of the composite ceramics are smaller than alumina and zirconia ceramics. Hence it is unlikely that micro-ploughing

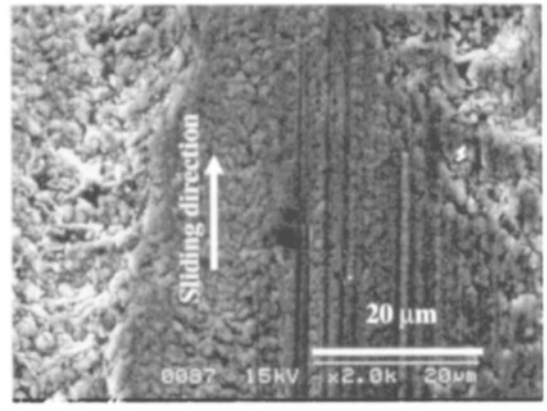


Figure 6. Micro-ploughing on zirconia surface after wear test with end face apparatus.

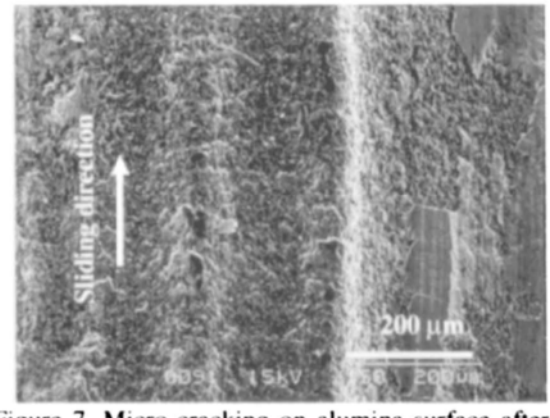


Figure 7. Micro-cracking on alumina surface after wear test with end face apparatus.

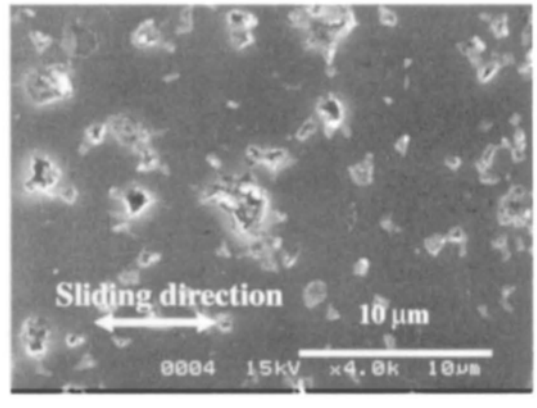


Figure 8. Wear surface of alumina after wear test with pin on flat apparatus under concentrated contact pressure.

observed on the zirconia surface or the micro-cracking on the alumina surface takes place on the surface of alumina/zirconia nano-composite ceramics. During the hot press sintering at 1500 °C, alumina and zirconia do not dissolve to each other. In the nano-composite ceramics, the crystals of alumina and the crystals of zirconia coexist and their sizes are much smaller than those of alumina and zirconia ceramics in monolithic conditions. After the experiment with the end face apparatus, wear surface of CAZ (composite of alumina (70%) and zirconia (30%)) and CZA (composite of zirconia (70%) and alumina (30%)) are smooth because the size of the wear particles detached from the surface is small. We suppose that the alumina nano-scale crystals prevent the crystal transformation of zirconia in the zirconia based nano-composite ceramic. In summary of the result, alumina/zirconia nano-composite ceramics show lower wear rate than alumina and zirconia monolithic ceramics irrespective of lubrication condition and stress distribution.

5. CONCLUSION

According to the result of the wear test, alumina/zirconia nano-composite ceramics show lower wear rate than alumina and zirconia monoliths in the wide range of lubrication condition and contact pressure. As hardness and strength of the composite ceramics are higher than the values prediction by the monolithic ceramics, a joint prosthesis of the composite material seems to have longer life and higher reliability than the existing alumina/alumina total hip prostheses.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support of Japanese Ministry of Health, Labour and Welfare. We would also like to acknowledge Kyocera Corporation for providing the medical grade alumina components as the reference material.

REFERENCES

1. Y.S. Zhou, M. Ohashi and K. Ikeuchi, Wear, 210 (1997) 112-119.
2. Y.S. Zhou, K. Ikeuchi and M. Ohashi, Wear, 210 (1997) 171-177.
3. J. Kusaka, K. Takashima, D. Yamane and K. Ikeuchi, Wear, 225-229 (1999) 734-742.
4. E. Enomoto, J. Tang, et al., J. Ceram. Proc. Res., 1 (2) (2000) 88-91.
5. J.R. Atkinson, K.J. Brown and D. Dowson, Wear, 100 (1978) 208-218.
6. A.Z. Szeri, Fluid Film Lubrication, Cambridge University Press (1998) 13.
7. K. Shimizu, M. Oka, et al., J. Biomed. Mat. Res., 27 (1993) 729-734.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Surface Characteristics and Tribological Behaviour of New Zirconia Ceramics for Cementless Knee Joint Prostheses

Y. Sawae[a], T. Murakami[a], T. Sawano[a], I. Noda[b] and T. Shimotoso[b]

[a] Faculty of Engineering, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

[b] Bioceram Division, Kyocera Corporation, 10-1 Kawai, Gamo-cho, Gamo-gun, Shiga 529-1595, Japan

Hydroxyapatite and Titanium on Zirconia (HTOZ) is a newly developed zirconia composite material for use in the femoral component of cementless knee prosthesis. The material has a tetragonal zirconia polycrystal (TZP) substrate covered with an arc-spray coating layer of titanium and a flame-spray coating layer of hydroxyapatite on the fixation side. These surface layers provide good conductivity of bone formation and enhance the integration of the ceramic femoral component to living bone. Although the mechanical properties of the TZP substrate is not affected by the manufacturing process of applying the coating layers on the fixation side, processing causes changes to the articulating surface TZP microstructure which makes the surface smoother microscopically and also hydrophobic. In this study, the friction and wear behaviour of the UHMWPE-on-HTOZ sliding pair for the cementless knee prostheses was evaluated using unidirectional pin-on-disk and multidirectional pin-on-plate tests.

Results indicated that the UHMWPE-on-HTOZ sliding pair produced favourable tribological characteristics compared to UHMWPE-on-TZP. Friction was lower than that of the conventional UHMWPE-on-TZP sliding pair under the same physiological lubrication conditions. Under multidirectional sliding the wear of polyethylene was, similar to the wear of UHMWPE-on-ceramic. It is postulated that the hydrophobic and microscopically smooth surface of HTOZ contributed to its improved friction and wear behaviour through enhanced fluid film formation compared to conventional TZP. The results support the use of the HTOZ-on-UHMWPE bearing combination for knee prostheses.

1. INTRODUCTION

Ceramic-on-UHMWPE (Ultra-High Molecular Weight Polyethylene) sliding pairs have recently been developed for use in artificial knee prostheses as an alternative to metal-on-UHMWPE sliding pairs. The ceramic femoral component of the artificial knee has several advantages over the conventional metal component, the most important of which is reduced wear of the UHMWPE tibial and patella components [1,2]. UHMWPE wear particles are generally believed to activate macrophages and lead to bone resorption and tissue necrosis [3-5]. These histological changes within the surrounding tissue and bone induce aseptic loosening and eventual failure of joint prostheses. Therefore, it is widely recognised that the amount of UHMWPE wear debris released into the joint capsule is a major factor limiting the long-tern clinical performance and the joint lifetime of artificial joint prostheses.

Current ceramic femoral components used in knee prostheses generally require bone cement to connect the implant to living bone. However, metal femoral components are available in cementless varieties which incorporate a plasma-spray hydroxyapatite (HA) coating or a porous bead coating on the fixation side to promote bone growth and avoid the potential adverse effects of bone cement, such as the hypersensitivity [6]. In addition, it is generally considered that hard bone cement particles released from the bone-prosthesis interface can increase the polyethylene wear significantly if they flow into the conjunction. Therefore, the lack of cementless fixation in ceramic knee prostheses has limited their use.

Recently, a new zirconia composite material (HTOZ) has been developed that can be used without bone cement [7]. Hydroxyapatite and Titanium On Zirconia (HTOZ) consists of a tetragonal zirconia polycrystals (TZP) substrate with the fixation interface covered with an arc-spray coating layer of titanium and flame-spray coating layer of hydroxyapatite. These surface layers provide good conductivity of bone formation and improve the integration of the ceramic femoral component with living bone. It is reported that the manufacturing process of HTOZ has little effect on mechanical properties of the TZP substrate [7]. However, changes in the surface characteristics may influence the tribological behaviour of the manufactured joint prostheses and also the resultant joint lifetime.

In this study, the tribological behaviour of the HTOZ-on-UHMWPE sliding pair for cementless knee joint prostheses was evaluated using unidirectional pin-on-disk tests and multidirectional pin-on-plate tests. Results were compared to the friction and wear properties of the TZP-on-UHMWPE sliding pair and the effects of the altered surface characteristics were examined.

2. MATERIALS AND METHODS

2.1. HTOZ

HTOZ is a ceramic composite material for the femoral component of cementless knee prostheses developed by Noda *et al* [7]. The substrate of the HTOZ femoral component is TZP, the outer surface of which is polished and articulates against the UHMWPE tibial and patella components (Fig 1). The inner surface of HTOZ femoral component is covered with a HA coating layer for cementless fixation.

Hydroxyapatite is a type of calcium phosphate ceramic. As it has good conductivity of bone formation and can make a firm chemical bonding to living bone, HA coating has been widely applied to metallic medical implants, such as joint prosthesis and dental implants, to improve the integration with bone [8]. However, HA coatings have never been applied to ceramics materials, such as alumina and zirconia, for attachment directly to bone.

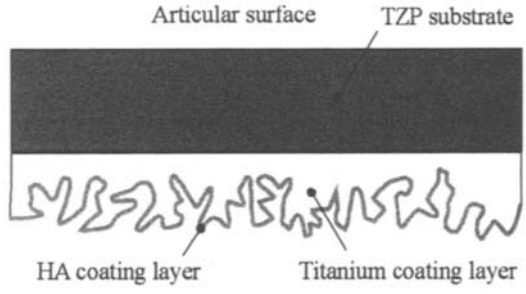


Figure 1. Schematic structure of HTOZ.

In previous studies, Noda *et al* successfully applied a HA coating layer on the surface of TZP by using a titanium intermediate layer. In the manufacturing process of HTOZ, a titanium layer was initially applied to the TZP substrate by arc-spraying. The HA coating layer was then deposited on this titanium foundation by flame-spraying. The large surface roughness of the initial titanium layer produced a mechanical anchoring effect for the HA coating which enhanced the integration with bone [7].

During the titanium arc-spray process, small numbers of oxygen atoms moved from the zirconia substrate and were deposited onto the titanium layer. As a result, the TZP substrate of HTOZ discolored from white to black. These manufacturing processes had little effect on material properties of the TZP substrate. Material properties of the TZP substrate

Table 1
Comparison of material properties

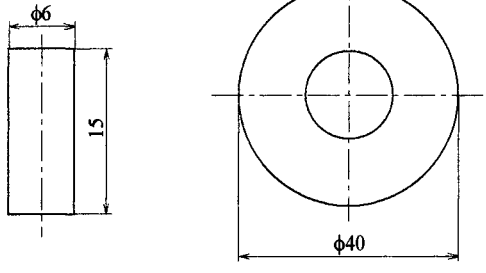
	HTOZ	TZP
Density [g/cm^3]	6.08	6.08
Average grain size [μm]	0.2	0.2
Hardness [HV]	1400	1400
Bending strength [MPa]	1214	1099
Fracture toughness [MN/m$^{1.5}$]	5	5
Thirmal conductivity [W/mK]	2.8	2.8
Contact angle [degree]	72.4	52.2

are compared with that of the original TZP in Table 1. There was little difference in both mechanical and thermal properties. However, the surface characteristic of the TZP substrate was slightly changed from the original TZP surface. The contact angle of distilled water became larger on the HTOZ surface compared with the original TZP surface. The articulating surface of the HTOZ, therefore, became significantly hydrophobic compared with the original TZP surface during the manufacturing process.

2.2. Unidirectional Sliding Pin-on-Disk Test

Friction characteristics of the HTOZ-on-UHMWPE sliding pairs were evaluated in unidirectional sliding tests using a two-channel pin-on-disk test apparatus [9]. Cylindrical UHMWPE pin specimens were fixed to the loading shafts and articulated under unidirectional motion against disk specimens, which were installed in the liquid baths and rotated by ring-cone motor.

The geometries of the UHMWPE pin specimen and the HTOZ disk specimen are shown in Fig.2. Cylindrical UHMWPE pin specimens were prepared from a GUR415 block. These polyethylene pin specimens were not sterilized. Prior to tests, both pin and disk specimens were washed in an ultrasonic cleaner with a detergent (polyoxyethylene *p-t*-octylphenylether) for 30min. Then the specimens were rinsed in a stream of distilled water and washed in the ultrasonic cleaner again for 30min with distilled water. Finally, the specimens were washed ultrasonically in ethanol for 15min and dried in a vacuum desiccator.



(a) Pin specimen (b) Disk specimen

Figure 2. Geometries of UHMWPE pin specimen and ceramic disk specimen.

All pin-on-disk tests were conducted under the room temperature. A constant axial load of 260N was applied to the pin specimen. It gave rise to a mean contact pressure of 9MPa. The sliding speed was set initially at 10mm/s and increased to 20, 30 and 40mm/s. These sliding speeds are in the range of the physiological sliding speed of joint prostheses. Friction between the pin and the disk specimen was measured by a load cell and data from the load cell were accumulated and analysed on a PC.

Frictional behaviour was examined under both lubricated and un-lubricated conditions. Distilled water was used as lubricant in the lubricated tests. In addition, silicone oil with a kinematic viscosity of $1mm^2/s$ and fetal bovine serum were also used as lubricants. Fetal bovine serum was diluted to 30vol% with distilled water to give the similar protein content to synovia, which played a role of lubricant for joint prostheses *in vivo*, and 0.3wt% sodium azide was added to retard bacterial growth. The viscosities of these lubricants are as small as distilled water.

Frictional characteristic of the UHMWPE-on-TZP sliding pair of current joint prostheses was also examined in this test for comparison. The TZP used in this test was the same material used in current ceramic joint prostheses and received hot isostatic pressing (HIP) treatment to reduce internal defects and improve the mechanical strength. The surface roughness of the HTOZ and TZP disk specimens measured by the stylus profilometer was 0.01μm (Ra).

2.3. Multidirectional Sliding Pin-on-Plate Test

The wear behaviour of UHMWPE sliding against HTOZ was examined quantitatively in a multidirectional sliding wear test [10]. A schematic drawing of the two-channel pin-on-plate test apparatus used in this wear test is shown in Figure 3. In this test apparatus, UHMWPE pin specimens are mounted on an X-Y stage with air cylinders and moved two-dimensionally. Air cylinders, controlled by regulators and solenoid valves, press the pin specimens against the plate specimens. The UHMWPE pin specimens are moved two-dimensionally by the X-Y stage and articulate

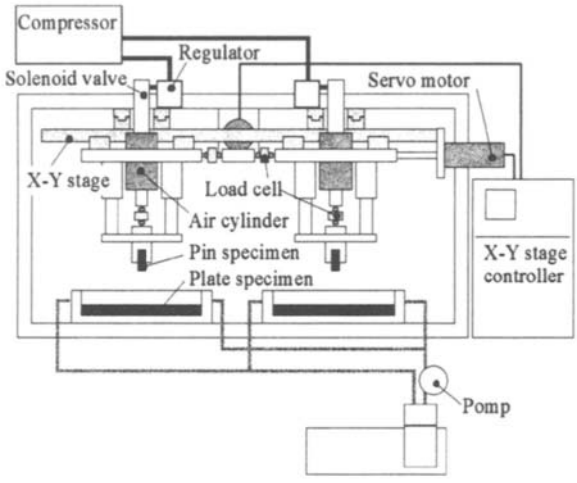


Figure 3. Schematic drawing of pin-on-plate test apparatus for the multidirectional sliding wear test.

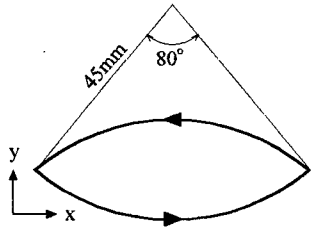


Figure 4. Two arcs shape of the sliding path way used in the multidirectional sliding wear test.

against plate specimens along a two-dimensional sliding pathway with various shapes under constant load. The shape of the sliding pathway used in this study consisted of two arcs with a diameter of 90mm (Fig.4) and length of 125.7mm. This shape is similar to the physiological shape of the sliding motion of an artificial hip joint [10-12].

The same UHMWPE pin specimens and the same specimen cleaning procedure were used in both the multidirectional sliding test and the unidirectional sliding test described previously. In all the tests, the mean contact pressure was 10MPa and the magnitude of the sliding velocity vector was 45mm/s. Diluted bovine serum (30vol% fetal bovine serum + 70vol% distilled water + 0.3wt% sodium azide) was used as the lubricant and its temperature was controlled at around 37C. Each test was run for a sliding distance of 50km.

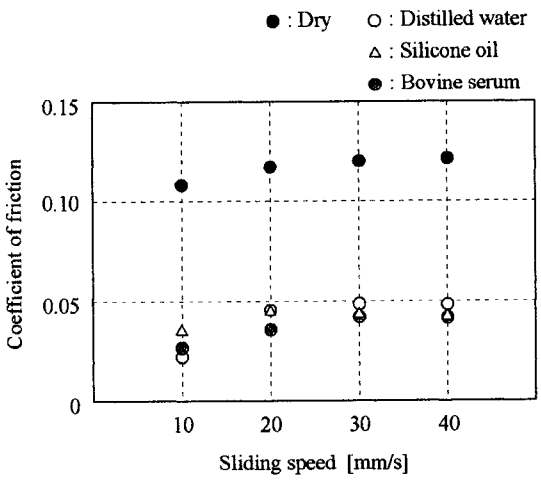
The wear rate of UHMWPE was determined from the weight loss of the pin specimen. At intervals of 10km sliding, the pin specimen was removed from the test apparatus and cleaned by the cleaning procedure described before. The polyethylene specimen was then dried in a vacuum desiccator for 2 hours under 50C and weighed to an accuracy of 0.01mg. A soak control pin was prepared and the measured weight change of the pin specimen was corrected with the weight change of the soak control pin to cancel the weight gain due to a fluid uptake from the lubricant.

Two kinds of ceramics, TZP and alumina ceramics, and one metal, cobalt-chromium alloy (CoCr), were used as plate specimens and the UHMWPE wear behaviour was compared to the results when sliding against HTOZ. The TZP used was the same medical grade material used in the unidirectional sliding test. The alumina ceramics used in this study were manufactured from medical grade high purity (99.9%) alumina. Both TZP and high purity alumina specimens have received a HIP treatment in the same way as current ceramic components for joint prostheses. CoCr specimens were made from medical grade material for the metal joint prostheses. The surfaces of all plate specimens were polished to a surface roughness (measured by stylus profilometer) of 0.01μm (Ra).

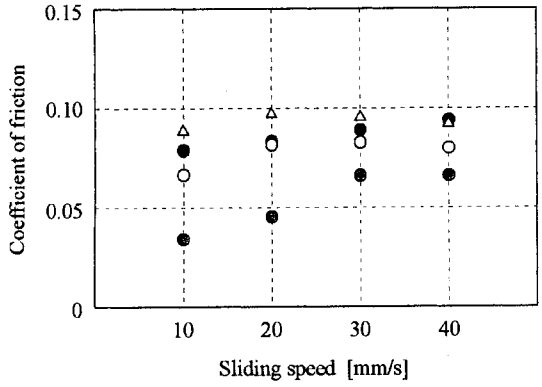
3. RESULTS

Results of the friction measurement in the unidirectional sliding test are shown in Fig.5. The coefficient of friction was measured during 8 minutes for each sliding speed. The data for the initial 3 minutes after changing the sliding speed were ignored to exclude the effect of the initial instability. The average value for the residual 5 minutes of measurements was plotted against the sliding speed.

In the majority of the unidirectional sliding tests, the coefficient of friction increased with increasing sliding speed except for UHMWPE-on-HTOZ and UHMWPE-on-TZP lubricated with silicone oil and UHMWPE-on-TZP with distilled water. In these three cases, the coefficient of friction reached to the



(a) UHMWPE / HTOZ



(b) UHMWPE / TZP

Figure 5. Comparison of coefficient of friction and effects of lubricants.

maximum value at the sliding speed of 20 or 30mm/s and then reduced.

The UHMWPE-on-HTOZ sliding pair showed apparently higher friction at all siding speeds compared with the UHMWPE-on-TZP sliding pair under the un-lubricated condition. However, friction of the UHMWPE-on-HTOZ sliding pair was successfully reduced by lubrication. The coefficient of friction was decreased over 50% under lubricated conditions and became less than 0.05. The three kinds of lubricants used in this study had a similar effect on friction between UHMWPE and HTOZ.

In contrast, there was a difference between the three lubricants in the case of the UHMWPE-on-TZP sliding pair. Silicone oil did not reduce the friction of UHMWPE-on-TZP. Distilled water showed only limited lubrication effects on TZP as the coefficient of friction lowered only slightly in the high sliding speed regime. Only bovine serum successfully reduced the friction between UHMWPE and TZP especially in the low sliding speed regime. However, the friction coefficient remained greater than 0.06 at sliding speeds of 30 and 40mm/s which was exceeded the coefficient of friction of UHMWPE-on-HTOZ under same test conditions.

Results under multidirectional sliding are shown in Figs. 6 and 7. The weight loss of the UHMWPE pin specimens as a function of the sliding distance is shown in Figure 6. The weight loss of the UHMWPE pin specimens increased linearly throughout the study except for one of the pin specimens sliding against the high purity alumina disk, which showed a high initial wear during the first 10km of sliding and then reduced to a lower steady state wear. The measured gravimetric wear was greater than 1mg by the end of the wear test. These values were large enough compared with the accuracy of weight measurement (0.01mg). Therefore, it was reasonable to conclude that the results of the wear tests were quantitatively reliable.

The gravimetric wear of the UHMWPE pin specimens sliding against the CoCr metal disk were

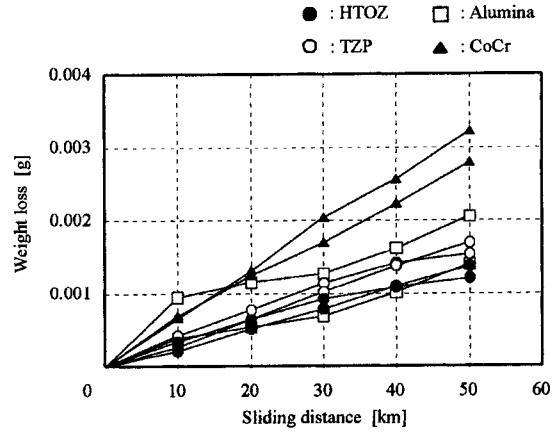


Figure 6. Weight loss of UHMWPE pin specimens as a function of the sliding distance.

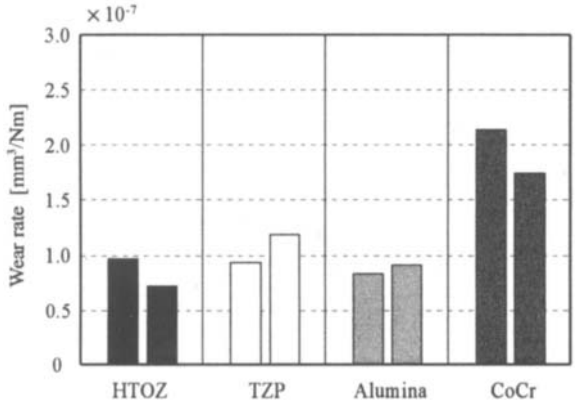


Figure 7. Comparison of steady state wear rate of UHMWPE.

3.23mg and 2.80mg at the sliding distance of 50km. This was greater than the wear of the UHMWPE-on-ceramic sliding pairs. Although one of the UHMWPE-on-alumina sliding pairs revealed relatively high wear (2.06mg), the wear of the other UHMWPE-on-alumina specimen was 1.41mg, which was slightly smaller than the wear of the UHMWPE-on-TZP sliding pairs which were 1.70 and 1.54mg respectively. The wear of the UHMWPE-on-HTOZ sliding pairs was 1.38 and 1.22mg respectively, and these values were the smallest in this wear test.

The comparison of the steady state wear rate, defined as the wear volume per unit load and unit sliding distance, is shown in Figure 7. In this study, the result of the initial 10km of sliding was not used for the calculation of the steady state wear rate to exclude any initially high wear.

The difference in the steady state wear rate of the UHMWPE-on-ceramic sliding pairs was not significant (P=0.56, 0.90 and 0.27 respectively). The wear rates of the UHMWPE-on-TZP sliding pairs were 1.18×10^{-7} mm^3/Nm and 0.93×10^{-7} mm^3/Nm respectively. These wear rates were slightly larger than the steady state wear rates of UHMWPE-on-alumina which were 0.91×10^{-7} and 0.83×10^{-7} mm^3/Nm respectively. One of UHMWPE-on-HTOZ sliding pairs also had a similar wear rate of 0.97×10^{-7} mm^3/Nm. However, the wear rate of the other UHMWPE-on-HTOZ specimen was 0.72×10^{-7} mm^3/Nm which was smaller than that of the other UHMWPE-on-ceramic sliding pairs. The wear rate of the UHMWPE-on-HTOZ specimens was less than half of the wear rate of the UHMWPE-on-CoCr specimens which was 2.13×10^{-7} and 1.74×10^{-7} mm^3/Nm respectively.

4. DISCUSIONS

The tribological characteristics of new prosthetic materials have to be examined, as excessive wear of the joint surface is a major cause of the failure of current joint prostheses. It has been widely recognised that UHMWPE wear debris released from the articulating surface of joint prostheses have an adverse effect on the surrounding tissue and can induce joint loosening [3-5]. Therefore, we evaluated in this study the friction and wear characteristics of the UHMWPE-on-HTOZ sliding pair, which was newly developed by our co-authors, to examine whether it was suitable to use in cementless knee joint prostheses.

The results of this study support the use of UHMWPE-on-HTOZ for use in artificial joint surfaces. The frictional behaviour of UHMWPE-on-HTOZ was less than UHMWPE-on-TZP when lubricated with physiological lubricant (diluted bovine serum). The wear of UHMWPE-on-HTOZ was similar to UHMWPE-on-alumina, which is used in current ceramic knee prostheses. The wear of UHMWPE-on-HTOZ was significantly less than the wear of UHMWPE-on-CoCr ($P<0.05$).

The friction coefficient of UHMWPE-on-HTOZ was less than UHMWPE-on-TZP in all lubricated friction tests. The inverse was true in the un-lubricated friction tests. In this study, three kinds of lubricants with similar viscosity and different boundary lubrication ability were used. Of these, only bovine serum reduced the friction of the UHMWPE-on-TZP sliding pair. Diluted bovine serum has been used widely as a lubricant in the wear testing of prosthetic joint materials, as it is reported to reproduce physiological wear behaviour [13]. It contains many kinds of biological molecules, such as proteins and lipids, and our previous studies indicated that its protein and lipid constituents could easily be adsorbed onto ceramics surfaces and the adsorbed molecules could play a role in boundary

lubrication [10, 14, 15]. Thus, even in this friction test, it can be considered that the friction between UHMWPE and TZP had been reduced by the boundary lubrication of these biological constituents of bovine serum. In contrast, silicon oil, which was very inert and had little boundary lubrication ability, could not lower the friction of UHMWPE-on-TZP at all. This suggests that the viscosity of lubricants used in this study was too small to form a hydrodynamic fluid film between UHMWPE and TZP, and that a boundary lubrication condition has been predominant.

However, the difference between bovine serum and silicone oil became very small with the UHMWPE-on-HTOZ sliding pair. Although the friction between UHMWPE and HTOZ became slightly higher in silicone oil compared with bovine serum at low sliding speeds of 10 and 20mm/s, the difference between these lubricants became almost negligible at higher sliding speeds of 30 and 40mm/s. Both lubricants successfully reduced the friction of UHMWPE-on-HTOZ from 0.12 in the un-lubricated condition to less than 0.05. We postulate that these results indicate that hydrodynamic lubrication was effective for the UHMWPE-on-HTOZ sliding pair especially under the higher sliding speeds when the effect of the boundary lubrication of serum constituents has been drowned out.

Despite the difference in the lubrication condition the steady state wear behaviour between UHMWPE-on-HTOZ and UHMWPE-on-TZP was not significant in the serum lubricated multidirectional sliding wear test ($P=0.56$). However, one of the UHMWPE-on-HTOZ sliding pairs revealed the lowest wear rate in this study. The UHMWPE-on-TZP sliding pairs had a relatively high wear rate compared to the UHMWPE-on-alumina sliding pairs. These results suggest that the difference in the lubrication condition between UHMWPE-on-HTOZ and UHMWPE-on-TZP induced not only the observed difference in the frictional behaviour but also the difference in the long-term polyethylene wear behaviour.

Although the preferable mechanical strength of zirconia ceramics was successfully retained in the TZP substrate of the HTOZ, some characteristics of HTOZ articulating surface had been consequentially altered from the original TZP surface after the manufacturing processes of the coating layers. For example, the articulating surface of HTOZ became hydrophobic compared with the original TZP surface. The hydrophobicity of the articulating surface could be considered to have various effects on the lubrication. Therefore, it may be presumed that the change in the characteristics of the articulating surface is one of the reasons for the difference in the friction and wear behaviours between the UHMWPE-on-HTOZ and UHMWPE-on-TZP sliding pairs.

In this study, atomic force microscopy (AFM) was used to observe the surface of the ceramic specimens in detail. AFM images of HTOZ and TZP are shown in Figure 8. The centreline average roughness (Ra) was also calculated from topographic data obtained by AFM. The calculated Ra values of HTOZ and TZP specimens are compared in Table 2. Both HTOZ and TZP specimens had been polished finely by the same procedure, and the Ra was measured using a stylus profilometer. AFM observation revealed that the microscopic structure of the HTOZ articulating surface was quite different from the TZP surface. The surface of the TZP specimen was conspicuously covered with very fine TZP granules, whose diameter was several hundreds nm or less. This type of granulated microstructure was not observed on the HTOZ surface as it looked very smooth even under AFM observation. Ra values also indicated that the surface of HTOZ was smoother than the TZP surface microscopically. Ra values of the TZP calculated from AFM data were consistent with the measurement using the stylus profilometer. However, the Ra of the HTOZ surface decreased with reducing the scan area of the AFM observation and became only a half of TZP surface roughness when using a 2μm scan.

The structural difference of the articulating surface described above presumably affected the contact condition of the UHMWPE-on-ceramic sliding pairs and also the frictional behaviour under the un-lubrication condition. In this study, friction of the UHMWPE-on-TZP sliding pair was smaller than that of UHMWPE-on-HTOZ in the un-lubricated test. This result suggests that the granulated

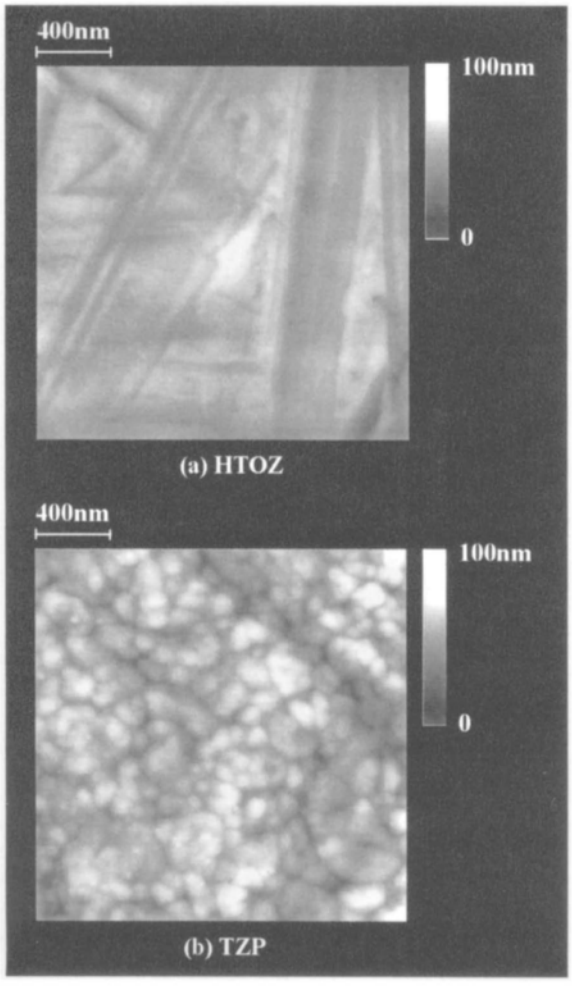


Figure 8. AFM image of HTOZ and TZP surfaces and surface roughness derived from AFM observation.

microstructure of TZP surface may consequently reduce the friction under the un-lubricated condition.

Furthermore, it can be considered that the microstructure of the ceramic surface had an important effect on the hydrodynamic fluid film formation under the lubricated condition. The detailed relationship between the microstructure of the ceramic surface and the fluid film formation is not fully understood. However, the microscopically smooth surface of HTOZ may have contributed to the preferable friction and wear behaviour of the UHMWPE-on-HTOZ sliding pair in the lubricated tests.

Table 2
Comparison of centreline average roughness (Ra, μm) measured by using stylus plofilometer and AFM

	HTOZ	TZP
Stylus plofilometer	0.01	0.01
AFM: 40μm scan	0.010	0.010
AFM: 10μm scan	0.009	0.011
AFM: 2μm scan	0.004	0.008

5. CONCLUSIONS

The tribological characteristics of a newly developed zirconia composite material (HTOZ) for cementless knee prostheses, was examined using unidirectional pin-on-disk and multidirectional pin-on-plate tests.

The results of these experiments indicated that the friction and wear characteristics of the UHMWPE-on-HTOZ sliding pair were potentially superior to that of the conventional UHMWPE-on-TZP sliding pair. The friction of UHMWPE-on-HTOZ was lower than that of the UHMWPE-on-TZP sliding pair under physiological conditions with a biological lubricant. The wear of polyethylene against HTOZ was, in addition, similar to other UHMWPE-on-ceramic sliding pairs tested under the same physiological multidirectional sliding motion. The relatively hydrophobic and microscopically smooth surface of HTOZ may have contributed to its preferable friction and wear behaviour, as its surface made it easier for fluid films to develop compared to the conventional TZP.

ACKNOWLEDGEMENT

The authors would like to thank Mr. M. Inoue, Kyushu University. Surface observations of ceramic specimens in this study were conducted using atomic

force microscopy at the Centre of Advanced Instrumental Analysis, Kyushu University. This study was financially supported by the Grant-in Aid for Encouragement of Young Scientists of the Ministry of Education, Science, Sports and Culture of Japan.

REFERENCES

1. D. Dowson, Instn Mech Engrs, Part H, 215 (2001) 335.
2. J.G. Lancaster, D. Dowson, G.H. Isaac and J. Fisher, Proc. Instn Mech Engrs, Part H, 211 (1997) 17.
3. H.G. Willert and M. Semlitsch, J. Biomed. Res., 11 (1977) 157.
4. P.A. Revel, B. Weightman, M.A.R. Freeman and B. Vernon Roberts, Arch. Orth. Traum. Surg., 91 (1978) 167.
5. E. Ingham and J. Fisher, Proc. Instn Mech Engrs Part H, 214 (2000) 21.
6. F.S. Haddad, N.J. Levell, P.M. Dowd, A.G. Cobb and G. Bentley, J Bone Joint Surg, 77B (1995), 329.
7. I. Noda, S. Masuda, H. Kitano, J. Ikeda, Y. Yoshihara, K. Mukai and T. Shimotoso, Bioceramics, 14 (2001), 577.
8. R.G. Geesink, Clin. Orthop., 395 (2002), 53.
9. Y. Sawae, T. Murakami and J. Chen, Wear, 216 (1998) 213.
10. Y. Sawae and T. Murakami, Proc. International Tribology Conference Nagasaki, 2000, (2001) 1521.
11. C.R. Bragdon, D.O. O'Connor, J.D. Lowenstein, M. Jasty and W.D. Syniuta, Proc Instn Mech Engrs Part H, 210 (1996) 157.
12. P.S.M. Barbour, M.H. Stone and J. Fisher, Proc Instn Mech Engrs Part H, 213 (1999) 455.
13. A. Wang, A. Essner, V.K. Polineni, D.C. Sun, C. Stark and J.H. Dumbleton, New Directions in Tribology, (1997) 443
14. T. Murakami, H. Higaki and S. Doi, Bioceramics, 9 (1996) 109.
15. T. Murakami, Y. Hayakawa, H. Higaki and Y. Sawae, Proc JSME International Conference on New Frontiers in Biomechanical Engineering, (1997), 233.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

The relationship between steady-state wear rate and theoretical film thickness in metal-on-metal total replacement hip joints

D. Dowson

Institute of Tribology, School of Mechanical Engineering,
The University of Leeds, Leeds, LS2 9JT

Metal-on-Metal total replacement hip joints are thought to operate primarily in the boundary, or mixed lubrication regimes, but recent studies[1] have suggested that fluid-film lubrication alleviates the severity of contact and hence wear under some conditions. It is shown that the iso-viscous / elastic mode of fluid-film lubrication applies in such implants. Measurements of steady-state wear rates for different prostheses, in different simulators and under different operating conditions reported from independent tests in laboratories in the U.K., Canada and the U.S.A. are analysed and related to the theoretical elasto-hydrodynamic film thickness. The results exhibit reasonable consistency, even though the test conditions were quite different. A transition from boundary lubrication through mixed lubrication, representing a change in wear rate of some two orders of magnitude, is evident in the wear results as the film thickness increases. At film thicknesses beyond about (15 nm –20 nm) the wear rate varies little as the film thickness increases over a large range. It is shown that very low wear rates of (0.05 $mm^3$to 0.10 mm^3) / 10^6 cycles are achievable beyond this film thickness range through careful tribological design. It would take 9-18 years to generate one pin-head of metallic wear under these conditions.

1. INTRODUCTION

The mechanism of lubrication in total replacement hip joints has been a subject of considerable interest and debate for many years [2].

Four main approaches have been adopted to ascertain the dominant mode of lubrication.

(a) *Measurement of Friction in Simulators;-*
- value of friction coefficient (or factor).
- Stribeck plot.

(b) *Calculation of Fluid-Film Thickness for smooth surfaces;-*
- comparison with composite surface roughness; lambda ratio.

(c) *Detection of voltage drop and surface separation in Joint Simulators;-*
- indicates presence or absence of lubricating film (fluid-film or boundary).

(d) *Wear Measurements in Simulators;-*
- relationship between wear rate and principal design features (eg. diameter, clearance).

However, as with any lubricated, dynamically loaded, tribological components, the lubrication mechanism might well change throughout each cycle of operation and the search for a single mode of operation might be fruitless or even misleading.

The long established metal-on-polyethylene implant, represented mainly by the 7/8" (22.225mm) diameter 'gold standard' Charnley joint, operates predominantly in the boundary, or severe mixed, lubrication regime, such that substantial direct interaction occurs between the surfaces of the opposing femoral head and acetabular cup. Indeed, the implant was designed to minimise both the friction at the interface between the femoral head and acetabular cup and frictional torque on the cup consistent with maximum wear life. The former was achieved through the use of a 'self-lubricating' polymeric material, initially PTFE but from 1962 UHMWPE, rather than a metallic cup rubbing against a metallic femoral head. The second objective was achieved by selecting a nominal femoral head diameter of

22mm. The essential features responsible for the outstanding success of the Charnley joint were thus tribological, through the use of a low friction pair of sliding materials and the optimisation of head diameter to balance the requirements for minimum torque and maximum wear life. Both features were soundly based upon the concepts of boundary lubrication and a near constant coefficient of friction throughout each cycle of operation.

Much has been done [3,4,5,6] to reveal the mode of lubrication in such joints through the measurement of friction in joint simulators of increasing sophistication. Each simulator was capable of measuring the friction to high accuracy. The recorded values of friction were generally similar to those associated with mixed or boundary lubrication. When plotted in the form of a Stribeck curve, the friction factors revealed variations consistent with the expected characteristics for mixed lubrication, with suggestions of either boundary or fluid-film lubrication when relatively low or high viscosity lubricants were used respectively.

The minimum film thickness in the natural hip joint has been estimated to be in the range 0.5 μm - 1.5 μm [7]; a range that did little to encourage the view that fluid-film lubrication could play a significant role in determining the excellent tribological characteristics of such bearings, since the roughness of cartilage was known to be 2 μm - 5 μm. However, it was later demonstrated [8] that micro-elastohydrodynamic action could effectively iron out the asperities passing through the loaded conjunction to leave a film of about 0.5 μm separating the articulating surfaces. This restored confidence in the concept that fluid-film lubrication contributed to the performance of natural joints

In due course a full numerical solution to the elastohydrodynamic problem for Charnley type, metal-on-polymer joints was reported [9]. It was confirmed that the theoretical film thickness in natural joints was about 1 μm , while films of only about 0.2 μm were predicted for the 'gold-standard' form of joint replacement The latter films were too small to ensure separation of the metallic femoral heads and polyethylene acetabular cups and it was concluded that boundary, or at the best mixed lubrication, would prevail. This theoretical finding was entirely consistent with the experimental findings noted earlier [3,4,5,6].

The same boundary and mixed lubrication regimes are found in small diameter *'hard-on-hard'* metallic implants [10,11,12]. However, beyond a certain diameter, or below a certain clearance, the severity of mixed lubrication diminishes, as hydrodynamic action becomes increasingly significant [1]. The severity of surface interaction and wear can thus be controlled through careful design and manufacture to optimise fluid film lubrication effects.

The extent to which fluid-film action can relieve the severity of asperity loading should be related to the ratio of the theoretical fluid-film thickness between smooth surfaces to the composite surface roughness, known as the lambda ratio (Λ). For a specified form of Total Hip Replacement and a given lubricant, the film thickness is determined by the diameter of the femoral head and the clearance between the head and the acetabular cup alone. These are the dominant design features of hard-on-hard Total Hip Replacements.

However, the measurement of surface features on femoral heads and acetabular cups throughout simulator tests is a difficult, expensive and time absorbing exercise. This is particularly true for the cups. The composite surface roughness in the steady state wear regime achieved after running-in is not always recorded, thus making it impossible to calculate the appropriate lambda ratios. It is nevertheless interesting and informative to relate the reported steady-state wear rates for metal-on-metal total hip replacements to film thickness, rather than the more meaningful lambda ratio. The results of an initial study of metal-on-metal total hip replacements of various materials, forms, diameters and clearances, tested in different joint simulators in various laboratories is reported in the present paper.

2. THEORETICAL BACKGROUND

Ideally, the wear rate should be related to the ratio of film thickness to some effective composite surface roughness, but as noted earlier, the latter is

rarely recorded in the literature following the establishment of steady conditions. Furthermore, it has been noted that the correlation between steady-state wear rate and the lambda ratio might be poor if the mode of lubrication changes due to different geometrical features of the prostheses, such as diameter and clearance [12]. It is significant, however, that reasonable correlation between total volumetric wear and the lambda ratio has been noted in extensive simulator studies of metal-on-metal prostheses [13].

Further information on the nature and quantification of worn surfaces in metal-on-metal prostheses is awaited. In the present study the relationship between steady-state wear rate and theoretical minimum film thickness for smooth surfaces will be explored. This will be achieved by plotting the steady-state wear rates recorded in several laboratories against the calculated minimum film thickness.

The minimum film thicknesses in metal-on-metal total hip replacements are of the order of nano-metres or a few tens of nano-metres (10^{-9} to 10^{-8} m). A simple, Hertzian, calculation of the local elastic deformations shows that they are of the order of a few microns (10^{-6} m). Since the elastic deformations are typically (10^{2} to 10^{3}) times greater than the film thicknesses, the mode of lubrication between smooth surfaces is clearly elastohydrodynamic.

The viscosity of fluids like water and synovial fluid is known to be little affected by pressure [14]. The same observation is likely to apply to the serum solutions used in many joint simulators in laboratory testing. The lubricant in total hip replacements, in vivo and in simulators, can thus be treated as an iso-viscous fluid.

These two observations on the magnitude of the elastic deformations relative to the theoretical film thicknesses and the iso-viscous nature of the lubricants indicate that metal-on-metal joints can theoretically operate in the iso-viscous/elastic regime of elasto-hydrodynamic lubrication.

The essential theoretical requirement in the present study was therefore to be able to estimate the appropriate elastohydrodynamic minimum film thickness in total replacement metal-on-metal hip joints operating in the iso-viscous/ elastic regime of lubrication.

2.1 The minimum film thickness.

The generation of a full solution for film thickness in a dynamically loaded spherical bearing, like a metal-on-metal total hip replacement, can be a time consuming process [9,15]. However, it has been demonstrated that the film thickness throughout the cycle in synovial joints does not vary as much as the dynamic conditions of load and entraining velocity might suggest. Furthermore, a fair estimate of the mean minimum film thickness throughout the cycle can be achieved by adopting average values for the principal external variables in the well established, empirical elastohydrodynamic lubrication equations for the iso-viscous/elastic regime. This fortunate finding is, of course, a consequence of the impressive ability of squeeze-film action to maintain hydrodynamic films, once they have been established.

The empirical equation for the minimum film thickness in elastohydodynamic conjunctions operating in the elastic-isoviscous regime can be written as;

$$H_{min} = \frac{h_{min}}{R_X} = 2.80 \left(\frac{\eta u}{E' R_X} \right)^{0.65} \left(\frac{w}{E' R_X^{\,2}} \right)^{-0.21} \qquad (1)$$

where,

$$u = \frac{\left(u_c + u_h \right)}{2} = \frac{\left((d/2\Omega) + 0 \right)}{2} = \frac{d\Omega}{4} \qquad (2)$$

and, for a spherical femoral head of diameter (d) in an acetabular cup with a diametral clearance (c_d)

$$R_X = \frac{d}{2} \left(\frac{d}{c_d} + 1 \right) \qquad (3)$$

If representative values are adopted for the mean lubricant viscosity (η), load (w) and angular velocity (Ω), the above equations can be used to calculate the minimum elastohydrodynamic film

thickness between smooth surfaces for total joint replacements of any specified diameter and clearance.

The following representative values will be adopted for the present analysis.

lubricant viscosity (η)	0.0009 Pa.s
load (w)	2,500 N
angular velocity (Ω)	1.5 rad/s

Equation (1) then reduces to;-

$$h_{min} = 1.759\,(d)^{1.42}\left(\frac{d}{c_d}+1\right)^{0.77}(E')^{-0.44}\,x10^{-3} \qquad (4)$$

For CoCrMo materials representative values of Young's modulus and Poisson's ratio are 230 GPa and 0.3 respectively, enabling the film thickness equation to be further reduced to

$$h_{min} = 0.929\,(d)^{1.42}\left(\frac{1000\,d}{c_d}+1\right)^{0.77} x10^{-3}\ nm \qquad (5)$$

- where (d) is measured in (mm) and (c_d) in (μm).

2.2 Steady-State Wear Rate.

A general representation of the penetration trace for a metal-on-metal total joint replacement is shown in Figure 1.

Initially a relatively high wear rate is experienced in a 'running-in' or 'bedding-in' period. During this period the wear rate gradually decreases as the surfaces accommodate each other, until a relatively low, steady-state wear rate is achieved. In clinical studies the penetration rate, defined as (p_2 / T), is usually measured from sequential x-ray films. It should be noted that this differs from the steady state penetration rate of ((p_2-p_1) / T).

In joint simulator testing it is usual to plot wear (mm^3) against the number of test cycles (10^6), and the curve is similar in form to that displayed in

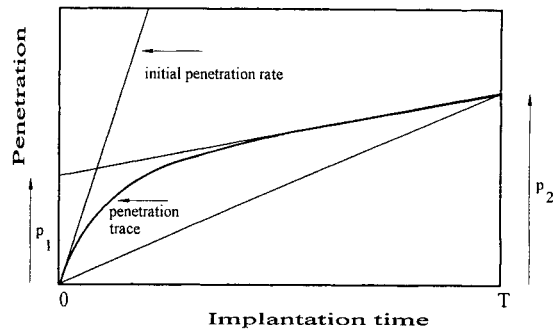


Figure 1. Penetration Trace for Metal-on-Metal Total Replacement Hip Joint.

Figure 1. In such work, the steady- state wear rates are usually expressed in units of ($mm^3/10^6$) cycles

3. RESULTS

Test results for steady-state wear rates from five different laboratories in the U.K., Canada and the U.S.A have been collected from the published literature. Some 70 separate data points were collected, representing different prostheses, simulators and lubricants. The corresponding theoretical minimum film thicknesses were calculated from equation (1), taking these variations into account. The mean loads representative of test conditions in the several laboratories ranged from 891N to 2500N, but it can be seen from equation (1) that changes in load exert only a modest influence upon film thickness. The entraining velocities, which are more influential, varied from about 0.011m/s to 0.021m/s. Serum in various dilutions was used as the lubricant in all the tests and estimates of the viscosity at high shear rates ranged from 0.0009 Pas to 0.0015 Pas.

It was, of course, possible to include only those results for which the steady-state wear data and full details of the metal-on-metal joint diameter and clearance together with the simulator operating conditions were available.

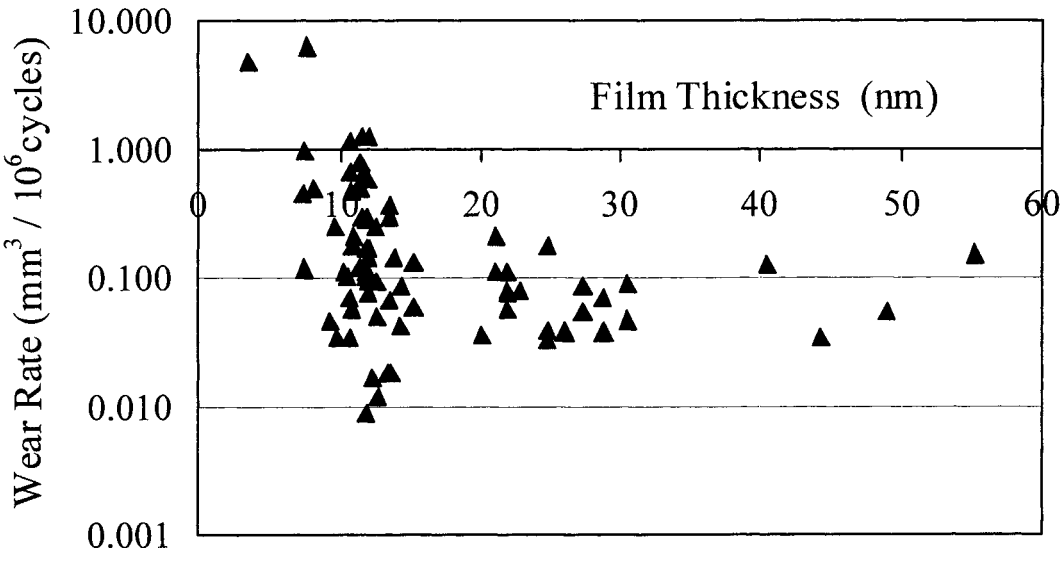


Figure 2 Steady-State Wear Rates v Theoretical Elastohydrodynamic Film Thickness

The relationship between steady-state wear rate for ideally smooth surfaces is shown in Figure 2.

It is clear from this Figure that most metal-on-metal total replacement hip joints operate with mean theoretical film thicknesses of the order of nanometres or a few tens of nanometres. Furthermore, there is a fair correlation between wear rate and film thickness, irrespective of the type of prosthesis and the form of simulator and test conditions selected.

There is, of course, considerable scatter in the recorded wear rates, particularly at low values of the theoretical film thickness, where mixed or even boundary lubrication prevails. In this region the recorded wear rates decrease by about two orders of magnitude as the film thickness increases from about 5 nm to 15 nm. Thereafter, the wear rates decrease only slightly as the film thickness increases to several tens of nanometres. It should be recalled that the two distinct regions are linked almost entirely to changes in femoral head diameter or clearance.

In the low wear region the wear rates are generally of order (0.05 mm^3 to 0.1 mm^3) / 10^6 cycles. This represents the limiting range of wear rates that can reasonably be achieved in metal-on-metal total replacement hip joints. The form of Figure 2 strongly suggests that careful tribological design and sound modern manufacturing techniques can lead to these remarkably low rates of wear. If the theoretical film thickness is less than about 15 nm, such that relatively severe wear takes place, modest changes in film thickness can result in substantial changes in wear. The influence of tribology, directly related to film thickness, is likely to mask changes associated with differences in mechanical design and even the selection of materials.

4. DISCUSSION

It has been demonstrated elsewhere [10,13] that the mode of lubrication (boundary; mixed; fluid-film) plays a major role in determining the steady-state wear rate in metal-on-metal total replacement hip joints.

Direct evidence of lubricating film formation in metal-on-metal joints tested in a hip joint simulator has been presented earlier [1]. It is now evident that the estimated film thickness has a major effect upon steady-state wear rate for a wide range of conditions in such joints. As the film thickness increases from about 5 nm to 15 nm the steady-state wear rate decreases by some two orders of magnitude. This correlation between steady-state wear rate and film thickness is initially somewhat surprising, since it might reasonably be expected that the strongest correlation would be with the lambda ratio if the phenomenon is attributable directly to asperity interactions. However, the latter point could not be investigated directly from the test results available from the published literature, since surface roughness in the steady-state wear region was rarely reported. It is, of course possible that modern production techniques ensure that, after bedding-in, the surface form and roughness of the various implants evaluated became similar, irrespective of material, design, diameter and clearance. Although unlikely, this would explain why reasonable correlation was evident between steady-state wear rate and film thickness.

The rapid decline in steady-state wear rate as the theoretical elastohydrodynamic film thickness increases to about 15 nm is followed by a low rate of change over a wide range of film thicknesses. From the point of view of steady-state wear, there appears to be little to be gained through the use of exceedingly large diameter implants, or through the maintenance of exceedingly tight clearances once the theoretical minimum film thickness has exceeded, say, 20 nm.

Another remarkable feature of the relationship evident in Figure 2 is that the wear rate is exceedingly low, typically (0.05 mm^3 to 0.10 mm^3/ 10^6 cycles), for theoretical film thicknesses in excess of about (15 nm-20 nm). This at once suggests that a good tribological design guide for the minimisation of steady-state wear in metal-on-metal implants is to ensure that the theoretical film thickness, calculated on the basis of the procedure outlined in Section 2, equals or exceeds this range.

In order to appreciate the exceedingly small wear rates involved in the lifetime of metal-on-metal prostheses designed to operate with theoretical film thicknesses in excess of (15 nm-20 nm), the volumetric wear can be related to the size

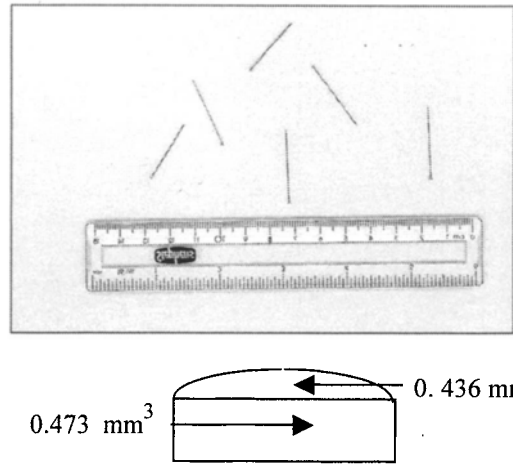


Figure 3 Volume of a Pin-Head (0.909 mm^3)

of a pin-head. The random selection of pins, of varying lengths, shown in Figure 3, was taken from a co-operative embroiderer and patchworker's box and measured on a shadowgraph.

The pin heads were found to be in the form of short circular cylinders with domed heads, as shown in Figure 3. The mean volume of the pin-heads was found to be about 0.9 mm^3.

It is well known that subjects take about one million steps each year and this means that for a wear rate of (0.1 mm^3 / 10^6 cycles), it would take 9 years to produce one pin-head of wear debris. This duration would be doubled to 18 years if the lower wear rate of (0.05 mm^3 / 10^6 cycles) applied.

These exceptionally long durations required to yield one pin-head of metallic wear debris would, of course, be shortened if frequent stops and starts, or unusual activities, resulted in repeated running-in wear.

Representative wear rates for a range of different material pairs and diameters of total replacement hip joints suggest that the periods required to produce one pin head of wear are as shown in Table 1.

Further data will be added to develop the relationship between steady-state wear rate and calculated elastohydrodynamic film thickness illustrated in Figure 2 as results become available. However, the trend portrayed by the results recorded from some 70 individual tests from five centres using different simulators in three different

Table 1.
Implantation Period to Yield One Pin Head of Wear

Prosthesis Type	Dia. (mm)	Wear Rate	Implantation Period
Metal-&-Polymer*[2]	22	(0.1-0.2) mm/year	4-8 days
Ceramic**&-Polymer*[2]	22	(0.05-0.1) mm/year	8-17 days
Metal-on-Metal[10]	22	6.3 mm^3/yr.	52 days
Metal-on-Metal[10]	28	0.54 mm^3/yr.	1.7 years
Metal-on-Metal[10]	36	0.07 mm^3/yr	18 years

* UHMWPE **Alumina

countries provides some confidence in the overall conclusions summarized below.

5 CONCLUSIONS

The 70 data points collected together for steady-state wear rates for different designs of metal-on-metal prostheses tested in different joint simulators located in five laboratories in three countries enable the following conclusions to be drawn.

1 The reduction in severity of asperity contact achieved by effective lubrication in metal-on-metal prostheses has a major effect upon the steady-state wear rate achieved in such joints.

2 A clear and dramatic decrease of some two orders of magnitude in steady-state wear rate is evident in metal-on-metal prostheses as the theoretical elastohydrodynamic film thickness increases from about 5 nm to 15 nm.

3 Some relatively small diameter joints, which generated only modest film thicknesses, less than about 7 nm, operated in the boundary lubrication regime, whilst some of relatively large diameter and small clearance enjoyed the benefits of fluid-film action. However, the majority of the implants tested operated in the mixed lubrication regime, where the steady-state wear rate varied substantially.

4 Only modest further decreases in steady-state wear rate occured once the theoretical film thickness had attained values of film thickness in excess of 15 nm –20 nm. The wear rates at film thicknesses above this range are exceedingly low, being typically (0.05 –0.10) $mm^3/10^6$ cycles

5 The very rapid decrease in steady-state wear rates as the severity of asperity contact is relieved by more effective lubrication in the mixed lubrication regime illustrates the importance of correct tribological simulation in laboratory evaluations of joint performance. Changes in implant wear rates, which may previously have been attributed to the use of different material combinations, might well have been influenced primarily by the role of the elastohydrodynamic film thickness in determining the severity of contact.

6 The data suggest that a good design guide is to ensure that the theoretical elastohydrodynamic film thickness achieves, or preferably exceeds, the range of 15 nm – 20 nm.

REFERENCES

1. Dowson D, McNie C M, Goldsmith A A J (2000) Direct experimental evidence of lubrication in a metal-on-metal total hip replacement tested in a joint simulator, Proc. Instn. Mech. Engrs, Part C, 214, 75-86.
2. Dowson D (2000) New joints for the millenium: wear control in total replacement joints. (The James Clayton Memorial Lecture 2000), Proc. IMechE, Part 'H', Engineering in Medicine, Volume 215, No.H4, pp 335-358.
3. Unsworth, A, Dowson, D, Wright, V and Koshal, D. (1975), The frictional behaviour of human synovial joints - Part II. artificial joints, Trans ASME, Journal of Lubrication Technology, , Vol. 97, Series F, No. 3, pp 377-382.
4. Unsworth, A., Hall, R.M., Burgess, I.C., Wroblewski, B.M., Streicher, R. M. and

Semlitsch, M., (1994), Frictional resistance of new and explanted artificial hip joints, Austrib, '94, Perth, Western Australia, Conference Proceedings, 383-390.

5. Hall, R.M., Unsworth, A, Wroblewski, B.M. and Burgess, I.C., (1994), Frictional characteristics of 5 explanted Charnley hip prostheses, WEAR , 175, 159-166.
6. Scholes, S.C. and Unsworth, A., ,Comparison of friction and lubrication of different hip prostheses, *Proc.* Instn. Mech.Engrs., Part 'H', Engineering in Medicine, , Vol 214, No. H1, 49-57.
7. Dowson, D, Unsworth, A, Cooke, A F and Gvozdanovic, D., (1981), Lubrication of joints, Chapter 13, pp 120-145, in 'An Introduction to the Biomechanics of Joints and Joint Replacement', Edited by Dowson, D and Wright, V., Mechanical Engineering Publications, Bury St. Edmunds.

8 Dowson, D and Jin, Zhong-Min, (1986), 'Micro-elastohydrodynamic lubrication of synovial joints', Engineering in Medicine, Volume 15, Number 2, pp 63-65.

9 Jalali-Vahid, D., Jagati, M., Jin, Z-M and Dowson, D., (2001), 'Prediction of lubricating film thickness in a ball-in-socket model with a soft lining representing human natural and artificial hip joints', Proceedings of the Institution of Mechanical Engineers, Part 'J', Engineering Tribology, Volume 215, No. J4, pp 363-372.

10 Smith, S. L., Dowson, D. and Goldsmith, A. A. J. (2001), ' The lubrication of metal-on-metal total hip joints: A slide down the Stribeck curve', , Proceedings of the Institution of Mechanical Engineers, Part 'J', Engineering Tribology, Volume 215, No. J5, pp 483-484.

11 Jin, Z.M., Dowson,D. and Fisher, J., (1997), ' Analysis of Fluid Film Lubrication in Artificial Hip Joint Replacements with Surfaces of High Elastic Modulus', Proceedings of the Institution of Mechanical Engineers, Part 'H', (ISSN-0954-4119), Volume 211, No. H3, 247-256

12 Chan, F W, Bobyn, J D, Medley, J B, Krygier, J J and Tanzer,M, (1999), Wear and lubrication of metal-on-metal hip implants, Clinical Orthopaedics and Related Research, Number 369, pp10-24.

13 Dowson, D, Goldsmith, A A J, McNie, C M and Smith, S L, (2003), A tribological study of metal-on-metal total replacement hip joints, Ch. 4, pp 41-58, in 'Friction, Lubrication and Wear of Artificial Joints, Ed. I M Hutchings, published by Professional Engineering Publishing Limited, Bury St. Edmunds, Suffolk, U.K.

14 Cooke, A F, Dowson, D and Wright, V, (1978), The pressure-viscosity characteristics of synovial fluid, Bio-Rheology, Vol. 15, pp 129-135.

15 Jagatia, M and Jin, Z M, (2001), Elastohydrodynamic lubrication of metal-on-metal hip prostheses under steady-state entraining motion, Proceedings of the Institution of Mechanical Engineers, Part 'H', Engineering in Medicine, Volume 215, pp 531-541.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

On the role of the lambda parameter in simulator wear of metal-on-metal hip implants

P. Paré[a], J.B. Medley[a], F.W. Chan[b], S.K. Young[b]

[a] Department of Mechanical Engineering, University of Waterloo
200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

[b] DePuy Orthopaedics Inc., a Johnson & Johnson company
700 Orthopaedic Drive, Warsaw, Indiana 46581, United States of America

The purpose of the present study was to quantify the change in the lambda parameter (λ) during continuous orbital bearing simulator testing of metal-on-metal hip implants and relate λ to wear. The topographies of eight implants were measured periodically throughout four million cycles (Mc) of testing. A typical initial run-in wear period was followed by lower wear rates. The final values of volumetric wear were low (< 0.9 mm^3). The values of λ stayed nearly constant for all implants throughout testing. However, the micro-geometry of the implant surfaces was subject to changes that were not part of the λ assessment. Superficial scratches of a depth less than about 0.10 μm and some deeper scratches (e.g. 65 μm width by 5 μm depth) were observed mainly on the cups after the first 1 Mc and persisted throughout testing. Although λ might indicate the transition to continuous fluid film protection for surfaces with a Gaussian distribution of surface heights, it may not be as effective in interpreting the extent of surface protection for non-Gaussian surfaces under mixed film conditions. Further work on the evolution of wear in metal-on-metal hip implants is recommended.

NOTATION

E'	effective elastic modulus [Pa]
F_{AVG}	average load [N]
h_C	central film thickness [m]
R	reduced or effective radius [m]
R_H, R_C	radius of head, cup [m]
Rq	root mean square (RMS) surface roughness of a trace [m]
u	entrainment velocity [m/s]
ε	time difference between the simulator motion and applied load period [s]
η	viscosity [Pa·s]
λ	lambda
ν	Poisson's ratio
σ	combined surface roughness [m]
σ_H, σ_C	average Rq of traces on head, cup [m]

1. INTRODUCTION

Wear is one of the major concerns in modern total hip arthroplasty (THA). It is generally accepted that polyethylene wear particles, released from the artificial joint, can cause peri-prosthetic osteolysis leading to implant failure by loosening. Since the long term success of THA depends, in part, on low volumetric wear, renewed interest in alternative bearing technologies has resulted in a new generation of metal-on-metal (MM) hip implants [1].

These MM implants had low volumetric wear in some simulator studies [2, 3, 4] and somewhat higher wear in others [5, 6, 7, 8]. *In vivo* volumetric wear (Fig. 1) was estimated from retrieval studies [9, 10] by Medley et al [11] with a conversion from the reported linear wear to volumetric wear using a formula. This formula was based on a curve fit of analytical results from three studies [12, 13, 14] and related volumetric wear (V) in mm^3 to linear wear (L) in μm as follows:

$$V = 0.0283\, L^{1.69} \qquad (1)$$

The *in vivo* volumetric wear from retrieval studies was very scattered (Fig. 1). However, for one year approximately equal to two million cycles (Mc) [15], the wear from simulator studies [2, 3, 4, 5, 6, 7, 8] fell within the range of the in vivo data [9,10].

The results of Rieker et al [10] suggested that a

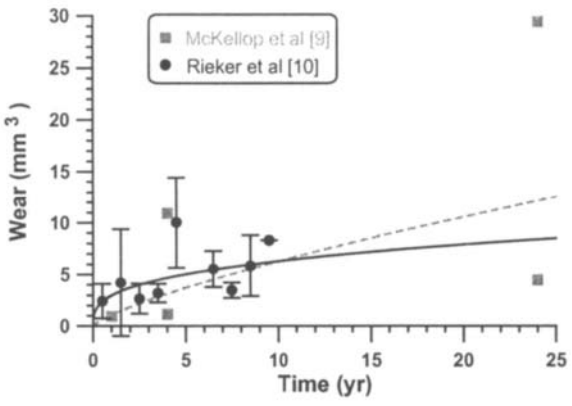


Fig. 1: Prediction of clinical volumetric wear (scatter bars equal one standard deviation).

higher initial wear rate decreased to a somewhat lower "steady state" wear rate after about 0.5 – 1 years *in vivo* (Fig. 1). Similar behaviour also occurred in some simulator studies [2, 3, 4] but tended to be slightly less distinct when simulator testing was performed with the cup on top in the anatomical orientation [5, 6, 7].

Self-polishing of the metal surfaces by a mild run-in wear process might explain the decrease in the initial wear rate [16] and elastohydrodynamic lubrication (EHL) might be intimately involved in the process [17]. The run-in wear may contribute to the development of the steady state wear rate. The actual quantitative changes in surface geometry and their dependence on EHL may be very important.

As noted in previous research [14, 17, 18], EHL can be quantified by a theoretical central film thickness (h_C) and the extent of surface protection can be represented by lambda (λ), the ratio of h_C to the initial measured combined surface roughness (σ) as shown below.

$$\lambda = \frac{h_C}{\sigma} \qquad (2)$$

where $$\sigma = \sqrt{\sigma_H{}^2 + \sigma_C{}^2} \qquad (3)$$

and $$h_C = 5.083 \frac{(\eta u)^{0.660} R^{0.767}}{(E')^{0.447} F^{0.213}} \qquad (4)$$

By using the initial σ, a useful prediction of long-term wear performance can be made prior to THA with a λ assessment. Continuous films are considered to provide good long-term surface protection provided $\lambda > 3$ (Fig. 2) and mixed film lubrication, in which the load is shared between the fluid film and asperity contacts, can also contribute to surface protection in the range of about $1 < \lambda < 3$ [14]. However, if σ changes after implantation as wear proceeds, and this σ is used in Eqn (2), a new "operating" λ value may provide insight into the evolution of surface protection and thus the long-term effectiveness of EHL in THA.

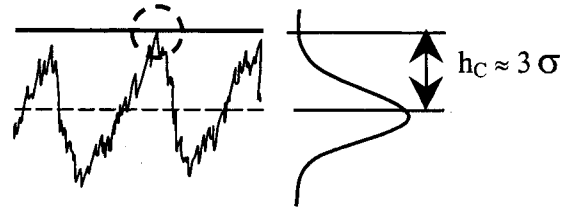


Fig. 2: The fundamentals of the λ assessment. (The contact of rough curved surfaces is represented by a perfectly smooth plane in contact with a rough surface having a parallel mean plane and a Gaussian distribution of surface heights, characterized by σ. Asperity tip contact begins when $\lambda \approx 3$ or $h_C \approx 3\sigma$.)

For example, Chan et al [2] correlated λ (based on the initial σ) with simulator wear after 3 Mc and found a general trend of lower wear for higher λ but some individual implants did not support this trend (Fig. 3). The findings of Firkins et al [6] were included in Fig. 3 and demonstrated some low wear volumes. However, with their small range of λ values, a distinct relationship between λ and wear was not apparent. In both studies, monitoring the evolution of λ might provide some insight into wear.

While high λ values seem optimal to minimize wear, many commercial implants, such as those of Firkins et al [6], seem to operate within a mixed film lubrication regime. If lubrication protection improves as wear proceeds, it may not be essential to manufacture implants to produce high initial λ.

Thus, it may be very useful to relate the wear mechanics of the bearing surfaces in the important early run-in phase and the subsequent lower wear steady state phase to the real-time surface topography, in addition to the initial λ. The amount of surface protection from EHL probably changes as wear proceeds. The purpose of the present study was to explore such changes by monitoring λ as it

evolved during simulator testing and attempt to relate this operating λ to wear.

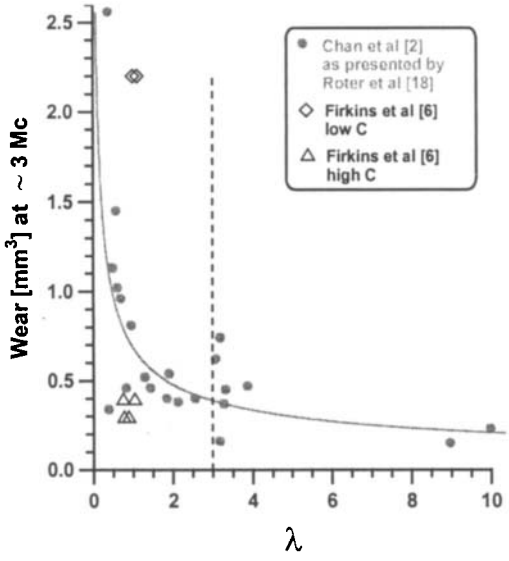


Fig. 3: Simulator volumetric wear vs λ for 28 mm diameter MM hip implants.

2. MATERIALS AND METHODS

Eight MM implants (Table 1) were tested in an orbital bearing hip simulator (Fig. 4) for a total of 4 Mc. All components were manufactured from medical grade CoCrMo alloy (ASTM F1537) and had high carbon (HC) content, except for the Ultima™ heads that had low carbon (LC) content.

Table 1. Implant description and initial geometry

No.	Type	Dia. (mm)	Dia. Clear. (μm)	σ (nm)
1	Ultamet™	28	59.69	19.69
2	Ultamet™	28	60.19	22.06
3	Ultamet™	28	59.94	33.38
4	Ultamet™	36	115.57	16.54
5	Ultamet™	36	117.09	18.01
6	Ultamet™	36	116.84	30.15
7	Ultima™	28	65.80	21.16
8	Ultima™	28	66.96	40.19

The implants were tested using a Paul-type load cycle [19] at a frequency of 1.14 Hz with a maximum load of 2100 N. While the implants were tested in a non-anatomic orientation, Medley et al [20] had shown that the kinematics of the Matco simulator reproduced an in vivo crossing path motion. Following Chan et al [2], testing was performed in a 95 % bovine serum solution that was changed every 0.25 Mc while the implants were cleaned.

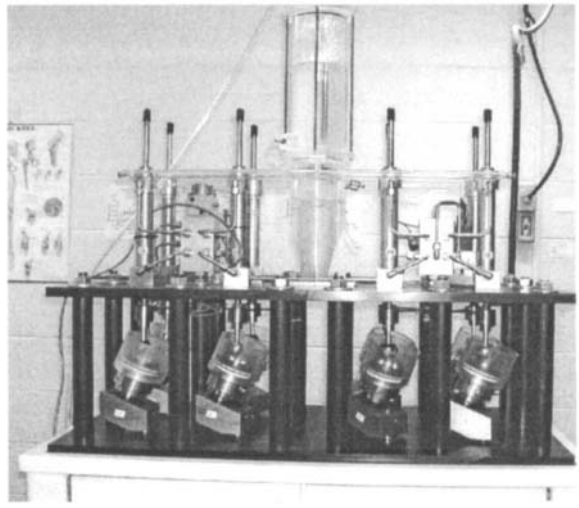


Fig. 4: Orbital bearing simulator (EW08 MMED, Matco, La Canada, CA, USA).

To produce an isotropic wear scar, the synchronization between the load cycle and the simulator motion was initially disengaged so that the load cycle gradually shifted its position over the cup surface (Fig. 5). The time difference (ε) between the periods of the motion and load cycles was 9.6 ms/cycle for the first 1 Mc followed by 6.3 ms/cycle until 3 Mc. In Fig. 5, the magnitude of ε was exaggerated for visualization purpose. A synchronized test was performed for the final 1 Mc cycles in order to reproduce the usual simulator test conditions.

A load cell (FL2.5U - 2SGKT, Strainsert, PA, USA) was used to verify the accuracy of the simulator control system. There was good agreement between measured load and the imposed load cycle.

The gravimetric wear of the components was measured every 0.25 Mc for the first 2 Mc followed by every 0.5 Mc for the remainder of the test using an analytical balance (OHAUS AP250D, Mettler, Toledo, OH, USA). The volumetric wear was calculated from the weight loss using a density of 8.28 mg/mm^3 for the CoCrMo alloy.

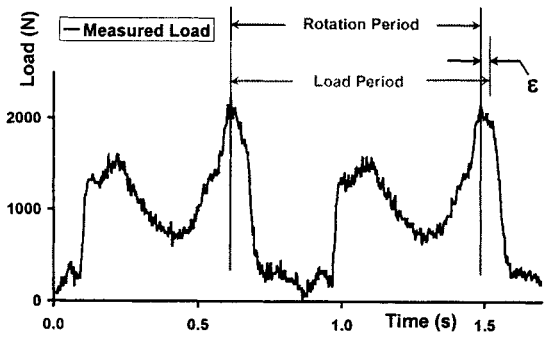


Fig. 5: Load curve showing disengaged synchronization of load and motion.

Surface roughness traces were obtained with a profilometer (Form Talysurf Series II, Taylor Hobson, Leicester, England) with a cut-off length of 0.25 mm and trace length of 1.25 mm. The root mean square (RMS) roughness values of the cup (σ_C) were the average of an RMS surface roughness value (Rq) from each of four traces (except at 0 Mc when it was two to four traces) measured at an "azimuthal" angle of 23° from the apex at various "longitudes" around the cup surfaces. The head RMS roughness values (σ_H) were also the average of an Rq from each of four traces (except at 0 and 1 Mc where only one trace was available) and they were measured at the apex of the heads. Unfortunately, σ_H values were not available for analysis at 1.9 Mc. These measurement locations on both head and cup components were selected because they were within the wear zone [20].

When an occasional Rq at a particular location changed by at least an order of magnitude, it was omitted from the averaging that gave σ_H and σ_C because it was assumed to possess deep scratches. Such scratches might reduce local protection from EHL, but be less influential on global EHL protection. Thus, including traces with scratches would misrepresent the overall surface roughness in the calculation of λ. Although the assumption that high roughness measurements corresponded to the presence of deep scratches was not confirmed in detail, several interferometer measurements (Wyko NT2000, Veeco Process Metrology, Tucson, AZ, USA) performed over visible scratches revealed that they were deep enough to cause high local RMS roughness values.

The implant diameters were measured at 0, 1, 1.9 and 4 Mc using a coordinate measuring machine (CMM – Prismo, Zeiss, Minneapolis, MN, USA). The head diameters were extracted from the best-fit circle passing through a "cloud" of 50 points randomly located over the head surface while 20 to 30 points were used for the cup diameters. These measured "diameters" provided an indication of the on-going wear that was changing the macro-geometry but did not represented the geometry of the inlet zone.

The λ values were calculated during testing by using the current σ but the h_C values were based on the initial geometry. As mentioned above, the actual macro-geometry was changed slightly by the wear but it would have been very difficult to both quantify these changes in sufficient detail to include in a lubrication analysis and then to perform the analysis itself. These changes probably caused the inlet zone lubricant flow to encounter a more sudden decrease in the gap between the surfaces leading to a reduction in lubricant entrainment and a drop in h_C. However, for the small wear volumes of the present study, it was assumed that the likely drop in h_C due to changes in macro-geometry was small and thus the initial h_C would remain fairly constant.

Therefore, following Medley et al [20], the nominal diameters and initial diametral clearance (Table 1) were used to determine both the effective radius (R) and, along with simulator frequency of 1.14 Hz, the entrainment velocity (u). These quantities were then used in Eqn (4) along with F = F_{AVG} = 792 N, E' = 231 GPa and η = 1 mPa·s to calculate an h_C value. This h_C value approximated the minimum over the load cycle, as discussed in previous studies [18, 21].

An estimation of the experimental uncertainty was determined for both σ and λ. Since both of these parameters were related to measured quantities having their own level of uncertainty, an error propagation analysis [22] was performed to estimate the uncertainty range (at the 95% confidence level) of the computed values for σ and λ.

3. RESULTS

3.1. Volumetric Wear

Most of the implants exhibited low volumetric wear (Fig. 6) with initial run-in occurring within the first 0.25 Mc, similar to the results of other simulator studies [2, 3, 4, 6]. Wear rates were determined for

various test intervals (Fig. 7) and declined rapidly after 0.25 Mc. Very low wear and wear rates occurred for implant no. 3 throughout the testing. The occasional very low wear implants had also been found in previous studies [2, 6, 13].

A misalignment between the head and cup of specimen no. 7 (for the first 3 Mc) caused audible evidence of impact loading. This condition probably caused the higher wear rate observed for this implant. An alignment of the components of implant no. 7 was achieved at 3 Mc by re-mounting them. Because of this abnormality, implant no. 7 was not considered further in the analysis and discussion of the present study.

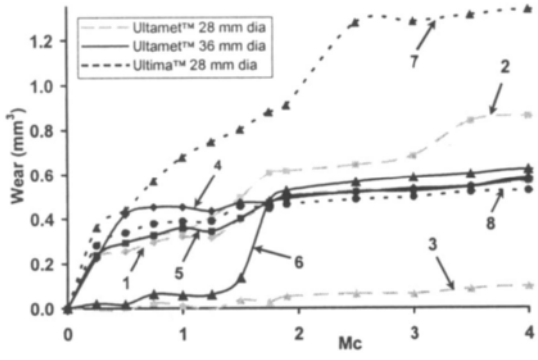


Fig. 6: Cumulative volumetric wear.

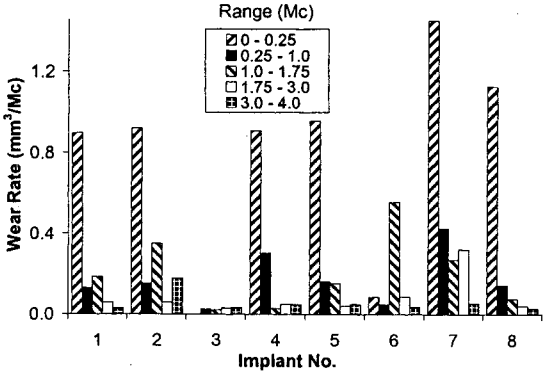


Fig. 7: Total implant wear rate over the indicated ranges, plotted sequentially from left to right.

The wear of the remaining seven implants throughout the simulator test was low and very similar to the findings of some other simulator studies [2, 3, 4]. Three of the seven implants had nominal diameters of 36 mm while the rest had 28 mm, but there was no apparent difference in the wear (Fig. 6).

3.2. Implant Macro-Geometry

The diameters of the heads, as measured by the CMM, tended to increase slightly, possibly because of wear. However, the CMM lacked the resolution to accurately determine any detail of the small changes in macro-geometry of the either the head or cup as wear proceeded. Nevertheless, it was interesting to note that the CMM results indicated a lack of change in the diametral clearance of the "low wearing" implant no. 3 (Fig. 6) that was not found for any of the other implants. This suggested that a different run-in wear mechanism may have acted on implant no. 3 (Figs. 6 and 7).

3.3. Implant Micro-Geometry

The polished as-received implant surfaces had a homogeneous topography as shown by a typical three-dimensional surface map generated by interferometry (Fig. 8). Except for implant no. 4, the σ values of the implants remained fairly constant throughout testing (Fig. 9) for which the average σ_H was 36% lower than the average σ_C.

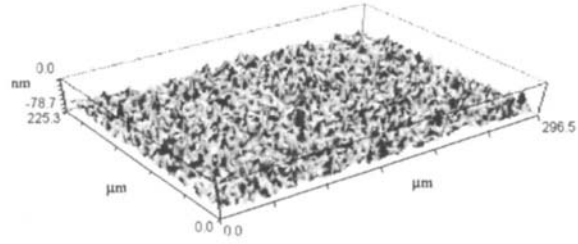


Fig. 8: Initial micro-geometry at the apex of the head of implant no. 4.

From the propagation of error analysis and the estimated uncertainties for σ_H and σ_C, the overall uncertainty range for σ at the 95% confidence level was ± 15 nm. This range meant that the scatter of σ shown in Fig. 9 could have been caused by random measurement errors and differences in the actual topography over the articulating surfaces due to the cascading and local nature of wear rather than reflecting distinctly different physics in the wear process. If wear was a more deterministic phenomenon and identical surfaces could be produced, the disengaged load synchronization would have given a uniform isotropic wear surface topography with little scatter or uncertainty.

However, it was important to recall that the occasional traces with "deep" scratches had been excluded from the σ calculations. Although the values of σ_H and σ_C were generally similar, both heads and cups exhibited some superficial scratches of depth less than 0.1 μm. Chan et al [23] showed similar scratches on implants from simulator testing and from retrievals.

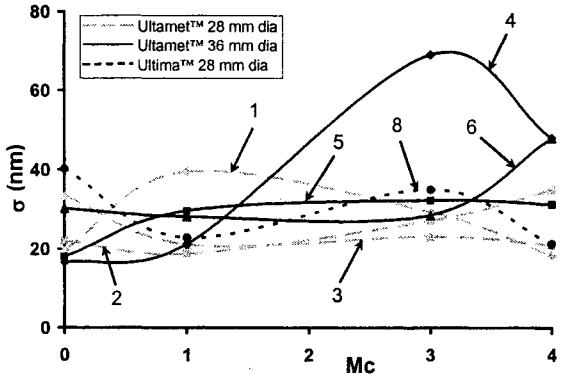


Fig. 9: Combined roughness versus Mc.

Deeper scratches also occurred but only on less than 5% of the traces (Table 2). These scratches could be very wide. For example, implant no. 1 had a scratch that was 5 μm deep and 65 μm wide (Fig. 10). Other studies [7, 14] found similar deep scratches. Interestingly, implant no. 3, that showed remarkably low wear, had neither deep nor superficial scratches.

Table 2: Number of surface roughness traces with deep scratches.

Mc	Number of Deep Scratches	Total Traces Measured	Freq. (%)
1	6	143	4.2
1.9	3	145	2.1
3	4	139	2.9

Profilometer traces provided a more quantitative record of surface topography. Typical traces were selected to illustrate the general trends (Fig. 11). Initially, the surface asperities protruded above the mean of the surface (Fig. 11a) but after 3 Mc both head (Fig. 11b) and cup (Fig. 11c) indicated a rounding of asperity tips. Such rounding of the asperity tips almost certainly occurred before 3 Mc but, unfortunately, the data for the earlier traces were unavailable. It is important to mention that the profilometer traces recorded surface heights directly and only indirectly indicated the asperity geometry.

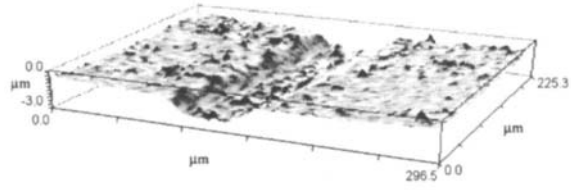


Fig. 10: A deep scratch on the cup of implant no. 1 after 2 Mc.

Typically, the superficial scratches were less common on the head (Fig. 11b) than the cup (Fig. 11c). As shown in Fig. 11d, a deep and wide scratch was found on cup no. 2 (80 μm width by 1.6 μm depth). While the dimensions of this particular case were similar to other deep scratches, the surrounding lips (height ≈ 0.6 μm) were unusual.

3.4. Progression of λ

As expected from the combined roughness plot (Fig. 9), there was no substantial change in λ over the 4 Mc of simulator testing, although the data were somewhat scattered (Fig. 12). The error analysis revealed an overall uncertainty range (95% confidence level) of about ±0.76. Thus, according to the conventional interpretation of λ ranges, the majority of the implants started and remained within a mixed film lubrication regime throughout simulator testing.

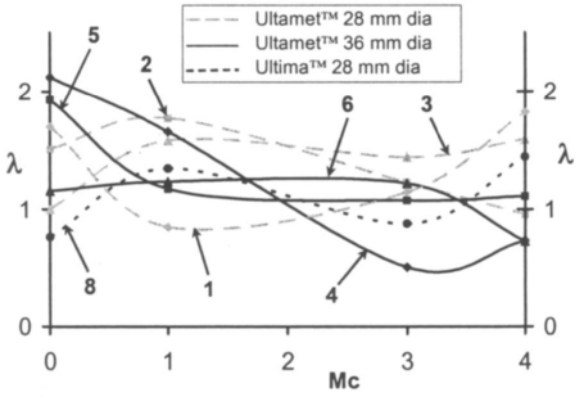


Fig. 12: λ versus number of cycles.

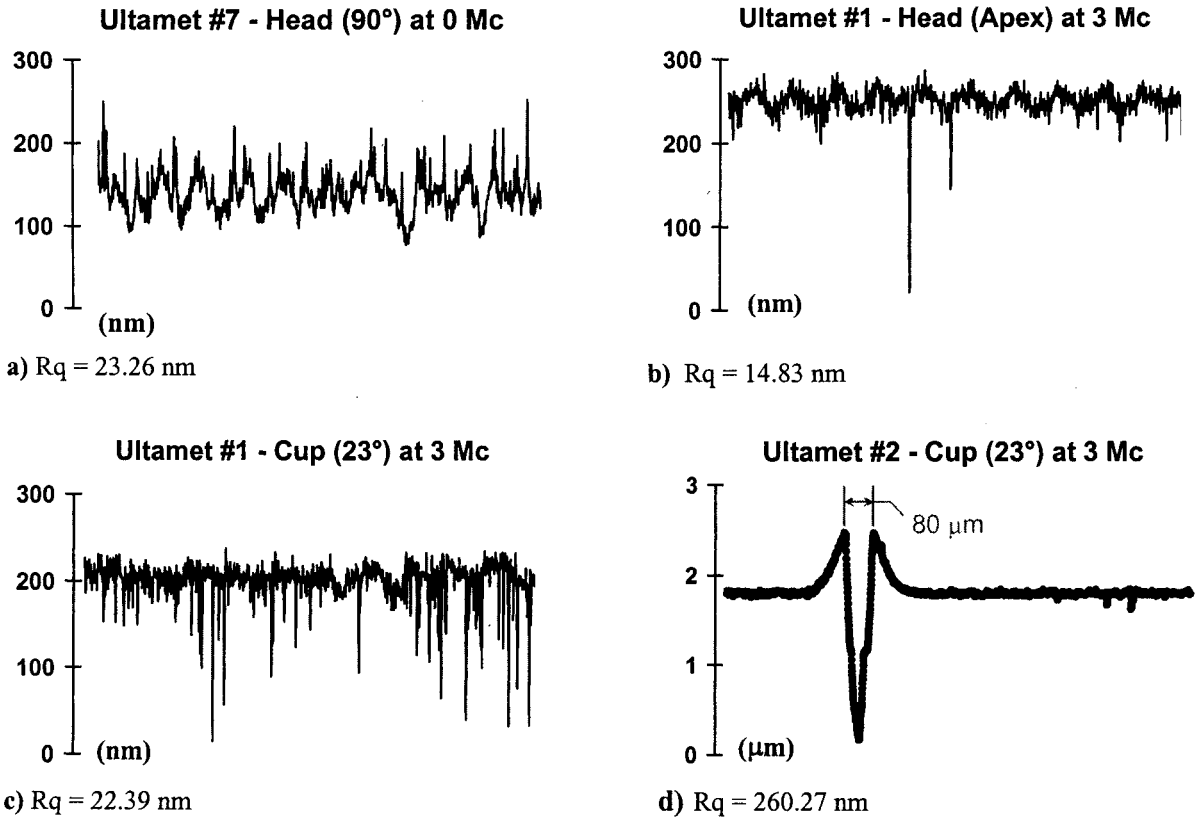


Fig. 11. Typical traces (length: 1.25 mm; angles quoted are measured from the apex pivoted about the geometric centres of the head or cup).

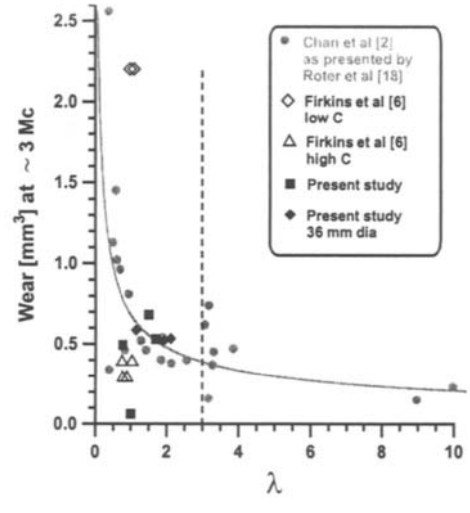


Fig. 13: Wear at 3 Mc versus the initial λ

All the implant diameters in Fig. 13 were 28 mm except as indicated.

3.5. Correlation between λ and Wear

The wear at 3 Mc was plotted versus the initial values of λ (Fig. 13). The initial values of λ were used because no substantial change during simulator testing was found, as mentioned above. The results of previous simulator work [2, 6] were superimposed and a generally consistent pattern emerged. In the present study, there was no obvious difference between the wear of implants with a nominal diameter of 28 mm and those with 36 mm. While this was contrary to the findings of Smith et al [8], the range of diameters and numbers of implants in the present study was small.

4. DISCUSSION

All implants exhibited low wear during the 4 Mc of testing. The wear was similar to the findings of some simulator studies [2, 3, 4, 6] but somewhat lower than that of other simulator studies [5, 7, 8]. The scatter in the wear of retrieved implants [9, 10] was large enough to include the wear of the present study and all of the other quoted simulator studies. This meant that some simulator wear studies could be missing features of *in vivo* wear and this omission would not have been noticed. It was impossible, based just on volumetric wear, to judge which simulator studies produced the most realistic wear. Overall, it was considered likely that simulator studies in general represented many of the characteristics of in vivo wear and could be used to provide more fundamental details of the phenomena but some caution was warranted. These fundamental details included wear during run-in and how the surface topography evolved.

It was not possible to tell if the small changes in σ and λ over the 4 Mc of testing resulted from different wear mechanisms or measurement uncertainties. However, the λ values were well below 3 and this suggested that the implants probably remained under mixed film lubrication conditions. This finding indicated that the run-in wear did not damage the surface enough to prevent a low steady state wear rate from being established. Furthermore, the actual scatter of the data did not exceed the uncertainty range thus providing some support for the precision of the uncertainty analysis.

However, various important features were not represented in the assessment based on λ. The superficial scratches, the rounding the asperity tips and the deeper scratches may have rendered the surfaces non-Gaussian and thus introduced some imprecision into the λ assessment. The deep scratches, although not found very often, might be capable of draining a lubricant film and causing surface damage. However, identifying a density of scratches that could result in significant damage was beyond the scope of the present study.

Implant no. 3 had exceptionally low wear and exhibited an interesting topography. Neither the head nor cup surfaces exhibited notable changes in topography during testing and both remained free of superficial and deeper scratches. The λ value apparently increased from 1 to 1.61 within the first 1 Mc corresponding to a decrease in σ from 33.4 nm to 21.1 nm. The λ value of 1.61 remained within the uncertainty range the other implants at 4 Mc. The surface heights stayed closer to a Gaussian distribution and thus the λ assessment could be considered more accurate than for the other implants. However, it was not clear why this implant, with its estimated λ value well below 3 was free of scratches and exhibited the lowest wear. One possible explanation was that λ was actually at the top of the uncertainty range and thus approaching a state of continuous elastohydrodynamic lubrication that would provide protection against surface wear. However, more data would be required to support this hypothesis.

5. CONCLUSIONS

The purpose of the present study was to explore possible changes in surface protection as wear proceeded by monitoring the values of λ based on the changing values of the combined surface roughness. The starting values of λ were all less than three and did not change substantially during simulator testing. It is possible that λ was not capable of indicating the change in surface protection that caused the transition from run-in to steady state wear. For example, implant no. 3 showed remarkably low wear for reasons that were not understood. However, the present study did begin to document, for the first time, the rounding of the asperity tips and the surface scratches that occurred in the early stages of simulator wear of MM hip implants.

Future work should include a similar exploration of the evolving surface topography for initial $\lambda > 3$ to compare with the findings of the present study. Furthermore, theoretical modelling and experimental research into the influence of both superficial and deeper scratches on EHL would be beneficial in interpreting the findings of the present study. Such wear might be influenced by EHL, alloy microstructure and lubricant chemistry, all of which should be the subject of future research.

6. ACKNOWLEDGMENTS

We would like to acknowledge the superb technical support provided by Ernst Hubert and Andrew Barber. This research was supported by

Depuy Orthopaedics Inc., a Johnson & Johnson company.

REFERENCES

1. **Streicher, R.M., Semlitsch, M., Schon, R., Weber, H.** and **Rieker, C.** Metal-on-metal articulation for artificial hip joints: laboratory study and clinical results. *IMechE J. Engng in Med.*, 1996, **210**, H3, 223-232.

2. **Chan, F.W., Bobyn J.D., Medley J.B., Krygier J.J., Yue S.** and **Tanzer, M.** Wear and lubrication of metal-on-metal hip implants. *Clin. Orthop.*, 1999, **369**, 10-24.

3. **Farrar, R.** and **Schmidt, M.B.** The effect of diametral clearance on wear between head and cup for metal on metal articualtions. *Proc. 43rd Ann. Meeting of the Orthopaedic Research Society*, 1997, 71-12.

4. **Scott, R.A.** and **Schroeder, D.W.** The effect of radial mismatch on the wear of metal on metal hip prosthesis: a hip simulator study. *Proc. 43rd Ann. Meeting of the Orthopaedic Research Society*, 1997, 764.

5. **Goldsmith, A.A.J., Dowson, D., Isaac, G.H.** and **Lancaster, J.G.** A comparative joint simulator study of the wear of metal-on-metal and alternative material combinations in hip replacements. *IMechE J. Engng in Med*, 2000, **214**, H1, 39-47.

6. **Firkins, P.J., Tipper, J.L., Saadatzadeh, M.R., Ingham, E., Stone, M.H., Farrar, R.** and **Fisher, J.** Quantitative analysis of wear and wear debris from metal-on-metal hip prostheses tested in a physiological hip joint simulator. *Bio-Med. Mat. and Eng.*, 2001, **11**, 143-157.

7. **St. John, K.R., Poggie, R.A., Zardiackas, L.D.** and **Afflitto, R.M.** Comparison of two cobalt-based alloys for use in metal-on-metal hip prostheses: evaluation of the wer properties in a simulator. *Cobalt-Base Alloys for Biomedical Applications,* (Eds J.A. Disegi, R.L. Kennedy and R. Pilliar), 1999, ASTM STP 1365, ASTM, West Conshohocken, PA, U.S.A., 145-155.

8. **Smith, S.L., Dowson, D.** and **Goldsmith, A.A.J.** The effect of femoral head diameter upon lubrication and wear of metal-on-metal total hip replacements. *IMechE J. Engng in Med*, 2001, **215**, H1, 161-170.

9. **McKellop, H., Park, S.H., Chiesa, R., Doorn, P., Lu, B., Normand, P. Grigoris, P. and Amstutz, H.** In vivo wear of 3 types of metal on metal hip prostheses during 2 decades of use. *Clin Orthop*, 1996, 329 Suppl., S128-S140.

10. **Rieker, C.B., Sch n, R.** and **Wyss, U.P.** In vivo tribological performance of 177 metal on metal hip arthroplasties. *World Tribology Forum in Arthroplasty*, (Eds. C. Rieker, S. Oberholzer, and U. Wyss), 2001, Hans Huber, 137-142.

11. **Medley, J.B., Bobyn, J.D., Krygier, J.J., Roter, G.E., Cheng, N.Y.** and **Chan, F.W.** Intermittent motion as a simulator protocol for hip implants made from high elastic modulus materials. *Proceedings of the International Conference: Engineers & Surgeons - Joined at the Hip,* Institution of Mechanical Engineers, London, UK, Paper C601/094/2002.

12. **Willert, H.G., Buchorn, G.H..H., G bel, D., K ster, G., Schaffner, S., Schenk, R.** and **Semlitsch, M.**. Wear behavior and histopathology of classic cemented metal on metal hip endoprostheses. *Clin Orthop*, 1996, 329 Suppl., S160-S186.

13. **Medley, J.B., Chan, F.W., Krygier, J.J.** and **Bobyn, J.D.** Comparison of alloys and designs in a hip simulator study of metal on metal implants. *Clin. Orthop.*, 1996, **329 Suppl.**, S148-S159.

14. **Medley, J.B., Dowling, J.M., Poggie, R.A., Krygier, J.J.** and **Bobyn, J.D.** Simulator wear of some commercially available metal-on-metal hip implants. *Alternative Bearing Surfaces in Total Joint Replacement*, (Eds.

J.J. Jacobs and T.L. Craig), 1998, ASTM STP 1346, ASTM, West Conshohocken, PA, U.S.A., 92-110.

15. **Silva, M., Shepherd, E.F., Jackson, W.O., Dorey, F.J.** and **Schmalzried, T.P.** Average patient walking activity approaches 2 million cycles per year. *J. of Arthroplasty*, 2002, **17**, 6, 693-697.

16. **Schey, J.A.** Systems view of optimizing metal on metal bearings. *Clin. Orthop.*, 1996, **329 Suppl.**, S115-S127.

17. **Medley, J.B., Bobyn, J.D., Krygier, J.J., Chan, F.W., Tanzer, M.** and **Roter, G.E.** Elastohydrodynamic lubrication and wear of metal-on-metal hip implants. In *World Tribology Forum in Arthroplasty*, (Eds C. Rieker, S. Oberholzer and U. Wyss), 2001, Hans Huber, 125-136.

18. **Roter, G.E., Medley, J.B., Bobyn, J.D., Krygier, J.J.** and **Chan, F.W.** (2002) Intermittent motion: a novel simulator protocol for the wear of metal-metal hip implants. *Tribology Series 40,* Elsevier, 367-376.

19. **Paul, J.P.** Forces transmitted by joints in the human body. *Proc. Inst. Mech. Engrs.*, 1967, Part J, **181**, 8-15.

20. **Medley, J.B., Krygier, J.J., Bobyn, J.D., Chan, F.W., Lippincott, A.** and **Tanzer, M.** Kinematics of the MATCO hip simulator and issues related to wear testing of metal-metal implants. *IMechE J. Engng in Med*, 1997, **211**, H1, 89-99.

21. **Chan, F.W., Medley, J.B., Bobyn, J.D.** and **Krygier, J.J.** Time-varying fluid film lubrication of metal-metal hip implants in simulator tests. *Alternative Bearing Surfaces in Total Joint Replacement*, (Eds. J.J. Jacobs and T.L. Craig), 1998, ASTM STP 1346, ASTM, West Conshohocken, PA, U.S.A., 111-128.

22. **Gellert, W., Küstner, H., Hellwich, M., Kästner, H.** The VNR concise encyclopedia of mathematics. VNR Editions, 1977, 616-619.

23. **Chan, F.W., Bobyn, J.D., Medley, J.B., Krygier, J.J., Yue, S.** and **Tanzer, M.** The engineering issues and wear performance of metal-on-metal hip implants. *Clin. Orthop.*, **333**, 96-107.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

ADVANCES IN SIMULATOR TESTING OF ORTHOPAEDIC JOINT PROSTHESES

T D Stewart [a], S Williams [a], J L Tipper [b], E Ingham [b], M H Stone [c], J Fisher [a]

[a] School of Mechanical Engineering, The University of Leeds, Leeds LS2 9JT , United Kingdom

[b] Division of Microbiology, The University of Leeds, Leeds LS2 9JT United Kingdom

[c] Department of Orthopaedic Surgery, Leeds General Infirmary, Leeds LS1 3EX, United Kingdom

This paper reports on recent advances in hip joint simulator wear testing which incorporate joint laxity and micro-separation/subluxation of the bearing surfaces into the swing phase of the walking cycle. Under these microseparation conditions the wear of ceramic on ceramic and metal on metal bearings increased by 4 and 25 times respectively, compared to conventional simulation. In contrast, the wear of ceramic on UHMWPE bearings decreased by 4 times possibly due to enhanced squeeze film lubrication brought about by the joint separation. The results show that microseparation simulator wear testing can reproduce clinically relevant wear rates, wear mechanisms and wear debris for hip prostheses.

1. INTRODUCTION

Joint replacement is a very successful treatment for degenerative joints caused by diseases such as arthritis or traumatic injury and approximately 1 million hip and knee prostheses are implanted world-wide per year. Pre-clinical testing has increased the life-span of these devices over the last 40 years through research into improved materials, design and manufacture. The current generation of joint replacements generally last in the region of 10-20 years when loosening of the joint eventually leads to failure. Loosening is believed to be caused by a combination of mechanical and biological factors that are closely linked to the wear debris that is generated from articulation of the joint during normal function. As a consequence, in the last decade extensive joint simulator wear testing, which reproduces a standard walking cycle within a clinically relevant lubricant fluid, has been used to evaluate the wear performance of prosthetic devices.

It has been discovered recently, however, that the bearing components of hip prostheses can separate (0.5-2 mm) in the body during normal walking [1]. Joint separation has been linked to the tension in the ligaments and soft tissues following joint replacement surgery and raises concern over the clinical relevance of current testing methods. The purpose of this study is to report on recent advancements in simulator testing which incorporate joint laxity into the walking cycle of the hip joint simulator to produce more clinically relevant test results [2,3].

2. MATERIALS AND METHODS

Ten wear studies were carried out using three bearing combinations and two hip joint simulators as summarised in Table 1. The three bearing combinations were alumina ceramic heads against GUR 1020 ultra high molecular weight polyethylene (UHMWPE) inserts gamma irradiated in a vacuum foil package (C/PE), high carbon cobalt chromium heads against low carbon cobalt chromium inserts (M/M), and alumina ceramic heads against alumina ceramic inserts (C/C). Each bearing combination was tested under standard conditions and under microseparation conditions. In addition, two simulators were utilised for evaluation of the C/C bearings, the Prosim simulator (I) and the Leeds II simulator (II). Geometrical models of the components were analysed using computer aided design software to validate the two microseparation testing techniques with respect to component geometry to ensure that dislocation of the bearings would not occur.

The duration of all studies was 5 million cycles conducted at a frequency of 1 Hz using 25% v/v bovine serum as a lubricant, which was replaced every 333,000 cycles. Wear was determined gravametrically for metal and ceramic materials and geometrically for UHMWPE every million cycles. Wear surfaces were analysed with a 3D form Talysurf and wear debris analysed using digestion, centrifugation and SEM/TEM. Analysis methods have been extensively reported in a previous publication [4]. All components were 28 mm in diameter and commercially available.

Table 1. Components Tested

Components	Method	Number
C/PE standard C/PE separation	I	5
M/M standard M/M separation	I	5
C/C standard C/C separation	I	5
C/C standard C/C separation	II Mild	3
C/C standard C/C separation	II Severe	3

In-vivo studies have shown separation during the swing phase with relocation at heel-strike [1]. The two methods used in this study to represent the separation *in-vivo* are described below.

2.1. Prosim Hip Simulator Biomechanics

In the Prosim simulator (see Figure 1) an inferior load was applied to the femoral head using an actuator, which, during the swing phase, produced a separation of (~0.7 mm) between the components [5]. This initially resulted in contact of the head with the inferior rim of the insert producing lateral displacement of the head (Position 1 to C). At heel-strike this resulted in superior translation of the head, which then contacted the superior rim of the insert (Position 2) before returning to it's normal articulating position (Position 1).

2.2. Leeds II Hip Simulator Biomechanics

In the Leeds II simulator (see Figure 2) a lateral to medial load was applied to the acetabular insert with a spring, which, during the swing phase produced medial (~0.4 mm) and superior translation of the insert as it tracked around the head (Position 1 to 2) [3]. At heel-strike (30° flexion) this resulted in contact between the head and the superior rim of the insert (Position 2) before returning to it's normal articulating position (Position 1). The Leeds II simulator had two test configurations, mild and severe.

Severity was altered by adjusting the swing phase load from 400 N for mild microseparation to 50 N for severe microseparation. The reduced swing phase load increased the efficiency of the separation spring thus producing an increased momentum and impact energy representative of greater laxity.

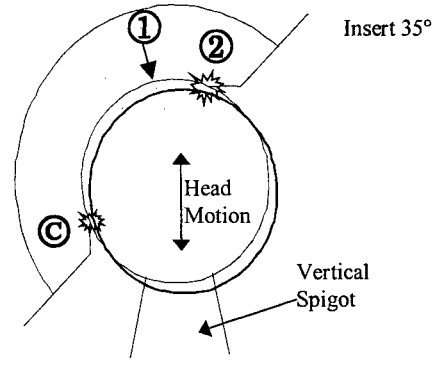


Figure 1. Prosim Microseparation

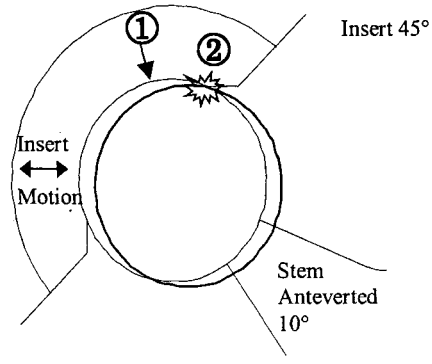


Figure 2. Leeds II Microseparation

3. RESULTS AND DISCUSSION

Both simulation techniques produced small separations of the joint centres, which resulted in contact of the femoral head with the internal chamfer at the superior rim of the acetabular insert at heel-strike. This produced a momentary stress concentration before the bearing returned to its normal articulating position. No dislocation of the

bearings occurred at any time during testing. The wear results were higher during the initial million cycles of testing (bedding-in) for all three bearing combinations. However, the wear rate then reduced to lower steady-state levels for the remaining 4 million cycles.

Average wear rates for the C/PE, M/M and C/C bearings are listed in Table 2 for standard and microseparation (separation) testing in units of mm^3/million cycles [3,5,6]. These are compared to *in-vivo* results in mm^3/year reported in a recent review paper by Tipper [4] for C/PE [7,8,9], M/M [10,11,12] and C/C [13,14]. The wear of C/PE in the Prosim simulator (superior/inferior separation ~0.7 mm) was found to reduce with microseparation by a factor of 4, to a lower level than observed clinically [5]. Plastic deformation was observed on the rim of the polyethylene inserts in the order of 0.5 mm, while no damage was observed on the alumina ceramic heads. Clinical observations of plastic deformation on polyethylene insert rims have not been reported. However, past wear studies have generally only consider the main bearing area.

Table 2. Average overall wear results.

Component	In vivo mm^3/ year	Standard Mm^3/mil cycles	Separation mm^3/mil cycles
C/PE (I)	25-100	25.6± 5.3	5.6 ± 4.2
M/M (I)	0.3-5	0.09± 0.05	2.3 ± 1.2
C/C (I)	0.1-1	0.05± 0.03	0.2 ± 0.03
C/C(II)Mild C/C(II)Severe	0.1-1	0.07± 0.02	0.1 ± 0.05 1.8 ± 0.4

Note: I – Prosim simulator, II – Leeds II simulator.

The wear of M/M in the Prosim simulator (superior/inferior separation ~0.7 mm) was found to increase with the introduction of microseparation by a factor of 25, similar to the levels reported clinically[5]. Wear stripes were observed on the femoral heads with matching wear to the rim of the acetabular inserts. This stripe wear pattern has been observed clinically with ceramic on ceramic bearings, but not with metal on metal bearings. It has been postulated, however, that the self-polishing action of metal articulating against metal may mask this stripe formation clinically. In addition, clinically the separation of the bearing will undoubtedly be more erratic than in the simulator where repeated actions produce more severe damage.

The wear of C/C was found to increase with the introduction of microseparation by a factor of 4 in the Prosim simulator and by a factor of 2 and 36 in the Leeds II simulator (medial/lateral subluxation ~0.4 mm) under mild and severe conditions respectivley[3,5]. Wear rates with microseparation were similar to those reported *in-vivo*. A localised worn stripe area was observed on the femoral heads with a matching wear stripe to the rim of the acetabular inserts in all components using both simulator techniques. Stripe areas were the same colour as the Alumina material. To assist in photographing the components the stripe areas were rubbed with a pencil, as shown in Figures 3-5. Stripe wear patterns were similar to wear stripes observed clinically with the same alumina material, as shown by the pencil rubbing in Figure 6. The component shown was retrieved after approximately 1 year and was well fixed prior to retrieval. Intragranular fracture of the ceramic was observed within the wear stripe as shown in Figure 7, similar to previously reported *in-vivo* studies [2].

Over the 5 million cycles test duration no severe runaway wear, as reported for some historical retrieved Mittelmeier ceramic on ceramic bearings occurred and a steady state wear was established following stripe formation with re-polishing occurring within the wear stripe [13]. These early devices, however, were manufactured from 1st generation alumina and cannot be compared to current designs with improved hot isostatic pressed material. In addition, these stripe areas should not be confused with metal transfer, which is also observed *in-vivo* during retrieval and appears as a black stripe.

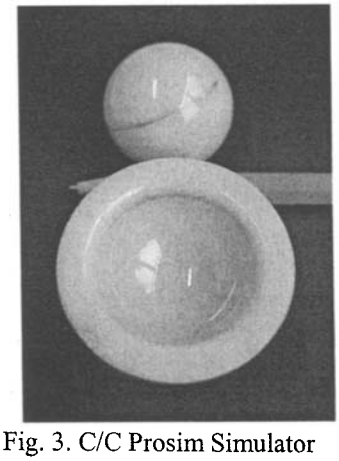


Fig. 3. C/C Prosim Simulator

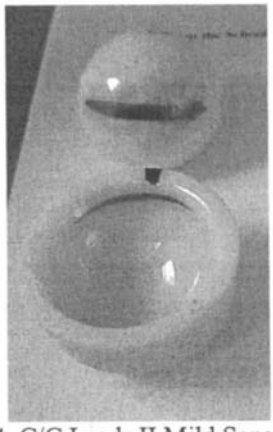


Fig. 4. C/C Leeds II Mild Separation
* felt tip alignment mark on superior insert face

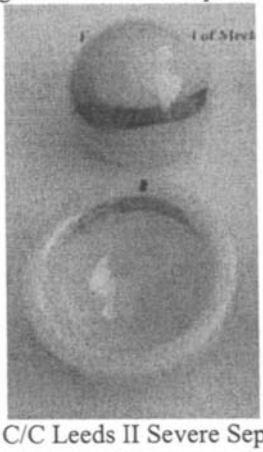


Fig. 5. C/C Leeds II Severe Separation
* felt tip alignment mark on superior insert face

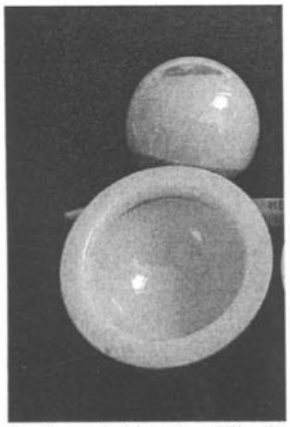


Fig. 6. Explanted Alumina Hip Prostheses

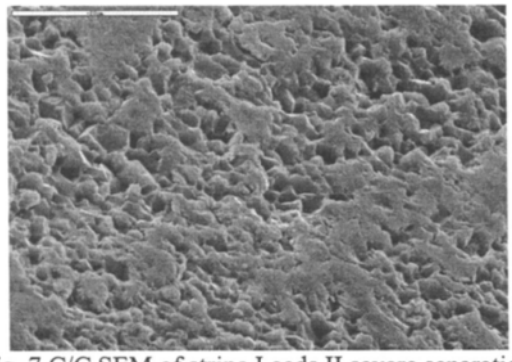


Fig. 7.C/C SEM of stripe Leeds II severe separation.

The results of CAD geometrical studies revealed that from the geometry alone, C/C and M/M prostheses with radial clearances in the order of 30-40 μm require only 0.1mm of relative motion between the head and insert for head/rim impact to occur. Figures 3 to 6 can be used to compare the biomechanics of the different simulation techniques. Typical head stripe dimensions for the C/C bearings are also shown in Table 3 for comparison. In the Prosim simulator the loading vector was along the axis of the femoral head (30° flexion). Impact was less direct with very little effect of momentum of the insert and its holder. Consequently, a long, thin, less pronounced stripe of wear resulted on the M/M and C/C bearings tested in the Prosim simulator. In the Leeds II simulator the loading vector was vertically downwards through the insert, essentially perpendicular to the contact. Impact was, therefore, very direct and included the momentum of the insert and its holder. A deeper and more pronounced stripe of wear resulted on the C/C bearings tested in the Leeds II simulator. Retrieval alumina C/C bearings show slightly shorter stripes to both simulator methods, however, the wear mechanisms clinically may be patient/implant/surgeon/technique related.

Table 3. Typical C/C head stripe dimensions

	Depth	Width	Length
Current Generation Explant	6 μm	2-4 mm	18mm
Leeds II Mild Separation	10 μm	2-3 mm	26 mm
Leeds II Severe Separation	90 μm	4-5 mm	28 mm
Prosim Separation	8 μm	1 mm	40 mm

The wear debris from M/M and C/C microseparation hip simulator wear testing was similar to debris from retrieved tissues [4,6,15]. M/M debris contained small nanometer sized particles (mean 12 nm)[16]. C/C debris contained

mainly small nanometer sized particles (mean 10-40 nm) and a few larger particles up to one micrometer in size associated with microseparation[17].

There is currently much debate as to the occurance of joint microseparation *in-vivo*. Fluroscopic studies have shown strong evidence of hip joint separation [1]. Conversley instrumented hip prostheses have shown that due to the actions of muscle forces a positive normal load allways exists throughout the gait cycle which prevents joint separation [18]. In the Prosim hip simulator (superior/inferior separation ~0.7 mm) a negative joint load was applied to the bearings which agrees with the fluroscopic evidence of Lombardi but contradicts the studies of Bergmann. The motion of the Leeds II simulator (medial/lateral subluxation ~0.4 mm) agrees with Lombardi as the joint centres are separated and also agrees with Bergmann as there is always a normal joint load acting on the bearings. The authors speculate that the joint subluxation produced by the Leeds II simulator is the most physiological as geometrically only 0.1mm of relative translation of the head and insert bearing centres is required for rim contact. Nevertheless, both experimental techniques are capeable of reproducing the stripe wear mechanisms that have been reported in retrieved C/C devices but not in conventional hip joint simulators [2,14]. The true cause of the stripe wear observed clinicaly is postulated to be related to an instability in the hip; a multifactoral problem associated with the relationship of the patient, implant design, surgical technique, and the prescribed physiotherapy.

4. CONCLUSIOINS

The introduction of laxity into the swing phase of the hip simulator cycle produced contact of the femoral head against the rim of the acetabular insert at heel-strike prior to the bearing returning to it's normal articulating position. The bearings did not dislocate at any time in the study. Microseparation produced an associated increase in the wear rate of metal on metal and ceramic on ceramic bearings by a factor of 4 and 25 times respectively, in the Prosim simulator (superior/ inferior separation ~0.7mm) compared to conventional simulator testing. Wear stripes were observed on the femoral heads with matching wear to the rim of the acetabular inserts. This pattern of increasing wear was repeated with the Leeds II simulator (medial/lateral subluxation ~0.4mm), where increases in wear rate of 2 to 36 times were found for C/C bearings under mild and severe conditions respectively, associated with the different simulator kinematics.

In contrast microseparation produced a decrease in the wear of ceramic on polyethylene bearings in the Prosim simulator by a factor of 4, where creep deformation and damage to the rim of the polyethylene insert was observed. It is postulated that the decrease in wear is due to squeeze film lubrication conditions brought about by the joint separation.

The wear debris from C/C and M/M microseparation hip simulator wear testing was similar to debris from retrieved tissues.

Bearing radial clearance defines the relative translation or subluxation required for head/rim contact and the associated stripe wear patterns to occur. The results have implications towards the design of implants that promote or discourage joint laxity through implant subsidence. In addition, for ceramic bearings, microseparation testing provides a clinically relevant test that can differentiate between the wear of different ceramic materials which is not possible using conventional hip joint simulators.

The results show that microseparation testing can reproduce clinically relevant wear rates, wear mechanisms and wear debris for hip prostheses which is essential for pre-clinical evaluation.

5. ACKNOWLEDGEMENTS

This work was supported by Action Research, The Engineering and Physical Sciences Research Council, The Arthritis Research Campaign, and The National Health Service.

6. REFERENCES

1. A V Lombardi, T H Mallory, D A Dennis, R D Komistek, R A Fada, E J Northcut, "An *in-vivo* determination of total hip arthroplasty pistoning during activity". Journal of Arthroplasty, 15, 702-709, 2000.
2. J E Nevelos, E Ingham, C Doyle, R Streicher, A B Nevelos, W Walter, J Fisher, "Microseparation of the centres of Alumina-Alumina artificial hip joints during simulator testing produces clinically relevant wear rates and patterns". Journal of Arthroplasty, 15, 793-795, 2000.

3. T Stewart, J L Tipper, R Streicher, E Ingham, J Fisher, "Long Term Wear of HIPed Alumina on Alumina Bearings for THR Under Microseparation conditions", Journal of Material Science: Materials in Medicine 12, 1053-1056, 2001.
4. J L Tipper, P J Firkins, A A Besong, P S M Barbour, J Nevelos, M H Stone, E Ingham, J Fisher, "Characterisation of wear debris from UHMWPE on zirconia ceramic, metal-on-metal and alumina ceramic-on-ceramic hip prostheses generated in a physiological anatomical hip joint simulator". Wear 250, 120-128, 2001.
5. S Williams, T D Stewart, E Ingham, M H Stone, J Fisher, "Influence of microseparation and joint laxity on wear of ceramic on polyethylene, ceramic on ceramic and metal on metal total hip prostheses". Proceedings of the British Hip Society, Feb. 2002.
6. J Fisher, J Bell, P S M Barbour, J L Tipper, B Mathews, A A Besong, M H Stone, E Ingham, "A novel method for the prediction of the functional biological activity of polyethylene wear debris". Proc Instn Mech Engrs, H215, 127-132, 2001.
7. M Semlitsch, H G Willert, "Clinical wear behaviour of ultra-high molecular weight polyethylene cups paired with metal and ceramic ball heads in comparison to metal-on-metal pairings of hip joint replacements". Proc. Inst. Mech. Eng.,H211, 73-88, 1997.
8. H Minakawa, M H Stone, B M Wroblewski, J G Lancaster, E Ingham, J Fisher. " Quantification of third-body damage and its effect on wear with different types of femoral head". Journal of Bone and Joint Surgery, 80-B, 894-899, 1998.
9. I C Clarke, A Gustafson. "Clinical and hip simulator comparisons of ceramic-on-polyethylene and metal-on-polyethylene wear". Clin. Orthop. Rel. Res. , 379, 34-40, 2000.
10. L D Dorr, Z Wan, D B Longjohn, B Dubious, R Murken". "Total hip arthroplasty with use of the Metasul metal-on-metal articulation". Journal of Bone and Joint Surgery, 82, 789-798, 2000.
11. H MeKellop, P Doorn, R Chiesa R, S H Park, H Amstutz. "Twenty-year retrieval analysis of metal-on-metal hip prostheses". Proc. 42nd Orthopaedic Research Society, No. 456, Atlanta, USA, 1996.
12. P Roberts P Grigoris, "Metal on metal articulation in total hip replacement", In: Interfaces in Total Hip Arthroplasty, I D Learmouth Ed., 121-133, Springer-Verlag, 1999.
13. J E Nevelos, F Prudhommeaux, M Hamadouche, C Doyle, E Ingham, A Meunier, A B Nevelos, L Sedel, J Fisher, "Comparative analysis of two different types of alumina-alumina hip prostheses retrieved for aseptic loosening". Journal of Bone and Joint Surgery, 83-B, 598-603, 2000.
14. T Shishido, I C Clarke, P Williams, A Gustafson, H Shoji, M Boehler, K Keggi, A Imakiire, "Clinical wear and simulator study of Ceramic-Ceramic THR to 20 years and beyond". Proc. 48th Orthopaedic Research Society, Poster 1019, Dallas, Texas, 2002.
15. A Hatton, E Ingham, J E Nevelos, J Fisher, A B Nevelos,"Histological analysis of tissue from revised uncemented ceramic on ceramic total hip arthroplasties". Proc 6th World Congress of Biomaterials, 1188, 2000.
16. S Williams, J L Tipper, T D Stewart, E Ingham, M H Stone, J Fisher, "Effect of swing phase microseparation on the wear and biological activity of metal on metal total hip replacements". Proceedings of Institute of Materials Biomaterials Day, April 2002, London UK.
17. J Fisher, S Williams, T D Stewart, J L Tipper, B Mathews, M H Stone, E Ingham, "Functional biological activity and osteolytic potential of wear debris generated in artificial hip joints". Proceedings IMechE Joined at the Hip, June, 2002, London UK.
18. G Bergmann, R Graichen, A Rohnmann "Investigation of potential separation between head and cup of total hip implants". Proc. 48th Orthopaedic Research Society, Poster 0982, Dallas, Texas, 2002.

SESSION VIII
TRIBOLOGICAL COATINGS - ADVANCES IN SURFACE ENGINEERING (2)

Chairman : Dr P A Dearnley

Paper VIII (i) **Rolling Contact Fatigue of Nitrided 32CrMoV13 Steel**

Paper VIII (ii) **Effect Hf Addition on the Tribological Performance of ZrN Coating**

Paper VIII (iii) **Tribological Properties of TiAlSiN Film Coated by AIP Equipment onto SKD61 Substrate**

Paper VIII (iv) **Wear Mechanisms in Thermal Spray Steel Coatings**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Rolling Contact Fatigue of Nitrided 32CrMoV13 steel

C. Jacq[a], D.Nélias[b], G. Lormand[c], A. Vincent[c], G. Dudragne[d]
[a] Direction Technique, Snecma Moteurs, Centre de Villaroche, 77550 Moissy-Cramayel, France
[b] INSA de Lyon, LaMCoS, UMR CNRS 5514, 69621 Villeurbanne Cédex, France
[c] INSA de Lyon, GEMPPM, UMR CNRS 5510, 69621 Villeurbanne CEDEX, France
[d] SNR Roulements, BP2017, 74010 Annecy, France

In this paper, dent initiated Rolling Contact Fatigue (RCF) of nitrided 32CrMoV13 steel is studied, both experimentally and numerically, from the indentation process to the RCF failure. First, a method based on nano-indentation testing has been set up to identify the local elastic-plastic behaviour of nitrided 32CrMoV13 steel, hence enabling the characterisation of the micro-yield shear stress profile. Fatigue tests have then been conducted to evidence the influence of different operating conditions on RCF and to compare nitrided 32CrMoV13 steel to a reference material. Last, a numerical study is presented. An indent based endurance limit, H1I is determined. The comparison of H1I with experimental results show that it is a coherent definition of the endurance limit.

1. INTRODUCTION

Rolling contact fatigue (RCF) is major cause of rolling bearing failure [1]. RCF due to the repeated contact between rolling elements is drastically accelerated when surface defects, such as dents, exist on the surface of the contacting bodies [2]. Dents can have two origins. First, they can be created during handling, transporting and mounting operations. Care during these operations should avoid dent occurrence. Second, dents can be due to wear debris or ingested contaminant from the environment and carried by the lubricant. Solid particles with diameter one or two orders of magnitude the EHL film thickness may be entrapped between contacting bodies and thus plastically deform their surface [3]. Dents created by the lubricant contamination can not be totally removed since perfect filters or seals do not exist. Therefore rolling element bearings should be designed to bear surface dents. To reduce the number of failures due to these surface defects, that act as local stress raisers, it is required to increase the tolerance of rolling bearings to surface defects. For this reason, the behaviour of nitrided 32CrMoV13 steel (AMS 6481) has been investigated. The nitriding hardens the material surface, not only in the area concerned by the surface defect by also up to the Hertzian depth, and introduces compressive residual stresses in the nitrided layer [4] that are known to be benefit for RCF. These two points make nitrided 32CrMoV13 steel a good candidate when dent initiated RCF is concerned.

The first step to improve understanding of the RCF behaviour of nitrided 32CrMoV13 steel is to measure its mechanical properties. Lamagnère et al. [5] have shown that micro yield shear stress (yield shear stress for a plastic deformation of 0.0002%) is a physical criteria of endurance limit as far as RCF is concerned. However, its determination in the case of surface hardened material is more complex since it requires a local measurement by nano-indentation function of depth. Moreover the denting process

involves plasticity, therefore it is also necessary to determine the hardening behaviour of the material to understand phenomena induced by indentation. Rolling fatigue tests have then be performed on the twin discs machine at the Mechanical Contact Laboratory (LMC) of INSA-Lyon [6]. Geometry and load conditions have been carefully chosen to reproduce the contact stress profile encountered in an aeronautic rolling bearing and especially its evolution with regards to micro yield shear stress profile. Last, a numerical study of dent initiated RCF has been conducted. Most numerical studies consider the dent only by its geometry and neglect plastic hardening of the material and residual stresses introduced during the indentation process, although they can be very significant [7]. The whole process that leads to failure, from indentation to the over-rolling of the dent has been simulated. RCF endurance limit has then been calculated for dents tested experimentally and compared for nitrided 32CrMoV13 steel and for M50 (AMS 6490), a reference steel for aeronautic rolling bearings.

2. MECHANICAL PROPERTIES OF NITRIDED 32CrMoV13 STEEL

The material studied is 32CrMoV13 steel. Its chemical composition is given in table 1. Before nitriding at 550°C, 32CrMoV13 steel was austenised at 950°C, quenched in oil and then annealed at 630°C. The compound layer, which formed at the surface of the material during nitriding, was removed by grinding. The penetration depth of nitrogen obtained after this treatment is about 600 µm.

The study of nitrided 32CrMoV13 steel dent initiated RCF requires the knowledge of (i) the micro yield shear stress profile, required to define the endurance limit; (ii) the hardening law of the material, necessary to describe the indentation process that involves plasticity ; (iii) residual stresses induce by the nitriding treatment. They contribute to stresses undergone by the material during indentation and over-rolling.

Micro yield shear stress and uniaxial hardening behaviour are usually measured by global compressive tests made on material samples. This process gives average values over the section of the sample. Therefore, to be relevant, this test must be made on uniform samples. Two uniform samples (A and B) have been tested. The first sample, A, is made of nitrided 32CrMoV13 from which the nitrided layer has been totally removed. It represents then the core material. The second sample, sample B, has been obtained by nitriding homogeneously thin tubes of 32CrMoV13 steel [8]. It represents the behaviour of a uniform nitrided layer. Results obtained in this two tests are plotted in figure 1. Two conclusions arise from these tests. First, the elastic plastic behaviour can be represented by the Swift law ($\sigma = B(C+10^6\times\varepsilon^p)^n$) [8], where B, C and n are three parameters of the law. This law describes the uniaxial hardening and gives the micro-yield shear stress τ_e ($2\tau_e = B(C+20)^n$). Second, both tests can be described with the same C and n parameters, here equal to 16 and 0.067, respectively. Therefore, we can conclude that the variation of the mechanical behaviour in the nitrided layer can be described by the variation of τ_e.

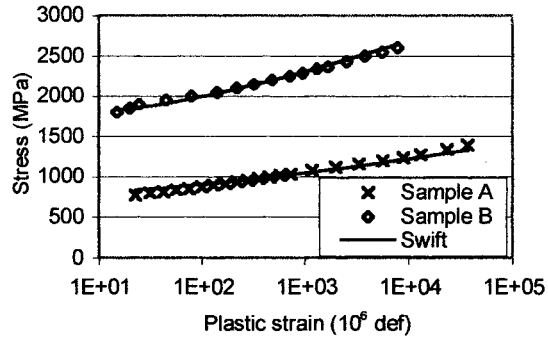


Figure 1. Stress-plastic strain curve for samples A and B

Table 1. Chemical composition of the nitrided 32CrMoV13 steel

C	Si	Mn	Ni	Cr	Mo	V
0.334	0.24	0.48	0.10	2.93	0.9	0.32

The characterisation of the mechanical behaviour profile requires a local measurement that can not be obtained by a global compressive test. Nano indentation test has then been used to stress locally the material. Experiments have been conducted on Nano-instruments model II nano-indenter, that has been fitted with a spherical diamond tip of 105 µm radius. This apparatus, (schematically described in figure 2) allows the simultaneous record of the applied load W (ranging form 0 to 700 mN) and of the displacement of the indentation head. El Ghazal

et al. [8] have shown that for nitrided 32CrMoV13 steel, an applied load of 650 mN deforms plastically a volume that is contained in cube of 20 μm edge. This volume can be considered as local with regards to the nitrided layer depth, but large enough for continuum mechanic modelling to be valid [8,9].

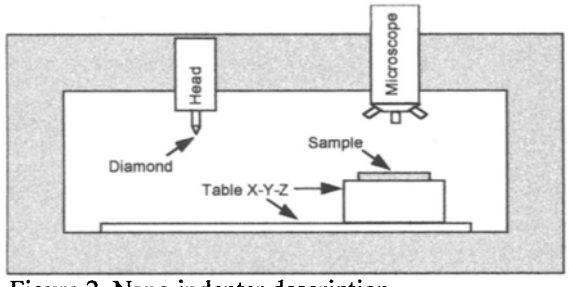


Figure 2. Nano indenter description.

The depth of the print, created during nano-indentation test, hereafter called residual displacement, e_r, is a function of the applied load W, and of the local elastic plastic behaviour of the material, that can be represented by the value of τ_e, E and ν. Therefore, since the load is known and the elastic parameters are constant in the depth of the nitrided layer (E=210 GPa, ν = 0.3 [8]), there is a direct relation between e_r and τ_e. To measure the depth of the print, the indentation head is loaded and unloaded at increasing maximum load, following the procedure shown in figure 3. The displacement is recorded at the same time. The residual displacement is calculated as the difference between displacement for a low load after unloading and before the first load. In a single test, e_r is then determined for several loads W and for the same local elastic plastic behaviour.

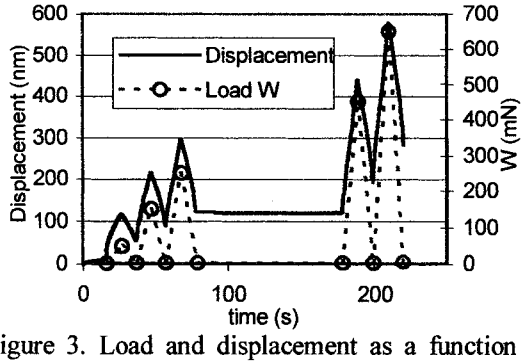


Figure 3. Load and displacement as a function of time during nano-indentation test.

To identify the relation between the load W, the micro-yield shear stress τ_e and the residual displacement e_r, nano-indentation test has been simulated, using a 3D elastic plastic contact code [10]. The code was first validated by simulating nano-indentation test on sample B (figure 4), where the elastic plastic behaviour is known by global compressive test. Simulations have been carried out for several values of τ_e. The variation of e_r with W and τ_e (figure 5) can be closely approximated by the function given in equation 1, whose parameters are given in table 2. In presence of residual stresses, a correction of the residual displacement must be made to remove the influence of residual stress on e_r. This correction, included in equation 1, has been determined by Jacq et al. [10] for nitrided 32CrMoV13 steel.

$$\text{if } \tau_e < \tau_{e0}$$
$$e_r(\tau_e, W, \sigma_r) = tr(\sigma_r)\,[c_1(W)\ln(\tau_e)+c_2(W)] + c_3(W)[\tau_{e0}(W)-\tau_e]^{c4(W)}$$
$$\text{else} \quad e_r = 0 \qquad (1)$$

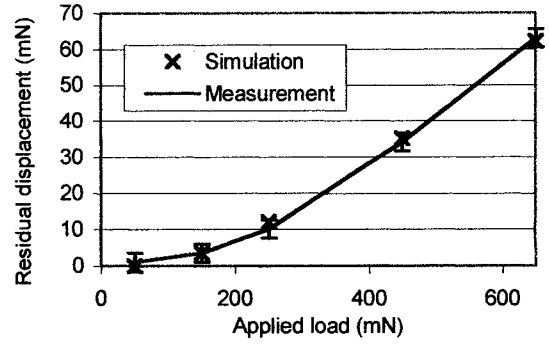


Figure 4. Variation of residual displacement with maximum applied load for sample B. Measurement and simulation.

This process has been applied to the measurement of the elastic plastic behaviour of the nitrided layer of 32CrMoV13 steel. Nano-indentation tests have been made at different depth of the nitrided layer, on a cut perpendicular to the nitrided surface.

Residual stresses introduced by the nitriding treatment have been measured by El Ghazal et al. [8] with the layer removal technique. Residual stresses are highly compressive on surface (-400 MPa), reaches a maximum at 500 μm depth (-500 MPa) and then decreases below the nitrided layer (figure 7). A finite element method has been used to determine from this residual stress profile the trace

of the residual stress profile along the cut perpendicular to the nitrided surface.

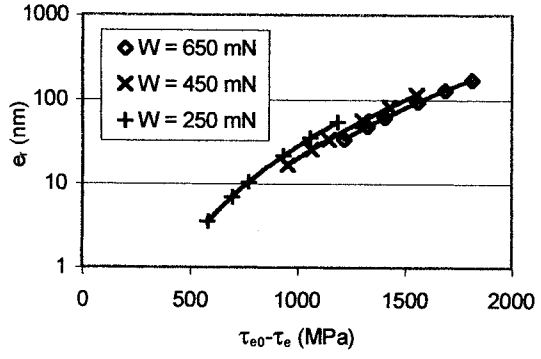


Figure 5. Variation of residual displacement with τ_e for different loads.

The micro yield shear stress profile has thus been obtained and is presented in figure 6. It decreases continuously from 1170 MPa at surface to 420 MPa at core.

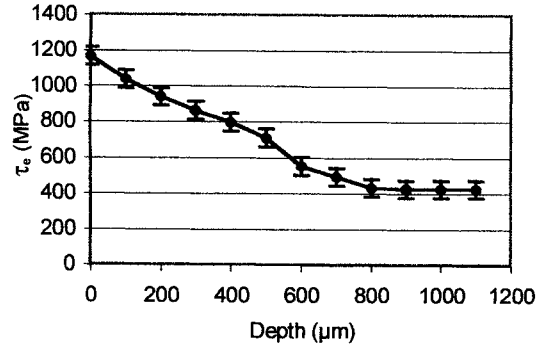


Figure 6. Nitrided 32CrMoV13 steel micro yield shear stress profile.

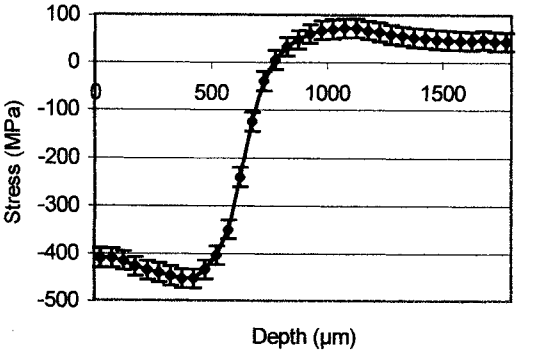


Figure 7. Residual stress profile of nitrided 32CrMoV13 steel [8].

Table 2. Parameters of equation 1.

W (mN)	650	450	250
c_1 (W)	-0.0325	-0.0272	-0.0204
c_2 (W)	0.257	0.212	0.156
c_3 (W)	$8.56\ 10^{-13}$	$2.36\ 10^{-12}$	$4.83\ 10^{-12}$
c_4 (W)	4.02	3.92	3.87
τ_{e0}(W) (MPa)	2320	2060	1690

3. RCF EXPERIMENTAL INVESTIGATION

Experimental investigations have been conducted on dent initiated RCF for nitrided 32CrMoV13 steel. Fatigue tests have been done on a twin discs machine [6] whose principle is recalled in figure 8. Two rotating discs are brought in contact with an imposed load. The rotating speed of discs is controlled independently, hence enabling various contact conditions, such as pure rolling or controlled sliding. Nitrided 32CrMoV13 steel is surface hardened. Therefore, the stress level undergone in the material should not be described only by Hertz pressure because it does not give any depth indication for the superposition of the contact stress profile, of the residual stress profile and of the micro yield shear stress profile. This problem has been stated by Vincent et al. [11], who have described the variation of the RCF endurance limit Hl for a spherical point contact as a function of the contact equivalent radius of curvature in the case of nitrided 32CrMoV13 steel. Subsequently, special care has been brought in the choice of operating conditions so that the stress profile undergone by the discs (noted TD in figure 9) is close to the one encountered in the aimed application, i.e. a typical aeronautic rolling bearing (ARB) (indicated by ARB in figure 9). This consideration has led to the choice of an equivalent point contact radius of 15 mm, obtained with a specimen disc (dented) of 24 mm radius and the opposite disc (smooth) of 40 mm radius. Two loads have been considered, 2.5 GPa and 3 GPa. The radius of the contact zone is then 510 and 610 µm, respectively. Nitrided 32CrMoV13 steel specimen discs have been tested again M50 discs, since in bearing application, M50 roller will be used with nitrided 32CrMoV13 rings. For reference, tests have also been made with M50 specimen discs. The contact is lubricated with an MIL-L-23699 turbine oil at 80°C. The linear velocity at contact is set to 24 $m.s^{-1}$. To evidence the influence of sliding, tests have been carried out in pure rolling, and with a sliding ratio of 1.5 %. In

presence of sliding, the slower disc is the nitrided one. Last, the influence of hoop stresses has been investigated, by adding 115 MPa of traction on some nitrided discs. Operating conditions corresponding to tests performed are summarised in table 3.

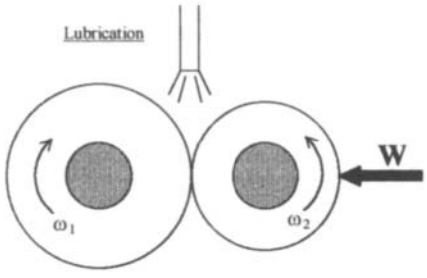


Figure 8. Twin discs machine scheme.

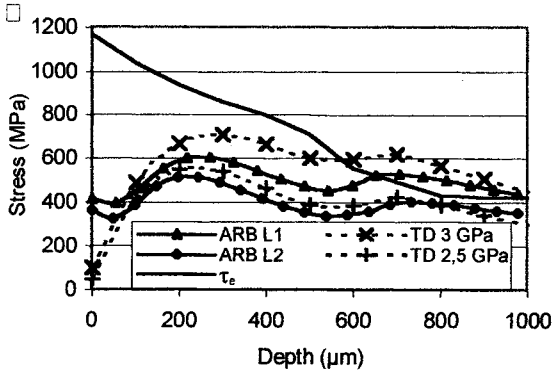


Figure 9. Von Mises stress profile and micro yield shear stress profile.

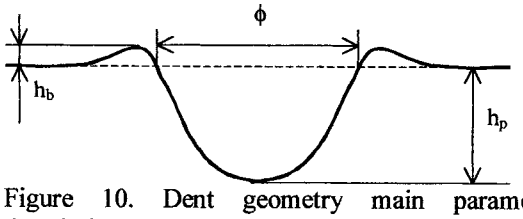


Figure 10. Dent geometry main parameters description.

Table 3. Dents geometry main parameters range.

	ϕ (µm)	h_p (µm)	h_b (µm)
From	85	2.5	0.4
To	290	32	4

Artificial controlled dents have been made with spherical diamond indenter of different radii R_p (200, 400 and 600 µm) at various loads W_i (W_i =5, 10, 20, 30, 50 kg). All dents have been made on the 24 mm radius disc. They are identified by the radius of the indenter and by the applied load. Hence, dent 200-20 is a dent made with R_p = 200 µm under load W_i = 20 kg. The range of dent characteristics thus obtained (diameter ϕ, depth h_p and shoulder height h_b, figure 10) are summarised in table 4. Up to 18 dents have been tested on each specimen.

Table 4. Fatigue tests operating conditions

Label	Specimen	Load (GPa)	Sliding	Hoop stress
C1	M50	2.5	0	no
C2	M50	3	1.5 %	no
ES1	Nitrided	2.5	0	no
ES2	Nitrided	2.5	0	yes
ES3	Nitrided	2.5	1.5 %	no
ES4	Nitrided	2.5	1.5 %	yes
ES5	Nitrided	3	0	no
ES6	Nitrided	3	0	yes
ES7	Nitrided	3	1.5 %	no
ES8	Nitrided	3	1,5 %	yes

To analyse experimental tests and determine the harmfulness of dents, a visual damage criteria has been set up. To determine a test duration, two tests, C1 and C2, have been run during 40 and 50 10^6 cycles. On this two tests, all damages visible at the end were initiated before 20 10^6 cycles (figure 11). Therefore a dent is considered as harmful for a given operating condition if micro spalls have appeared around the dent at 20 10^6 fatigue cycles. This strict criteria is focused on the apparition of damage, and not on its evolution to generalised micro pitting or deep spalling.

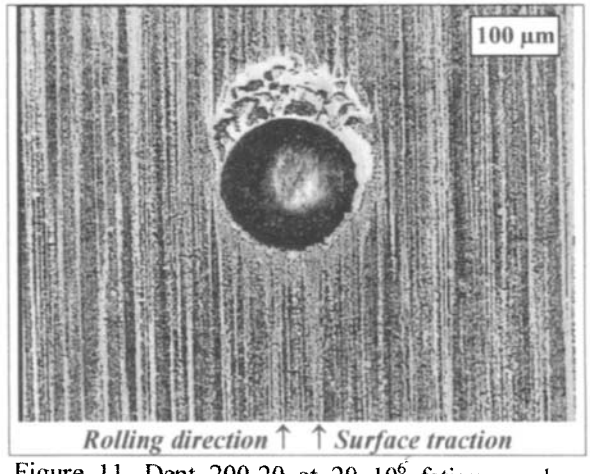


Figure 11. Dent 200-20 at 20 10^6 fatigue cycles. Operating condition ES3.

Results are presented on a table for each operating conditions. Each cell of the table corresponds to the mean behaviour of several (at least three) identical dents, with W_i on the line and R_p on the column. Dents that have been harmful for these conditions are represented by a shaded cell. A question mark means that the behaviour of the indent has not been evaluated. When the cell is blank, the dent is not harmful.

Influence of material. First, a comparison has been led between M50 and nitrided 32CrMoV13 steel. The only difference between operating conditions C2 and ES7 (tables 5 and 6) is the material of the dented disc. The comparison of these two operating conditions show that dents 400-20, 600-20 and 600-30 have not damaged nitrided 32CrMoV13 steel, but have been harmful for M50, showing then that nitrided 32CrMoV13 steel has a better behaviour than M50 as far as dent initiated RCF is concerned.

Table 5. Experimental results for conditions C2

C2	3 kg	5 kg	10 kg	20 kg	30 kg	50 kg
200 µm	?	?				
400 µm						
600 µm	?	?	?			

Table 6. Experimental results for conditions ES7

ES7	3 kg	5 kg	10 kg	20 kg	30 kg	50 kg
200 µm						
400 µm						
600 µm						

Influence of sliding. In pure rolling, every operating conditions leads to the same result, illustrated here by the case ES6, shown in table 7. The presence of sliding (cases ES8, ES7, ES4 and ES3) has changed the harmfulness of four dents (400-20, 400-30, 600-30 and 600-50). For all operating conditions, 11 cells have changed status in presence of sliding.

Influence of hoop stress. When hoop stress of 115 MPa exists, two dents became harmful (400-20, 600-30; ES7-ES8). For all operating conditions, 3 cells have changed status in presence of hoop stress.

Table 7. Experimental results for conditions ES6

ES6	3 kg	5 kg	10 kg	20 kg	30 kg	50 kg
200 µm						
400 µm						?
600 µm						

Influence of load. When the load is increased from 2.5 to 3 GPa, only one dent became harmful (600-30 ; ES4-ES8), and just one cell has changed status.

Table 8. Experimental results for conditions ES8

ES8	3 kg	5 kg	10 kg	20 kg	30 kg	50 kg
200 µm						
400 µm						
600 µm						

Table 9. Experimental results for conditions ES3

ES3	3 kg	5 kg	10 kg	20 kg	30 kg	50 kg
200 µm						
400 µm						
600 µm						

Table 10. Experimental results for conditions ES4

ES4	3 kg	5 kg	10 kg	20 kg	30 kg	50 kg
200 µm						
400 µm						
600 µm						

First, these experiments show that the load has little influence upon the dent initiated RCF. This remark must be limited to the load range encountered in the aeronautic domain (2 – 3 GPa). On one hand, if the load becomes higher, other phenomena might be activated, and on another hand, if no load is applied, no damage is likely to occur. This remark is also restrained to the nucleation of damage, since crack propagation depends upon the load. Second, they show that hoop stress of 115 MPa has little influence on dent initiated RCF. Last, they clearly show the major influence of sliding on dent initiated RCF and that an application with sliding, even low, is a dangerous application as far as dent initiated RCF is concerned.

4. RCF NUMERICAL STUDY

The RCF endurance limit, Hl, is defined as the maximum contact pressure that can undergo a contact without overtaking the micro-yield stress of the material [5]. In the experiment conducted in this study, the endurance limit of nitrided 32CrMoV13 steel in the smooth case, that depends upon the contact equivalent radius, is equal to 2.65 GPa [11]. When a surface is dented, pressure peaks arise around the dent which generate local stress concentration. The maximum contact pressure that can support the dented contact without overtaking

the micro yield shear stress locally around the dent is noted H1I. It is the endurance limit focused on the dent region. The endurance limit is of course the minimum value of H1 and H1I.
To determine this endurance limit, it is necessary to know the micro yield shear stress around the dent and the stress undergone during the over-rolling of the dent. Before indentation, residual stresses and micro yield shear stress profile of nitrided 32CrMoV13 steel have been determined (see section 2). During indentation, plastic deformation occurs, thus hardening the material and introducing residual stresses. Consequently a dent should not be considered as a simple geometric defect. The knowledge of the material hardening and residual stress state around the dent before over-rolling requires first the simulation of the indentation process.

The indentation process has been simulated using ABAQUS commercial FE code. The mesh used for this simulation is presented in figure 12. Since the punch is spherical, the problem is axisymmetric. The elastic behaviour of the diamond is represented by E=1140 GPa and ν=0.07. The friction coefficient between the punch and the specimen is taken to 0.1, and the plasticity model used for simulation is isotropic. The measured geometry of dent 400-20 is compared with the result of the simulation on figure 13. The agreement between both results validates the simulation of the indentation process.

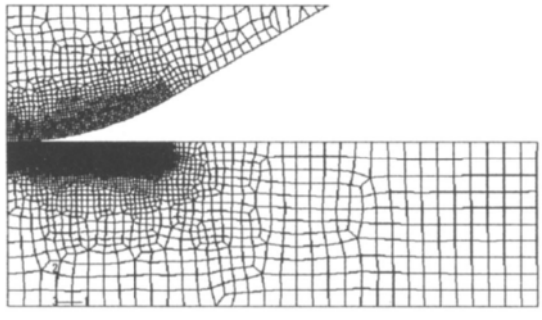


Figure 12. FE model used for indentation simulation.

The simulation of the dent over-rolling is a complex three-dimensional contact problem, that requires a mesh fine enough to describe the geometry of the dent and large enough to cover the area of contact. To solve this problem, a three dimensional semi-analytical contact code presented by Jacq et al. [10,12] is used. This code can solve 3D normal elastic plastic contact that respects Hertz hypotheses, i.e. when the dimension of the contact zone is small compared to the radii of curvature of contacting bodies, as it is the case in the present application (0.5 mm for 15 mm).

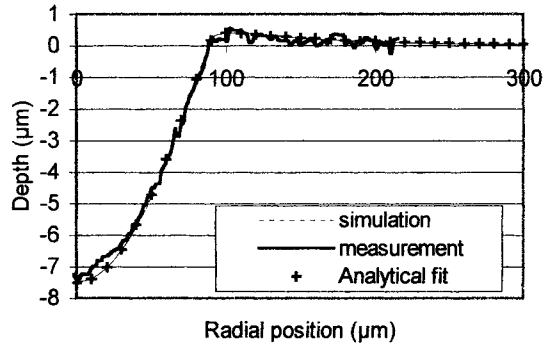


Figure 13. Simulated and measured 400-20 dent profile. Comparison with analytical shape function.

The dent is introduced in the contact code as an analytical shape function of five parameters (2), that are determined to match at best the result of the FE simulation. In this function, h_p is the depth of the dent, R its radius and r is the radial position. I_1, I_2 and I_3 are adjustable parameters that enables the reproduction of the shoulder height and slope (figure 13). The presence of the dent in the contact generates high pressure peaks, as shown for dent 400-20 and 600-20 in figure 14, under a load of 3.5 GPa.

$$\text{if } r<R \quad h(r) = h_p \cos\left(\frac{2\pi r}{4R}\right)\exp\left[\frac{r^2}{R^2}\ln\left(\frac{2R(I_1-I_2)}{\pi h_p}\right)\right]$$

$$\text{else } h(r) = I_1[\exp(-I_2(r-R)) - \exp(I_3(r-R))] \quad (2)$$

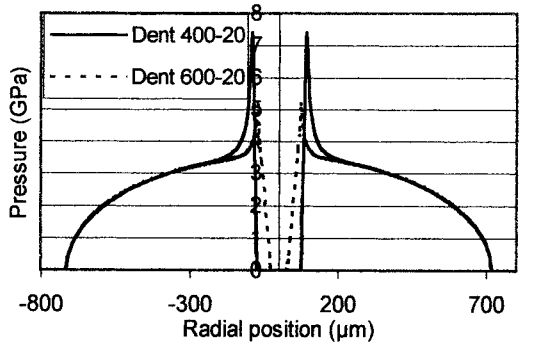


Figure 14. Contact pressure for two different dents.

Stresses due to this contact pressure are calculated by the semi-analytical contact code. Since the

determination of the endurance limit requires only an elastic simulation of the over-rolling of the dent, the contact pressure can be calculated independently from residual stresses and hardening parameters due to indentation. Contact stresses can then be added to residual stresses resulting from the nitriding and the indentation process. Von Mises stress is then calculated and compared to the local micro yield shear stress, modified by the hardening that occurred during the indentation process. The maximum load that can be applied during the over-rolling of the dent without over-taking the hardened micro yield shear stress is the endurance limit H1I. It is expressed as the Hertz pressure of the equivalent smooth contact case. The contact code does not take tangential effects into account. Therefore, the endurance limit H1I determined numerically is restricted here to pure rolling conditions. In the present paper, endurance limit H1I has been normalized by a Hertz pressure higher than Hertz pressures normally found in aeronautic applications. It is then noted H1Id. Results are given in table 11. H1Id has been investigated from 0.1 to 1. The lower bound has been chosen because the contact zone is becoming smaller than the dent. Dents can be classified in three groups, low (L), medium (M) and high (H). Dents of group L have a very low H1Id (<0.1). They undergo plastic deformation even for light loads and consequently are dangerous for most applications. Dents of group H have a very high H1Id. They are not dangerous in most aeronautic pure rolling applications. Last, group M is constituted of dents whose endurance limit H1Id is between 0.1 and 1. Their harmfulness is application dependent.

Table 11. Dimensionless H1Id endurance limit in pure rolling conditions.

Dent	H1Id (GPa)	Dent	H1Id (GPa)
200-5	0.107	400-20	0.112
200-10	< 0.1	400-50	< 0.1
200-20	< 0.1	600-5	> 1
200-50	< 0.1	600-10	> 1
400-5	> 1	600-20	0.485
400-10	0.292	600-50	< 0.1

The consistency of these numerical results can be compared with experimental data from §4. Dents are sorted in figure 15 by increasing H1Id along the vertical axis. An arbitrary horizontal offset is applied for seek of clarity. Numerical determination of H1Id has been made for pure rolling conditions without hoop stress. In figure 15 the behaviour of dents for operating conditions ES1 is illustrated by different symbols. A circle has been assigned when the dent has been harmful, a star for safe dents, and a square for undetermined dent behaviour. Two major observations can be made. First, dents whose H1I is higher than the load undergoes during tests have not caused any damage. Second, all dents that have been harmful have H1I lower than the load applied during the test. Hence, with predictions given by H1I criteria, all harmful dents have been identified, and every dents predicted as safe were effectively safe. So H1I seems to be an interesting definition of the local endurance limit.

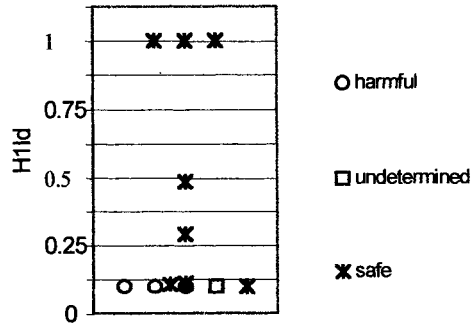


Figure 15. H1Id and ES1 experimental behaviour of dents.

The endurance limit H1Id has also been determined for dents created on M50. The elastic plastic behaviour of M50 has been determined by Lamagnère et al. [5]. Results are compared for both steels in figure 16 and 17, for dents made with the indenter of radii R_p=200 and 400 µm at different loads W_i. For all dents, H1Id is equal or higher for nitrided 32CrMoV13 than for M50, as it is observed experimentally. In addition it can be noted that R_p=400 µm indenter (figure 17) is more efficient to discriminate materials than R_p=200 µm indenter (figure 16).

The influence of residual stresses and hardening introduced during the indentation process can also be studied. The endurance limit H1Id has been determined for dents made on 32CrMoV13 steel surface in pure rolling while neglecting residual stresses and hardening due to the indentation process. This endurance limit, noted as

"geometry" in figure 18, based only on dent geometry and residual stresses due to nitriding has been compared to the real endurance limit, noted as "full" in figure 18. It can then be seen that residual stresses and hardening have an effect upon the endurance limit, and that they are not always protective, as it is often believed.

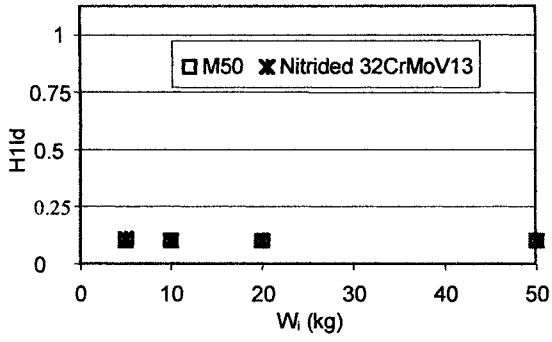


Figure 16. Comparison of H1Id for nitrided 32CrMoV13 and M50. R_p=200 µm

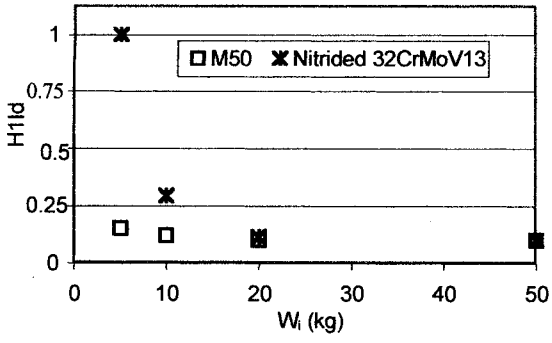


Figure 17. Comparison of H1Id for nitrided 32CrMoV13 and M50. R_p=400 µm

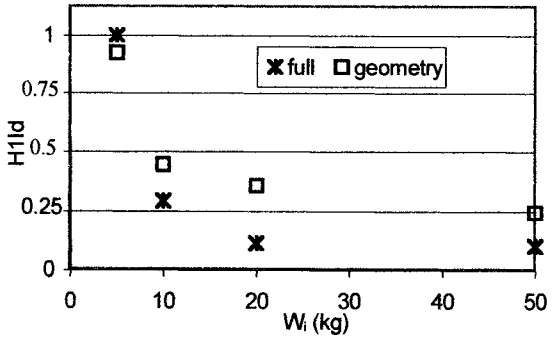


Figure 18. Influence of hardening and residual stresses due to indentation on H1Id.

For artificial dents made with spherical diamond indenter, the variation of endurance limit H1Id with the average slope of the dent is monotonic, as can be seen in figure 19. If the elastic unloading is neglected during indentation and the diamond indenter considered as rigid, the dent is a portion of a sphere, and its average slope can be expressed as a function of R, the radius of the dent and R_p, the radius of the indenter (3). During the indentation process, the contact pressure can be approximated as a constant, P_{ave}, because of the load redistribution due to plastic deformation. The radius of the dent can then be related to the applied load W (4). Because of the asymptotic evolution of stress-strain curve, and because of large plastic deformation undergone during indentation, P_{ave} can be considered as a constant for a given material. Therefore, for a given material, the average slope of the dent evolves linearly with I_R, ratio between the square root of the indentation load over the radius of the indenter (5).

$$h_p = \frac{R^2}{2R_p} \tag{3}$$

$$R = \left(\frac{W}{\pi P_{ave}}\right)^{\frac{1}{2}} \tag{4}$$

$$p_m = \frac{h_p}{R} \approx \frac{W^{\frac{1}{2}}}{R_p} \tag{5}$$

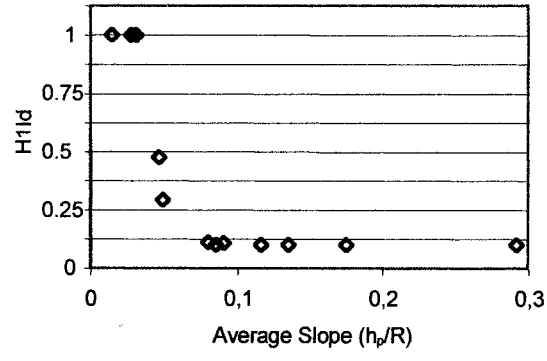


Figure 19. Monotonic evolution of H1Id with the average slope of dents.

Finally dents can be sorted by increasing I_R. For a given operating condition (ES8 in figure 20), experimental behaviour of dents is reported. Two domains, harmful and safe, appear. The transition between this two domains is noted I_{Rlimit}. It depends

upon operating conditions. For given operating conditions, it became an indicator of the quality of the tested material, as far as dent initiated RCF is concerned.

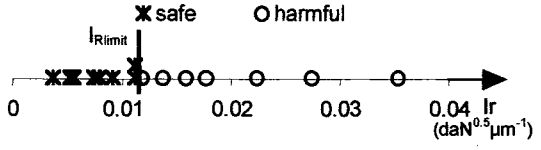


Figure 20. I_R of dents and their experimental behaviour for condition ES8.

5. Conclusion

The understanding and the simulation of the whole process that leads to dent initiated RCF requires the knowledge of the initial material mechanical characteristics, and of their evolution during the dent formation before the over-rolling of the dent. The use of the results of the nano-indentation local mechanical test enables the determination of the initial mechanical behaviour profile of nitrided 32CrMoV13 steel.

The indentation process and the over-rolling of the dent has then been simulated. From these simulations, a physically based endurance limit criteria focused on the dent, noted H1I, has been calculated for various dents. The comparison of experimental data with numerical results shows that this criteria can be used as definition of the local endurance limit.

This criteria has been calculated for identically made dents on nitrided 32CrMoV13 steel and M50 steel. Numerical results show a significant improvement of nitrided 32CrMoV13 steel to dent initiated rolling contact fatigue over M50, confirmed experimentally.

For a given operating condition, and for dents made with a diamond spherical tip, the limit between harmful and non harmful dents can be seen in terms of I_{Rlimit}. This factor depends upon operating conditions and material. For fixed operating conditions, this limit could become a representation of the material response for dent initiated RCF.

References.

[1] T.E. Tallian, Failure atlas for Hertz contact machine elements, 2nd ed. New York. ASME Press (1999), 486.

[2] D. Nélias, F. Ville. ASME Journal of Tribology vol. 122 n°1 (2000), 55-64.

[3] P.M.E. Cann, J.C. Hamer, R.S. Sayles, H.A. Spikes, E. Ioannides. Proc. of the 22nd Leeds-Lyon Symposium on Tribology, Elsevier Tribology Series vol. 31 (1996), 127-134.

[4] D. Ghiglione, C. Leroux, C. Tournier. Techniques de l'ingénieur, vol. MD M1222 (1996), 1-44.

[5] P. Lamagnère, R. Fougères, G. Lormand, A. Vincent, D. Girodin, G. Dudragne, F. Vergne. ASME Journal of Tribology vol. 120 n°3 (1998), 421-426.

[6] L. Flamand, D. Berthe, M. Godet. ASME Journal of Mechanical Design vol. 103 n°1 (1981), 204-209.

[7] G. Xu, F. Sadeghi, M. Hoeprich. Tribology Transactions vol. 40 n°4 (1997), 613-620.

[8] H. El Ghazal, G. Lormand, A. Hamel, D. Girodin, A. Vincent. Materials Science and Engineering A vol. 303 n°1 (2001), 110-119.

[9] C. Jacq, G. Lormand, D. Nélias, D. Girodin, A. Vincent. Materials Science and Engineering A, in press, (2002), 9.

[10] C. Jacq, D. Nélias, G. Lormand, D. Girodin. ASME Journal of Tribology vol. 124 (2002), 1-15.

[11] A. Vincent, R. Fougères, G. Lormand, G. Dudragne, D. Girodin. Bearing Steel Technology, ASTM STP 1419, J. M. Beswick, Ed. American Society for Testing Materials, West Conshohocken, PA, 2002, in press.

[12] P. Sainsot, C. Jacq, D. Nélias. Computer Modeling in Engineering & Sciences, in press, (2002), 9.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

EFFECT OF Hf ADDITION ON THE TRIBOLOGICAL PERFORMANCE OF ZrN COATINGS

Erdem ATAR[a], Harun MINDIVAN[b], Hüseyin CIMENOGLU[b], E. Sabri KAYALI[b]

[a] Gebze Institute of Technology, Material Science and Engineering Department, 41400, Gebze Kocaeli, Turkey

[b] Istanbul Technical University, Metallurgy and Materials Engineering Department, 34469, Maslak Istanbul, Turkey

ABSTRACT

In this study, tribological performance of 2 μm thick ZrN and Zr(21% Hf)N coatings, deposited on hardened AISI D2 cold work tool steel by Physical Vapour Deposition technique were compared. Alloying of ZrN with Hf did not cause significant hardness increase, whilst increased the mean surface roughness. According to the results of reciprocating wear tests carried out by utilising ceramic (Al_2O_3) and steel (AISI 52100) balls, the wear resistance of the ZrN coating increases with the addition of Hf.

1. INTRODUCTION

Protection of surfaces by thin, hard ceramic coatings is widely used to improve the life and tribological performance of metal cutting and forming tools. One of the commercially used and successfully applied thin ceramic coating technique is Physical Vapor Deposition (PVD) [1-6].

Most of the PVD coatings consist of nitrides and carbides. In current industrial applications, TiN coatings are widely used. Since the wear of TiN coatings change dramatically with adjustment of parameters such as contact load, sliding speed, contact geometry and humidity, for certain specific applications alternative coatings such as CrN, ZrN and HfN, exhibit superior machining or tribological performance than that of traditional TiN coatings [7-18]. The goal of the development of new alternative coatings with different compositions and structures is to eliminate the weak points of individual coatings

In this study we investigated the tribological performance of ZrN and a new Zr(21% Hf)N coatings deposited on hardened AISI D2 cold work tool steel by the PVD technique.

2. EXPERIMENTAL

Quenched and tempered AISI D2 cold work tool steel discs (28 mm diameter and 10 mm height) were used as substrate materials. The hardness of the discs were 59 HRC. ZrN and Zr(21% Hf)N coatings were deposited on substrates by a PVD technique after final polishing with a 1μm diamond paste. The process parameters of PVD technique are listed in Table 1.

Table 1
Deposition conditions of ZrN and Zr(%21Hf,N) coatings.

Coating	Bias (V)	$Px10^{-3}$ (Torr)	T_{first} (°C)	T_{last} (°C)	Time (sec)
ZrN	-200	2.5	243	302	1800
Zr(%21Hf,N)	-200	2.5	305	331	3600

P: N_2 pressure (Torr)
T_{first}: Coating start temperature
T_{last}: Coating finish temperature
Time: Whole coating time

Characterizations of the coatings were made by thickness, roughness, and hardness measurements. Thickness of the coatings was measured by Wirttz Buehler Calotest tester, which uses the ball crater method. Surface roughness of the coatings was determined by an optical profilometer, Perthometer S8P. Hardness measurements were conducted on a Fischercope HP XY PROG ultramicrohardness tester using Vickers type indenter under the loads of 10mN, 15 mN and 25mN, which caused the indenter to penetrate less than 10 % of the coating thickness.

Dry sliding wear performance of ZrN and Zr(21% Hf)N coatings were examined with a reciprocating wear tester in ambient atmosphere conditions (20±1 °C and 50±5 %RH). Wear tests were carried out by applying a normal load of 1.5 N to the coated samples with 10 mm diameter ceramic (Al_2O_3) and steel (AISI 52100-55 HRC) balls whose surface roughness (Ra) values were 0.2μm and 0.1μm, respectively. During the tests, sliding speed was 0.02 m/s for the total sliding distance of 120 m.

After the wear test, the wear tracks developed on the coatings were detected by profilometer. Wear of coatings and the counterface balls were examined by using Jeol 5410 Scanning Electron Microscope (SEM) equipped with Electron Diffraction Spectrometer (EDS) and a Zeiss light optical microscope.

3. RESULT AND DISCUSSION

The characteristics of the ZrN and Zr(21% Hf)N coatings are listed in Table 2, as thickness, surface roughness and hardness. Both of the examined coatings have the same thickness. According to the surface roughness measurements in terms of Ra, which define the arithmetic mean of departure of a surface profile from a mean line, addition of 21 % Hf increased the surface roughness of ZrN coating. Ultramicrohardness measurements carried out under three different loads showed that, ZrN and Zr(21% Hf)N coatings have almost the same HV value.

Figure 1 and Figure 2 depict the low magnification SEM micrographs of the wear tracks produced on the ZrN and Zr(21% Hf)N coatings by Al_2O_3 and steel balls, respectively. The 2D - profile images of these wear tracks are given in Figure 3 and Figure 4. Wider and deeper wear tracks are developed on ZrN coating than those of Zr(21% Hf)N coating. The average depths of wear tracks produced by Al_2O_3 ball are about 2 μm for both coatings, which are approximately the coating thickness. The depths of the wear tracks formed by steel ball were measured as 1.6 μm for ZrN coating and 1.3 μm for Zr(21% Hf)N coating, which are less than the thickness of the examined coatings. When compared to Al_2O_3 ball, steel ball produced wider but shallower wear tracks on both coatings.

The area of the wear tracks,calculated from the 2D - profile images (Figure 3 and Figure 4) are presented in Figure 5. Alloying of ZrN with Hf resulted with formation of smaller wear track area and therefore increment in the wear resistance. From this point of view, Zr(21% Hf)N coating exhibited higher resistance to wear than ZrN coating. When compared to Zr(21%Hf)N coating, the areas of the wear tracks produced on ZrN coating by Al_2O_3 and steel balls are about 25 % and 90 % larger, respectively.

High magnification SEM micrographs of the wear tracks produced by Al_2O_3 and steel balls are given in Figure 6 and Figure 7, respectively. As mentioned previously the depth of the wear tracks formed by Al_2O_3 ball is approximately equal to the coating thickness (Figure 3) and SEM examination revealed that, wear progress by gradual removal of coatings. EDS analysis of the wear track (Figure 6) showed Fe and Cr contributions from the AISI D2 substrate. There was no Al element detected. Presence of oxygen indicates formation of iron oxide from the substrate and zirconium oxide from the coating during testing.

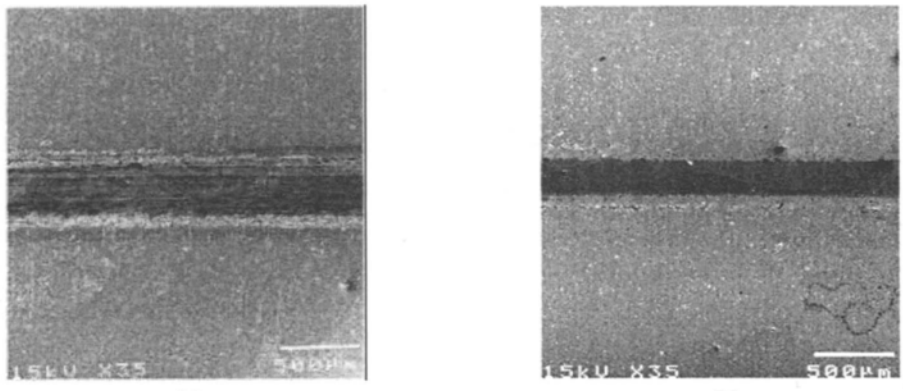
On the other hand, rubbing of steel ball caused adhesive and oxidative damage on coatings (Figure 7). The depth of the wear tracks formed by steel ball is less than the thickness of the coatings (Figure 4). EDS analysis show presence of Fe,Cr and O rich islands (eg. iron oxides) in the wear tracks. These oxide island were formed due to mass transfer from steel ball during testing.

During reciprocating wear tests balls were also subjected to wear, as shown in Figure 8 and Figure 9

Table 2
Surface properties of the ZrN and Zr(%21 Hf)N coatings.

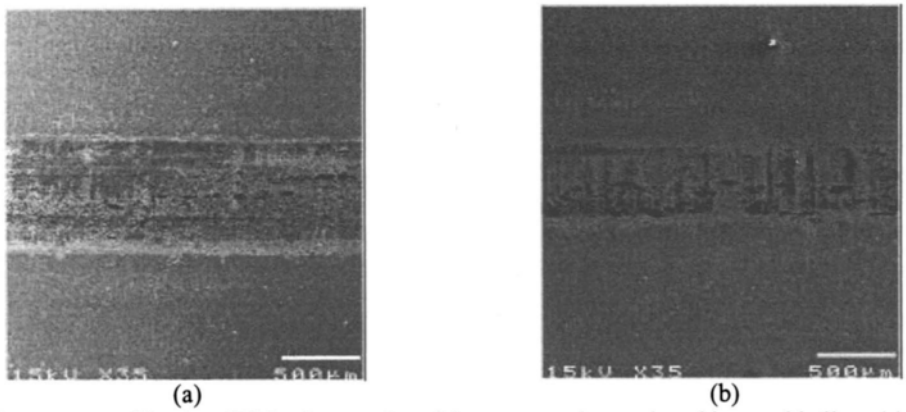
Coating	Thickness (μm)	Roughness (Ra, μm)	Hardness (HV, GPa)		
			10 mN	15mN	25 mN
ZrN	2±0,01	0.2±0,01	41±5	36±4	38±3
Zr(%21 Hf)N	2±0,01	0.3±0,03	40±4	35±3	40±3

The areas of the wear scars of the balls measured from the optical micrographs, are presented in Figure 10. The size of the wear scars formed on the contact surfaces of balls is proportional to the size of the wear tracks developed on the coatings, which indicate that the wear of the ball was controlled by the size of contact areas. It should be mentioned that steel balls were subjected to much more heavy wear than Al_2O_3 balls during reciprocating wear testing of ZrN and Zr(%21 Hf)N coatings.



(a) (b)

Figure 1. Low magnification SEM micrographs of the wear tracks produced by Al_2O_3 ball on (a) ZrN and (b) Zr(21% Hf)N coatings.



(a) (b)

Figure 2. Low magnification SEM micrographs of the wear tracks produced by steel ball on (a) ZrN and (b) Zr(21% Hf)N coatings.

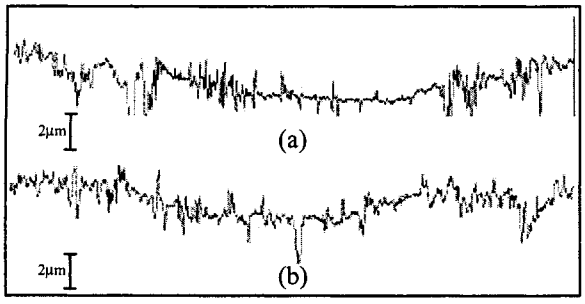


Figure 3. 2D - profile images of the wear tracks produced on (a) ZrN and (b) Zr(21% Hf)N coatings by Al_2O_3 ball.

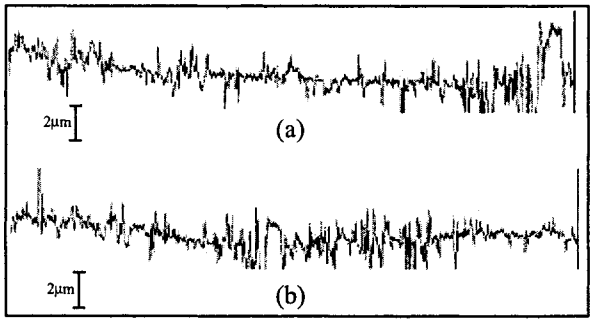


Figure 4. 2D - profile images of the wear tracks produced on (a) ZrN and (b) Zr(21% Hf)N coatings by steel ball.

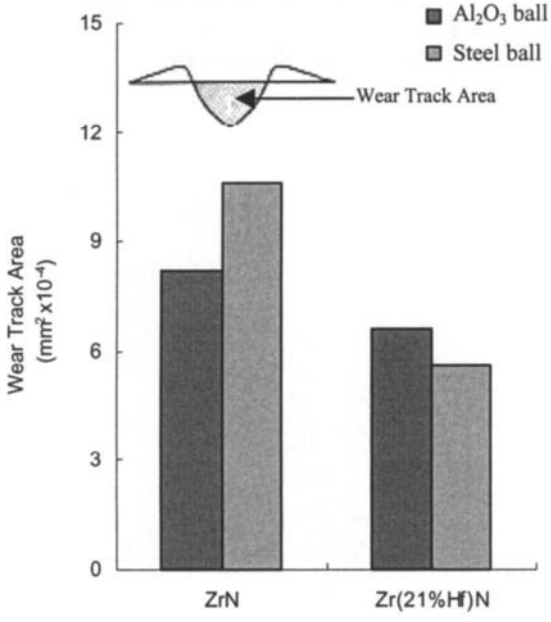
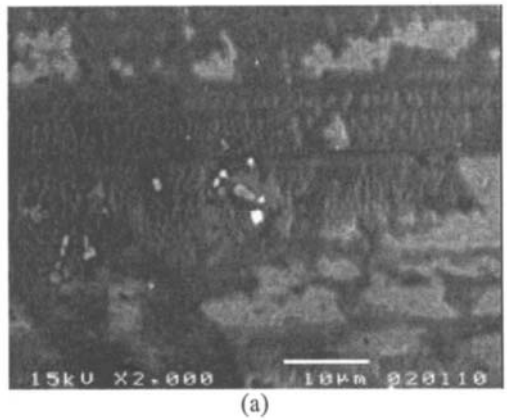


Figure 5. Wear track areas of the coatings against Al_2O_3 and steel balls.

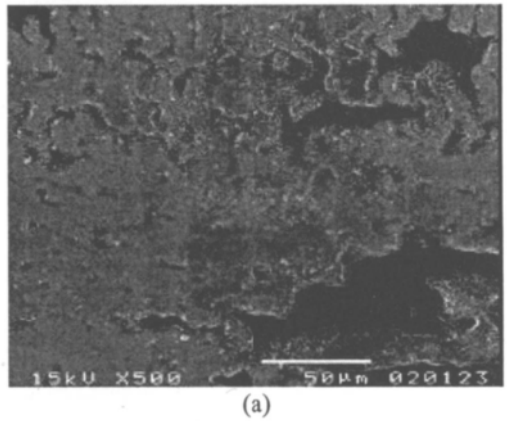


(a)

Element (wt.%)	Dark Grey Region	Light Grey Region
Zr	47,88	49,92
Cr	4,26	3,14
Fe	22,53	19,04
O	25,33	27,91

(b)

Figure 6. (a) High magnification SEM and (b) EDS analysis result of the wear track center of ZrN coating, produced by Al_2O_3 ball.



(a)

Element (wt.%)	Dark Grey Region	Light Grey Region
Zr	-	22,54
Hf	-	13,15
Cr	3,04	15,48
Fe	76,68	45,02
O	20,28	3,82

(b)

Figure 7. (a) High magnification SEM and (b) EDS analysis result of the wear track center of Zr(21% Hf)N coating, produced by steel ball.

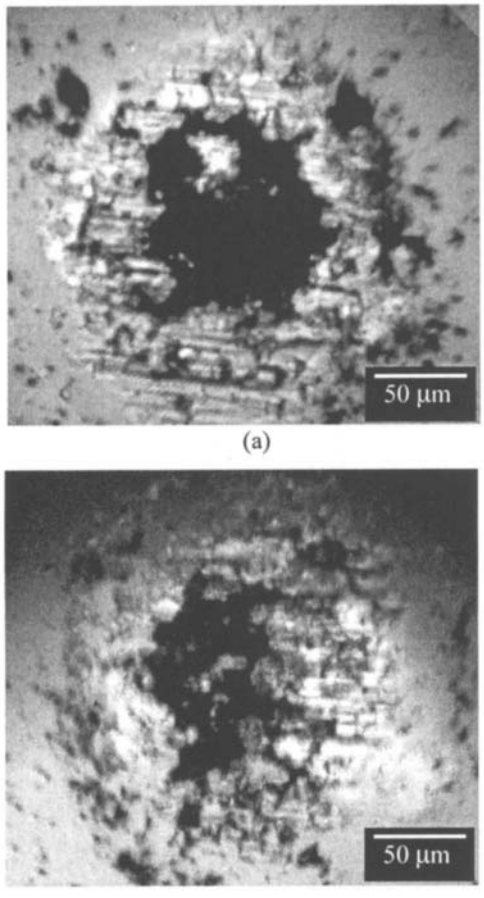


Figure 8. Wear scars formed on Al_2O_3 balls during test of (a) ZrN and (b) Zr(21% Hf)N coatings.

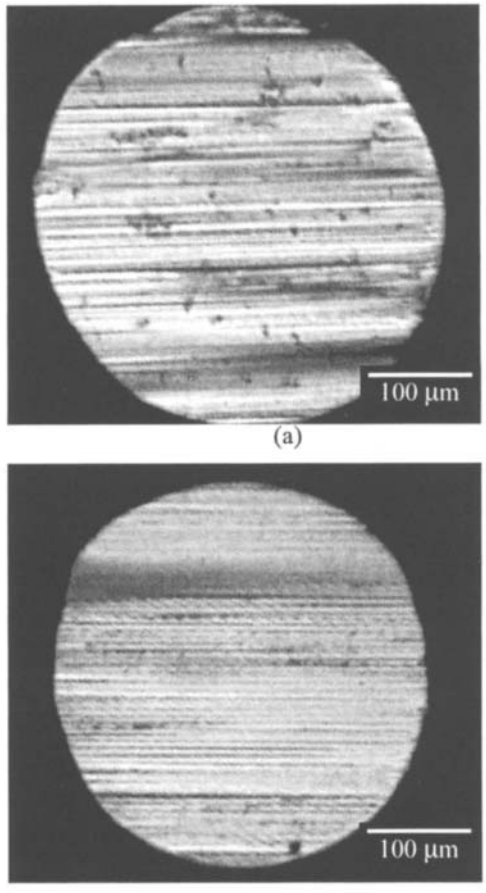


Figure 9. Wear scars formed on steel balls during test of (a) ZrN and (b) Zr(21% Hf)N coatings.

4. CONCLUSIONS

Following conclusions can be drawn according to the results of reciprocating wear tests conducted on ZrN and Zr(21% Hf)N coatings under dry sliding conditions by utilising Al_2O_3 and steel balls.

1. Addition of Hf caused a pronounced increase in the wear resistance of ZrN coatings. When compared to a Zr(21%Hf)N coating, the areas of the wear tracks produced on ZrN coating by Al_2O_3 and steel balls are namely about 25 % and 90 % larger, respectively.

2. During wear testing, Al_2O_3 ball removed the coatings and encouraged oxidation of the coating and the substrate. Wear of the coatings against steel ball is progressed by mass transfer from the ball and oxidation.

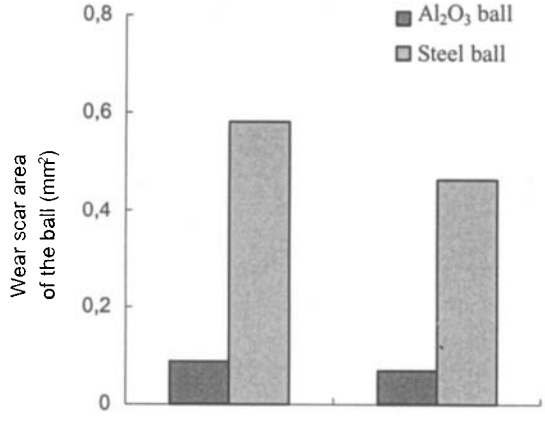


Figure 10. Wear scar areas of the balls.

3. Wear of counterface balls significantly depend on the type of coating being tested. Wear scar areas developed on Al_2O_3 and steel ball are 20% smaller when they rub on Zr(21% Hf)N coating instead of ZrN coating. During testing steel ball was subjected to much more heavy wear than Al_2O_3 ball.

REFERENCES

1. W. König, R. Fritsch, D. Kammermeier, Surface and Coatings Technology 49 (1991) 316-324.
2. K.H. Habig, G. Meier zu Köcker, Surface and Coatings Technology 62 (1993) 428-437.
3. O. A. Johansen, J. H. Dontje, and R. D. L. Zenner, 14 th Int. Conf. on Metallurgical Coatings, San Diego, California, March 23-27, 1987.
4. G. R. Fenske, Tribology Transactions, Vol. 32 (1989), 3, 339-345.
5. E. J. Bienk, H. Reitz, N. J. Mikkelsen, Int. Conf. on Metallurgical Coatings and Thin Films, App. 24-28, San Diego 1995.
6. H. Holleck, J. Vac. Sci. Technol. A 4 (6), Nov/Dec 1986,2661-2669.
   E. Lugscheider, O. Knotek, H. Zimmermann, S. Hellmann, Surface and Coatings Technology, 116-119 (1999) 239-243.
8. Y.L. Su, S.H. Yao, Z.L. Leu, C.S. Wei, C.T. Wu, Wear, 213 (1997) 165-174.
9. Y.M. Zhou, R.Asaki, K. Higashi, W.H. Soe, R. Yamamoto, Surface and Coatings Technology 130 (2000) 9-14.
10. C. Brugnoni, F. Lanza, G. Macchi, R.Müller, E.Parnisari, M.F. Stroosnijder, J. Vinhas, Surface and Coatings Technology 100-101 (1998) 23-26.
11. D.A. Glocker, S. Shah, Handbook of thin Film Process Technology, Chapters X1.5 and X1.6, IOP Publishing Ltd., Bristol and Philodelphia, 1995.
12. H. Çimenoglu, Wear, 210 (1997) 204-210.
13. Y. Sun, B. Li, D. Yang, T. Wang, Y. Sasaki, K. Ishii, Wear, 215 (1998) 232-236.
14. S. K. Biswas, Wear 245 178-189.
15. M. Z. Huq, C. P. Celis, Wear 212 (1997) 151-159.
16. W.M. Rainforth, Wear, 245 (2000) 162-177.
17. J. L. He, Y. H. Lin, K. C. Chen, Wear 208 (1997)36-41.
18. H.Ahlatci, M.C. Türküz, H. Çimenoglu, M. Ürgen, E.S. Kayalý, Metall 56 jahrgang 4/2002.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribological Properties of TiAlSiN Film Coated by AIP Equipment onto SKD61 Substrate

Ri-ichi Murakami[a] , Yuya Matsuda[b] and Daisuke Yonekura[a]

[a]Dept. of Mech. Eng., The University of Tokushima
2-1 Minami-josanjima-cho, Tokushima, Japan
[b]Graduate student, The University of Tokushima
2-1 Minami-josanjima-cho, Tokushima, Japan

TiAlSiN film was coated by Arc Ion Plating onto high-speed tool steel, SKD61. The wear behaviour of film against bearing steel, SUJ2, ball was investigated using the ball on disk testing machine under the condition of load of 10N, sliding distance of 5000m, and sliding speed of 250mm/s. Surface roughness of film was changed by hand polishing into four types of surface roughness. Tribological performance for the SKD61 specimen coated by TiAlSiN film was compared with that of uncoated specimen. Wear track of film was observed using a scanning electron microscope and optical microscope. The wear resistance of SKD61 was significantly improved by coating of TiAlSiN film, regardless of surface roughness. However, the wear rate of ball was greater for the unpolished film than for the polished film because many droplets existed on the unpolished film.

1. INTRODUCTION

There are many ways to deposit hard thin film onto metal substrate. Arc Ion Plating, AIP, is the equipment, which can deposit varies kinds of ceramic film onto tool under high coating rate [1,2]. When the ceramic film is coated using AIP equipment, a lot of micro particles such as droplet are usually observed on the film surface. The droplets are the particles, which are sputtered out and attached on the surface of ceramic film. The particle is quite large size and is a different crystal from the regular film [3]. Thus, the droplet results in the coarse surface roughness and should affect on the wear behaviour of ceramic film and against material.

In general, Al and Ti produce easily the intermetallic compound and its hardness increases without decrease of toughness [4-6]. Therefore, the TiAlN and TiAlSiN films must show an excellent wear resistance. In this study, TiAlSiN film was deposited by AIP equipment on high-speed tool steel, SKD61. The wear tests were carried out using ball-on- disk wear testing machine against bearing steel ball, SUJ2. When the TiAlSiN film was changed four kinds of surface roughness from R_a=0.17μm to R_a=0.89μm, the effect of droplet on wear behaviour was discussed.

2. EXPERIMENTAL

TiAlSiN film was coated by AIP equipment onto

*Corresponding author, Tel.: +81-88-656-7392
E-mail address: murakami@me.tokushima-u.ac.jp

high speed tool steel substrate, SKD61 (ϕ 20×10mm). Figure 1 shows the schematic diagram of AIP system. A target material was the sintered TiAlSi of which the ratio of Ti, Al and Si was 49 : 49 : 2 (atm%). The size of TiAlSi cathode was a diameter of 100mm. The substrate position didn't move so that the sample was always faced to the cathode to keep the constant distance of 200mm. After the chamber was vacuumed up to 4.0×10^{-3} Pa, the substrate was bombarded for 1 minute. After the bombarding, the chamber of AIP equipment was also vacuumed to 2.5×10^{-3} Pa. The TiAlSiN film was deposited onto the high-speed tool steel substrate, SKD61 having the chemical composition of C: 0.39, Cr: 5.13, Si: 0.97, Mn: 0.39, P: 0.017, S: 0.002, Mo: 1.22, V: 0.85 (wt%), and Vickers hardness of HV=420. The deposition time was 100 min, and the film thickness was about 40μm. The TiAlSiN film was analyzed by X-ray diffraction, fluorescence X-ray and energy dispersive X-ray equipment. Table 1 shows the coating condition of TiAlSiN film. The substrate was buff polished to the surface roughness of R_a=0.06μm. Then, the substrate was pre-cleaned ultrasonically in an acetone bath [6].

The TiAlSiN film surface was polished by the emery paper (#150 or #800) and the buff (Al_2O_3 powder of 0.06μm) to classify into four types of surface roughness; R_a=0.89μm, R_a=0.41μm, R_a=0.20μm and R_a=0.17μm. Figure 2 shows the topography of film surface for four kinds of surface roughness.

To estimate the tribological properties of TiAlSiN film against SUJ2 ball, the ball-on-disk wear test

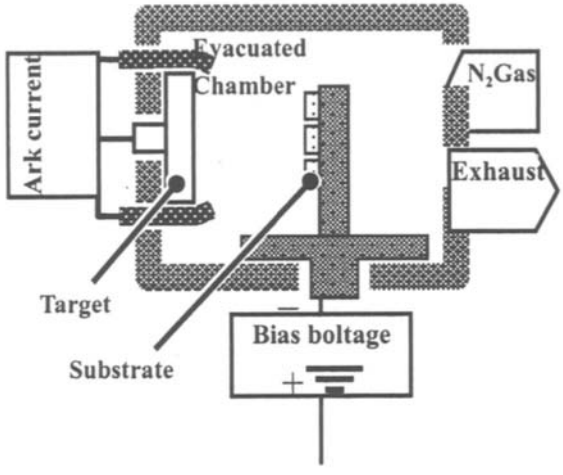


Figure 1. Schematic diagram of AIP equipment.

Table 1. Coating condition of TiAlSiN film in AIP.

	Bombard	Coating
Base pressure	4.0×10^{-3}Pa	2.5×10^{-3}Pa
Process gas (N_2) flow	53sccm	300sccm
Treatment time	1.0min	100min
Arc current	60A	200A
Vies voltage	700V	300V
Treatment pressure	0.51Pa	2.67Pa
Temperature of heater	773℃	573℃

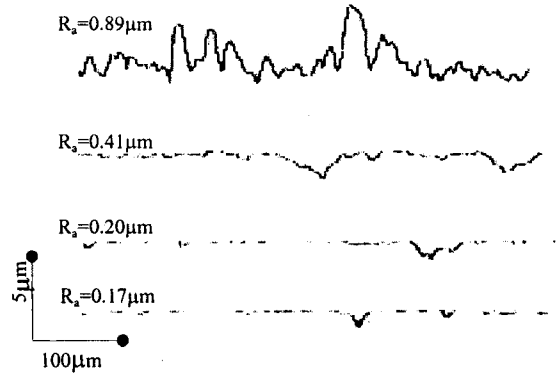


Figure 2. Surface topography of virgin and polished films

was done under the condition of sliding speed of 250mm/sec and vertical load of 10N until sliding distance of 5,000m [5,7]. Sample weight and coefficient of friction were measured for each the sliding distance of 100, 500, 1000, 2000, 3000 and 5000m. Mass loss was calculated by subtracting measured sample weight from original weight. Wear tracks of film and ball at the sliding distance of 5,000m were also observed by SEM.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Microstructure and surface morphology

Figure 3 shows the SEM images of virgin film and three polished films. For virgin film, many droplets, which widely distribute from small particle to quite large particle, are observed on the film surface. When the surface roughness of TiAlSiN film decreases, the number of droplet decreases. As shown in Figure 3, the droplet is seen as white circles on the film surface because the droplets can be polished. Also, the droplets are easily removed from the film surface when the film is polished by fine emery and the buff. Existence of droplet contributes to increase of surface roughness.

In the target material of TiAlSi, Si element contains only 2%. But in the measurement of fluorescence X-ray, Si element could not be detected from the TiAlSiN film. Because the melting point of Ti is greater than that of Al, the ratio of Ti to Al decreases from 1.8 in the target to 0.5 in the film. Figure 4 shows the XRD spectrum of TiAlSiN film. Because the peak of $Ti_7Al_5Si_{12}$ is seen in the XRD spectrum, Si exists in the TiAlSiN film as a compound of TiAlSi. In the fluorescence X-ray spectrum, it was difficult to detect Si element because of low Si content in the target material.

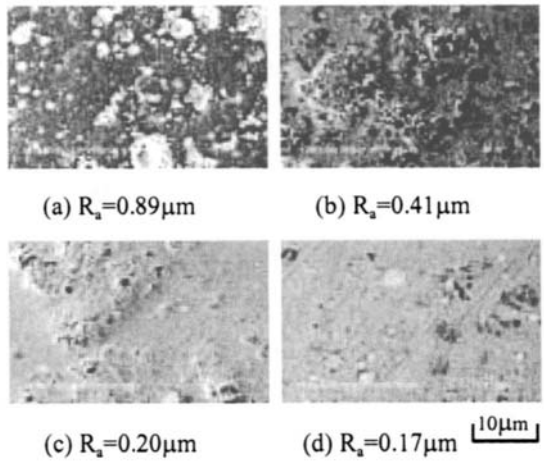


Figure 3. SEM images of TiAlSiN film surface.

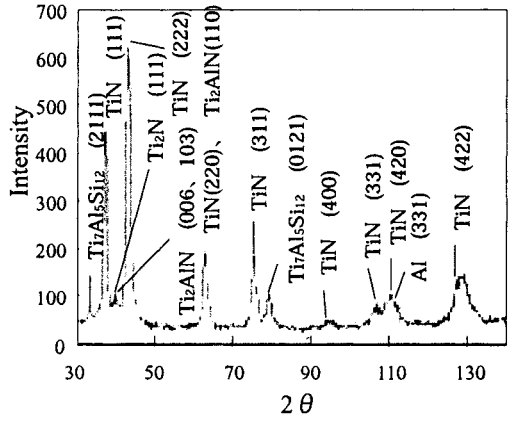


Figure 4. XRD spectrum of TiAlSiN film.

3.2 Friction and wear properties of TiAlSiN film

The wear test was carried out using the ball-on-disc tribometer for TiAlSiN film against SUJ2 ball under applied load of 10N. Figure 5 shows the comparison of the mass loss between the substrate and the films. It is clear that for the substrate, the

mass loss markedly increases with increasing the sliding distance. The severe wear for the substrate occurred because of the direct contact between the disk and the ball. The surface roughness is greater for the film than for the substrate as shown in Figure 3. However, the mass loss of the films is less than that of the substrate even though the surface roughness is greater for the film than for substrate (substrate roughness: R_a=0.06μm).

Figure 6 shows the wear results of TiAlSiN films. The negative mass loss means the increases of the specimen weight. The mass loss changes within only ±1mg even though the sliding distance increases till 5,000m, regardless of surface roughness. The mass loss means the change from unstable process to the stable process within the sliding distance of 1,000m. When the sliding distance is over 1,000m, the mass loss is small change. However, the mass loss for R_a=0.20μm is great among the four samples. The detail of such a wear behaiviour is not now clear.

This can be seen in Figure 7, which shows the relation between the wear coefficient and the sliding distance. The wear coefficient drastically decreases with increasing the sliding distance in the initial wear region (<1,000m) and is almost constant in the steady wear region (>1,000m). Thus, it assumes that the difference of surface roughness of film hardly affects the mass loss for steady wear process.

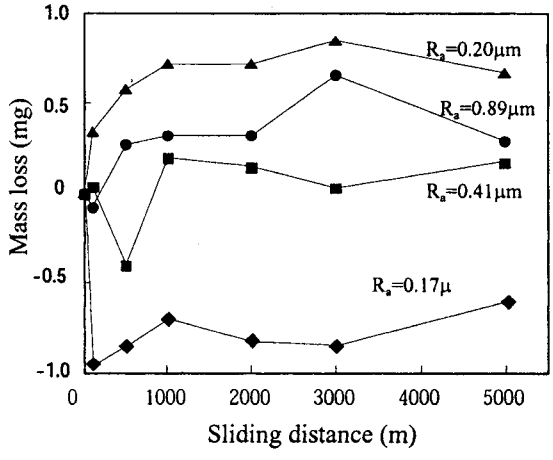


Figure 6. Relation between mass loss as a function of sliding destance for TiAlSiN film.

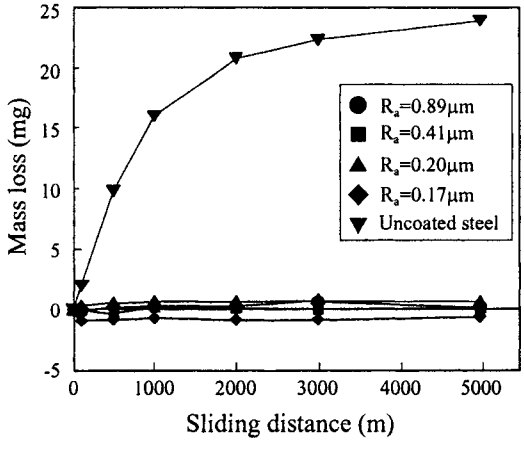


Figure 5. Comparison of mass loss between TiAlSiN film and substrate.

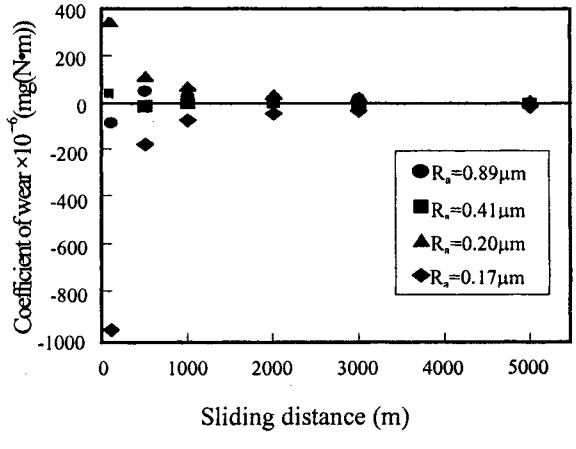


Figure 7. Relation between wear coefficient and sliding distance for TiAlSiN film.

We measured the wear track by using the surface roughness testing equipment. Figure 8 shows the topography of wear track for the film at the sliding distance of 5,000m. The surface roughness of the wear track is greater than that of original film surface. This indicates that in the wear track, a main wear process would be mass transfer from the ball to the film surface. The wear track of the film was also observed by SEM. The scattered stripe pattern, which seems the adhesive wear, was observed in the wear track, as shown in Figure 9 (a). The scattered stripe pattern area was analyzed by the energy dispersion X-ray equipment. Figure 9 shows the plane analysis results of wear track. It is clear that in the scattered stripe pattern area, Fe element exists. This suggests that the SUJ2 ball adhered toward the TiAlSiN film.

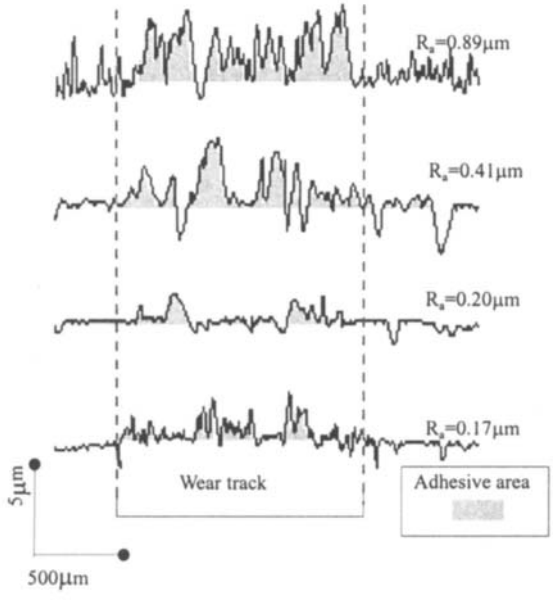


Figure 8. Morphology of wear track for four surface roughness films at the sliding distance of 5,000m.

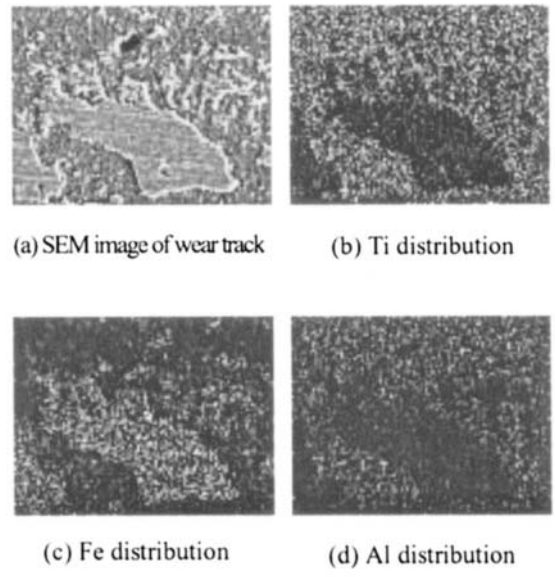


Figure 9. Plane analysis results of wear track for the film of R_a=0.17μm at the sliding distance of 5,000m.

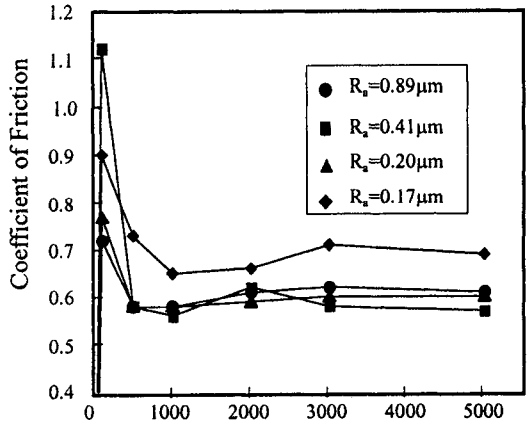


Figure 10. Relation between coefficient of friction and slidingdistance.

Figure 10 shows the relation between coefficient of friction and sliding distance for four surface roughness. There are no significant differences between four surface roughness films. But, the coefficient of friction initially increases and then is close to about 0.6 in the range of over the sliding distance of 1,000m.

3.3 Wear behaviour of ball

Figure 11 shows the relation between wear amount of ball and sliding distance against the TiAlSiN film. The mass loss of ball increases with increasing the sliding distance. When the surface roughness of film is great, the mass loss of ball increases. The mass loss of ball against R_a=0.89μm film is about 13mg at the sliding distance of 5,000m but the mass loss against the other three surface roughness films is less 5mg. For R_a=0.89μm film, there were many droplets, as shown in Figure 2. The mass loss of ball against the substrate is less than that of against R_a=0.89μm film. While, the mass loss of balls is less for the polished film than for the substrate. This suggests that if the droplet could be removed from the TiAlSiN film, the mass loss of ball can be remarkably improved. Thus, the great surface roughness of the film contributes to result in abrasive wear of ball.

Figure 12 is the optical microscope images of ball wear track against R_a=0.89μm film, R_a=0.41μm film, R_a=0.20μm film and R_a=0.17μm film. All wear tracks of ball show nearly a circle shape. If the wear of the film would happen, the wear track of ball should be ellipse shape. But, when the TiAlSiN film does not wear and then the ball alone wears, the wear track always keeps a circle shape. Thus, it is considered that for TiAlSiN film, the wear hardly occurs but the ball is reduced by the abrasive wear.

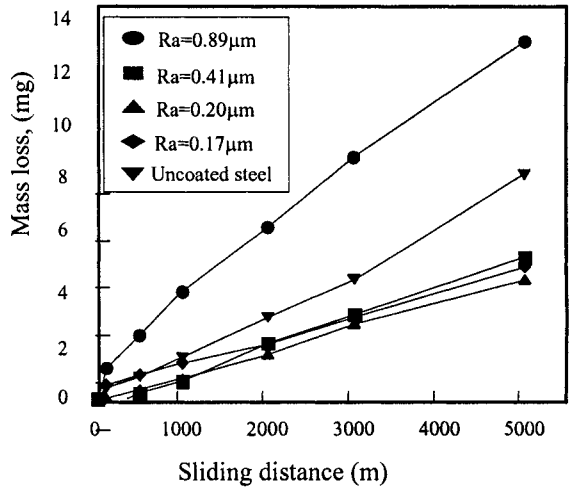
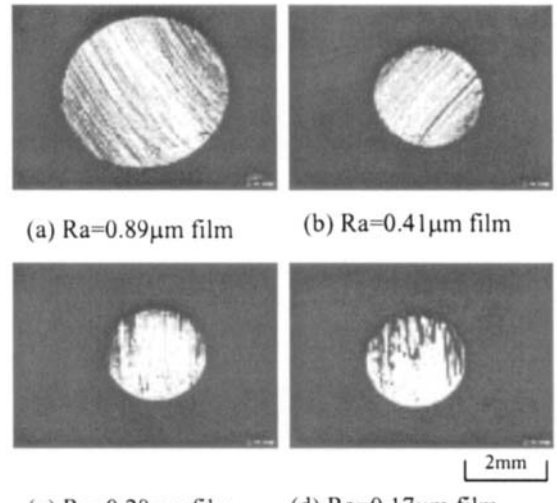


Figure 11. Relation between mass loss of ball and sliding distance against TiAlSiN film.



4. CONCLUSIONS

TiAlSiN film was coated by Arc Ion Plating onto high-speed tool steel, SKD61. The wear of TiAlSiN film against bearing steel, SUJ2, ball was investigated using the ball on disk standard testing machine. The results obtained were as follows.

(1) The existence of Si was detected as $Ti_7Al_5Si_{12}$ compound in the TiAlSiN film by XRD.

(2) For TiAlSiN film, the wear did mainly not happen, regardless of the film surface roughness. While, the substrate was occurred by the severe wear. Thus, the wear of SKD61 was improved by the coating TiAlSiN film.

(3) For SUJ2 ball, the wear linearly increased with increasing the sliding distance. It was the mass transfer process. The droplet on the TiAlSiN film resulted in the increases of wear weigh of ball. When the TiAlSiN film was polished to eliminate the droplet, the wear of ball was markedly reduced.

REFERENCES

[1] J. Almer and M. Oden, Microstructural evolution during tempering of arc-evaporated Cr-N coatings, American Vacuum Society, Vol. A 18(1), Jan/Feb, pp. 121-130, 2000.

[2] A. Tanaka, Deposition of Hard Thin Films and Their Tribological Properties, Tribologist, Vol. 46, No. 7, pp. 558-563, 2001.

[3] A. Okamoto, D. Yonekura, R. Murakami, The Effects of Macro Particle and film Thickness of Thin Film on Fatigue Strength, Procs. of APCFC & ATEM '01 (JSME No. 01-203), Oct., Vol. 20-22, pp. 450-454, 2001.

[4] A. Baba, Y. Yoshida, *Handbook of Mechanical Engineering* (Sanseido, Tokyo, 1986), p.229.

[5] J. D. Bressan, Abrasive wear behavior of high speed steel and hard metal coated with TiAlN and TiCN, Wear, No, 250, pp.561-568 2001.

[6] Xial-Ming He, Structure, hardness, and tribological properties of reactive magnetron sputtered chromium nitride films, American Vacuum Society, Vol. A, 18(1) pp.30-36, Jan/Feb, 2000.

[7] A.A. Torrance, I. Galligan, G. Liraut, A model of friction of a smooth hard surface sliding over a softer one, Wear, No. 212, pp.213-220, 1997.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

WEAR MECHANISMS IN THERMAL SPRAY STEEL COATINGS

A. Edrisy [a], T. Perry [b], Y. T. Cheng [b] and A. T. Alpas [a]

[a] Department of Mechanical, Automotive and Materials Engineering
University of Windsor, Windsor, Ontario, Canada N9B 3P4

[b] General Motors Research and Development Centre,
30500 Mound Rd. Warren, MI, US, 48090-9055

ABSTRACT

Low carbon steel coatings (0.2% C) deposited on 319 Al alloy substrates by a plasma transfer wire arc (PTWA) process, were subjected to dry sliding tests over a range of loads (5-75 N) and sliding speed (0.2-2.5 m/s). A wear map was constructed to summarize the measured wear rates and mechanisms that control the wear rates. Four regimes were identified for sliding wear under dry atmospheric conditions: at low loads (<20 N) and velocities (<1 m/s), surface oxidation to Fe_2O_3 was the main wear mechanism. The highest wear rates occurred at high loads and low velocities, and were associated with the fracture of the highly deformed steel splat tip. At high loads and velocities the wear rates decreased, as a result of the formation of thicker oxide rich tribo-layers and hardening of the sliding surfaces. The wear rates were the lowest at low load and high velocity conditions, where the surfaces were covered with oxide rich tribo-layers. The effect of atmospheric humidity was also studied by performing sliding wear tests in an environmental test chamber at humidity levels between 10% and 99% RH. As the atmospheric humidity increased above 50% RH, a tribo-polishing process became active, and the wear rates decreased.

Keywords: Thermal spray; Steel coatings; Wear map; Wear Mechanisms; Coatings for aluminum.

1. INTRODUCTION

Wear and scuffing resistances of cast aluminium components that are currently being used in automotive applications are generally inferior compared to the components traditionally made of cast iron. One possible way of improving wear resistance of aluminium castings in applications such as lightweight engine blocks is to deposit iron based coatings on the surfaces of the cylinder bores using a thermal spray process. The microstructures of thermal sprayed steel coatings consist of iron splats separated by thin oxide layers (FeO). It has been proposed that the oxide veins constitute the 'weak links' in the microstructure and could cause splat delamination during sliding wear [1, 2]. The effects of applied velocity, load and humidity on the wear behaviour of low carbon steel coatings produced using a PTWA process have been investigated by Edrisy et. al [3,4]. The main wear mechanisms under dry air atmosphere (RH = 10%) were associated by oxidation; fracture and fragmentation at the edges of splats; surface hardening and formation of thick and protective oxide layers. Welsh [5] found a transition from severe to mild wear at a critical sliding speed as a result of surface hardening of plain carbon steel. Papaphilippou et. al [6] studied the effect of humidity on wear of cast iron. They attributed the low wear rate at high humidity to the less abrasive character of debris. Lim and Ashby [7] studied the wear of steels using a wear map approach. The present study adopts a similar wear mapping approach. The wear rates of PTWA low carbon steel coated 319 Al alloys tested in dry sliding condition are presented on load vs. velocity axes. The effect of humidity on the sliding wear behaviour of the PTWA coatings is also discussed.

2. MATERIALS AND EXPERIMENTAL PROCEDURES

A PTWA gun with an AISI 1020 steel wire was used. The coating was built up by translating the gun over the 319 Al coupon surfaces (25mm × 25mm)

multiple times to produce coatings that were 250 ± 10 μm thick. SEM and XRD analyses confirmed that microstructure consisted of elemental Fe in the form of splats and FeO type iron oxide veins between them. The hardness of the PTWA coating was of 300 ± 10 kg/mm^2. The counterface pin material (5.0 mm diameter) was an AISI M2 high-speed tool steel (hardness = 910 ± 10 kg/mm^2).

The samples were polished to 0.38 ± 0.08 μm before wear test. The wear tests were performed using a pin-on-disc sliding wear apparatus enclosed in an environmental chamber. The wear tests were run to a constant sliding distance of 5×10^3 m. After each test the specimens were cleaned of loose debris and weighed to determine the amount of mass change during the test. SEM, EDX and XRD were used to delineate the wear mechanisms.

3. RESULTS AND DISCUSSION

3.1 Wear Map

Wear rates of the PTWA coatings for tests in the 5-75 N load and 0.2-2.5 m/s sliding speed ranges under dry air atmosphere (%RH = 10 ± 2) are presented in a map form on load versus sliding speed axes (**Fig. 1**). The measured wear rates are given numerically at different sliding speed and load combinations. The wear rates were high at low sliding velocities but decreased with increasing the sliding velocity. The map shows four wear regimes. These wear regimes are classified on the basis of differences in wear rates, as well as on the morphological and compositional differences of worn surfaces and loose debris.

a) Regime I

Regime I occurs at low sliding velocities and loads i.e., below approximately 1 m/s and 20 N. The back-scattered SEM micrographs of the worn surfaces of coatings tested at this regime shown in **Fig. 2**. EDS analysis of the grey dark patches showed the dominant presence Fe and O and light grey area were Fe. The oxide rich layers were uniformly distributed between the micro-scratches on the contact area. The metallic parts were subjected to plastic deformation and extruded in the sliding direction. Wear debris generated in this regime was orange in colour and powdery to the naked eye. The XRD spectrum of the debris at 10 N load and 0.5 m/s velocity, representing regime I, showed the peaks for hematite (Fe_2O_3) and few small peaks of ferrite [3]. The quantitative phase analysis confirmed that at low load and velocity conditions Fe_2O_3 was the predominant constituent (>99%) of the wear debris [3]. Therefore, the main wear mechanism in Regime I was surface oxidation to Fe_2O_3.

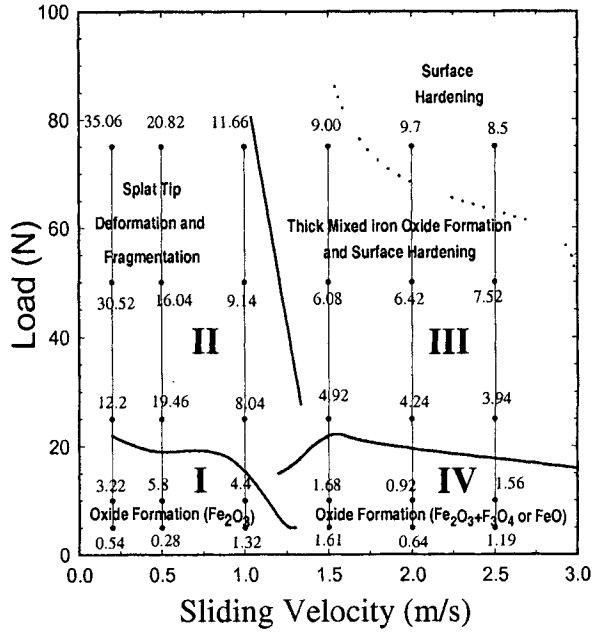


Fig. 1. Wear map for PTWA 0.2% C steel coatings in dry sliding under dry air (%RH = 10 ± 2). Multiply wear rates on the map by 10^{-6} to obtain the actual rates in units of g/m.

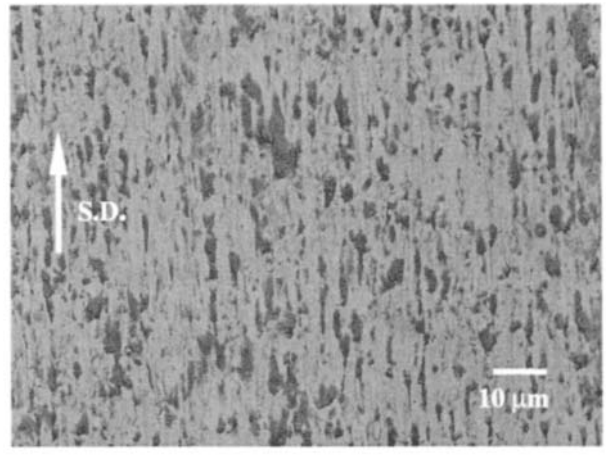


Fig. 2. The backscattered SEM micrograph of a sample worn at 10 N, 0.5 m/s and RH = 10% (regime I).

b) Regime II

Regime II is characterized by the high wear rates. This regime occurs at loads above approximately 20 N and sliding velocities lower than 1m/s. **Fig. 3** is a typical back-scattered SEM micrograph of a worn surface of a sample tested in this regime. About 20-30% of the wear tracks were covered by the oxide rich films (dark grey regions). The top surfaces of steel splats (medium grey regions) were flattened as a result of plastic deformation during wear and extruded in the sliding direction. An important aspect of wear at this regime is the fracture of the edges of the highly deformed steel splats. The wear debris collected under these conditions was magnetic, with a dark brown colour. The XRD spectrum of the wear debris of a sample tested at 50 N load and 0.5 m/s showed the presence of mixed oxides, namely Fe_2O_3, Fe_3O_4, and FeO [3]. It is important to emphasize that there was also significant amount of metallic ferrite fragments in the debris. The metallic fragments were formed as a result of fracture of iron splats. Iron splats adjacent to contact surfaces deformed became elongated in the direction of sliding. Exhaustion of ductility of the material at the splat tip eventually caused fracture of the splat tips, which were subsequently detached to form loose debris particles. Occasionally, repeated fracture and fragmentation of splats progressed to a degree where the entire splat became fragmented.

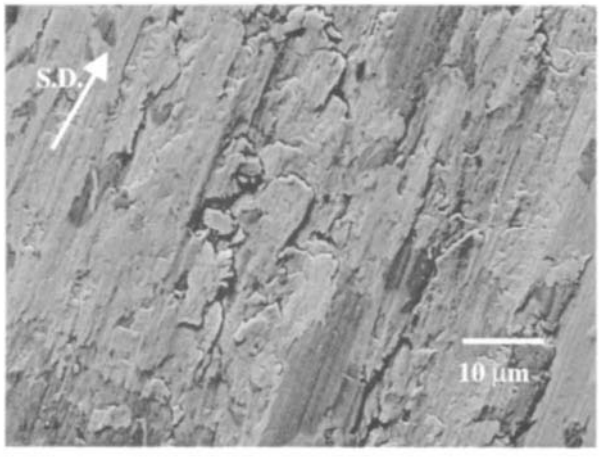


Fig. 3. The backscattered SEM micrograph of a sample worn at 75 N, 0.5 m/s and RH = 10% (regime II).

c) Regime III

Regime III exists at sliding velocities above those in regime II, and at the same load levels. At this regime the wear rates were about 50% lower than those measured in the regime II. About 70% of the wear tracks of the samples tested under regime III conditions were covered by relatively thick oxide rich layers, whose average thicknesses varied between 1-3 μm. The micrograph in **Fig. 4** is taken from a sample tested in this regime and shows evidence for surface damage, mostly in the form of surface grooving of both the oxide rich layers and the iron splats. The worn surfaces appeared to be less susceptible to splat fracture. The XRD spectrum of the wear debris of a sample tested at 50 N load and 2 m/s showed the same constituents of the worn debris of the sample tested in regime II, but the percentage of the ferrite in the debris was considerably smaller [3]. The friction results showed that by increasing velocity from 0.5 to 2.5m/s at 50 N load the coefficient of friction decreased from ~ 0.8 to ~ 0. 6. This is also consistent with the formation of thick oxide rich film. At this regime especially high loading conditions the hardness of wear tracks increased to exceptionally high values of over 700 kg/mm^2. Therefore the wear rates were reduced in regime III (relative to regime II) as a result of two mechanisms i) the production of thick oxide rich layers that reduced the coefficient of friction, and probably suppressing the splat tip fracture mechanism, ii) surface hardening during wear possibly due to local strain hardening or martensitic transformation during heating and cooling cycles that occur during sliding contact.

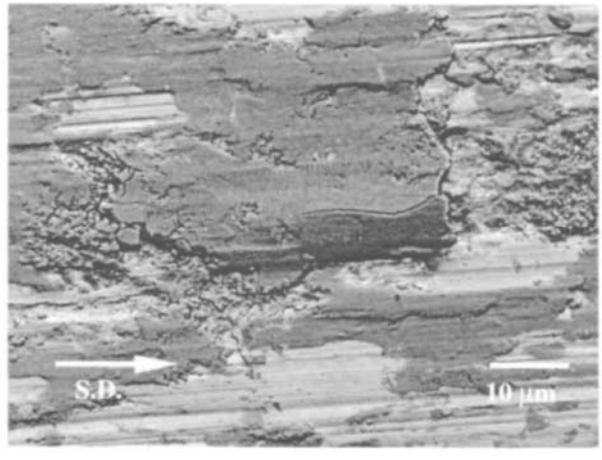


Fig. 4. The backscattered SEM micrograph of a sample worn at 50 N, 2.5 m/s and RH = 10% (regime III).

d) Regime IV

At sliding velocities above 1 m/s and loads approximately below 20 N, the wear rates were the lowest compared to those in the other regimes. The SEM micrograph of the worn surfaces of the coatings tested under these conditions (**Fig. 5**) showed a thin layer (0.5-1.0 μm) of oxide deposits. The coverage was more significant than at low load and low velocity conditions of regime I, and extended over the top of the steel splats. Because the wear rate was very low, insufficient wear debris could be collected for XRD investigation. However favourable factors, including the formation of almost continuous oxide rich films, hardening of the coating during wear, and lack of any significant iron splat fracture, are among the possible reasons why the wear rates are the lowest under these conditions.

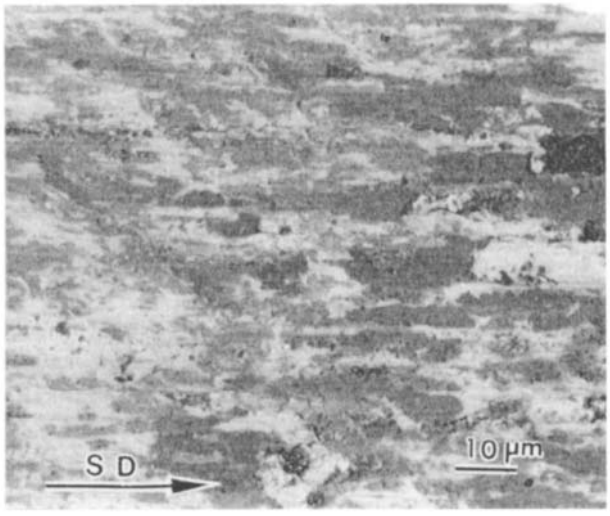


Fig. 5. The backscattered SEM micrograph of a sample worn at 10 N, 2.5 m/s and RH = 10% (regime IV).

3.2 EFFECT OF RELATIVE HUMIDITY ON THE WEAR

The variations of the wear rates of the PTWA coatings with relative humidity (RH) are shown in **Fig. 6**. The coatings tested at 5 N and 0.1 m/s showed that at 30% RH the wear rate decreased to zero and at humidity levels above 30%, the samples gained weight during sliding. The high humidity atmosphere could cause a weight gain by either increasing the rate of oxidation, or hydrating the nascent surface oxide. Similarly, for the testing conditions of 10 N and 0.5 m/s, the wear rates decreased as the relative humidity increased but no weight gain occurred. **Fig. 7** is the back-scattered SEM micrographs of the worn surface of coatings tested at a 10N load and 0.5 m/s sliding velocity (regime I) but under high humidity environments (80% RH). EDS analyses identified the light grey areas as metallic and the dark areas as oxygen rich layers. The metallic parts were polished to almost a mirror finish, so well that the oxide veins (FeO layers) between the splats can be clearly observed on the worn surfaces.

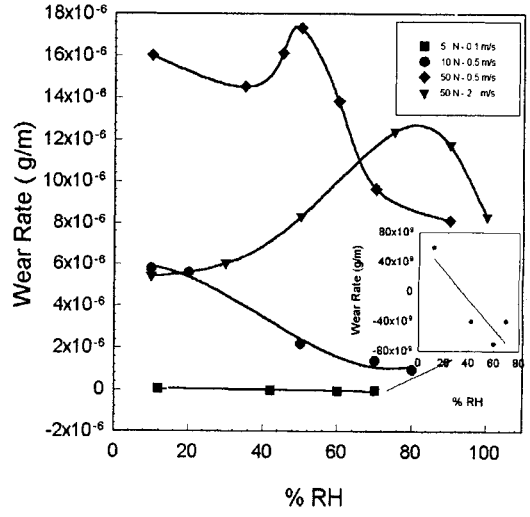


Fig. 6. The variation of wear rates with the relative humidity for four different testing conditions. The insert shows wear rates for tests at 5 N-0.1 m/s.

For sliding wear tests that were performed at the same sliding velocity of 0.5 m/s, but a higher normal load of 50 N, the wear rates were significantly higher within the humidity range of 10% and 50% (**Fig. 6**). SEM investigations of the worn surfaces of the coatings tested at this loading condition and relative humidity levels less than 50% indicated that the worn surface morphologies of the coatings were similar to the one shown in **Fig. 3**. The steel splats were flattened as a result of plastic deformation during wear and fractured. Above 50% RH a dramatic decrease in the wear rate was observed so that the wear rate at 90% RH is half that of the sample tested at 50% RH. At 90% relative humidity the worn surface was smoother with the less evidence of fracture (**Fig. 8**). The dominant wear mechanism at high humidity conditions was a chemical-mechanical polishing

process. This mechanism was active whenever there was a balance between chemical and mechanical wear.

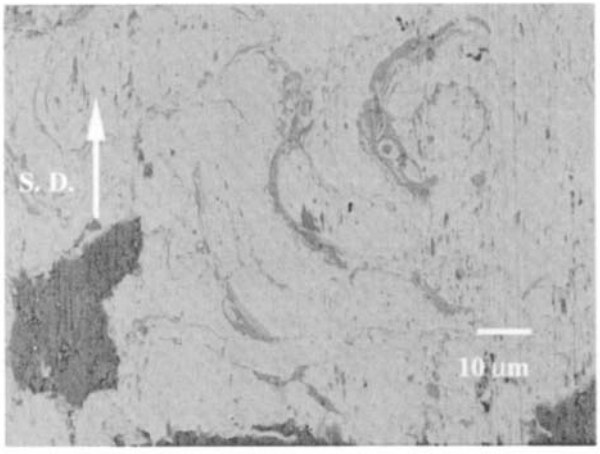


Fig.7. The backscattered SEM micrograph of a sample worn at 10 N, 0.5 m/s and RH = 80%.

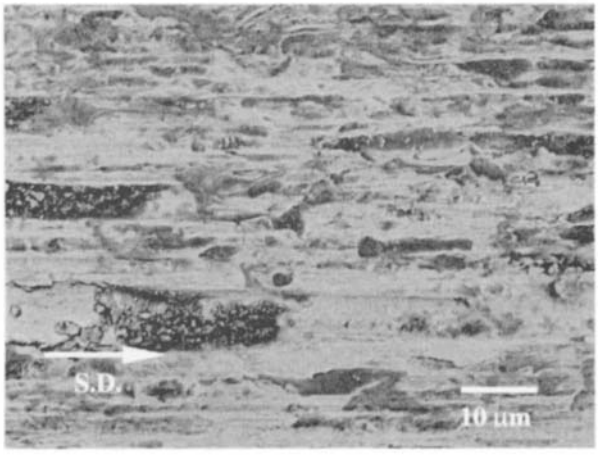


Fig.8. The backscattered SEM micrograph of a sample worn at 50 N, 0.5 m/s and RH = 90%.

4. CONCLUSIONS

A wear rates and mechanisms map was constructed for the PTWA low carbon steel coatings in dry sliding under dry atmosphere. Four wear regimes were identified. These were:

- Regime I oxidation or formation of Fe_2O_3 at low load and sliding velocity.
- Regime II severe deformation of the steel splat tip and splat fracture and fragmentation and oxidation at low velocity and high load.
- Regime III formation of a thick protective mixed oxide layer and a surface hardening at high load and sliding speed.
- Regime IV formation of mixed oxide layer and no evidence for the splat tip fracture at high velocity and low load.

At high humidity conditions the wear rates decreased. It was suggested that the lower wear rates at high humidity atmosphere were due to the effect of a tribo-chemical polishing wear mechanism.

REFERENCES

1. A. Rabiei, D. R. Mumm, J. W. Hutchinson, R. Schweinfest, M. Ruhle and A. G. Evans, Materials Science and Engineering A 269, (1999) 152-165.
2. S.E. Hartfield-Wunsch and S.C. Tung, in Thermal Spray Industrial Applications, C. C. Berndt and S. Sampath (Eds.), Materials Park, OH, ASM International, (1994) 19-32.
3. A. Edrisy, T. Perry, Y. T. Cheng and A. T. Alpas, "Wear of Thermal Spray Deposited Low Carbon Steel Coatings on Aluminum Alloys", Wear 8872 (2001) 1023-1033.
4. A. Edrisy, T. Perry, Y. T. Cheng and A. T. Alpas, The Effect of Humidity on the Sliding Wear of Plasma Transfer Wire Arc Thermal Sprayed (PTWA) Low Carbon Steel Coatings, Surface and Coatings Technology, Vol. 146-147, (2001) 571-577.
5. N.C.Welsh, Journal of Applied Physics, Vol. 28, No. 9, (1957) 960-968.
6. C. Papaphilippou, M. Vardavoulias and M. Jeandin, Wear, 177 (1994) 151-157.
7. S.C. Lim and M.F. Ashby, Acta Metallurgica. 35, (1987) 1-24.

This Page Intentionally Left Blank

SESSION IX
FRICTION, WEAR AND LIFE OF CONTACTS IN MACHINE ELEMENTS (1) – LIFE PREDICTIONS

Chairman : Professor A A Lubrecht

Paper IX (i) **The Role of Wear in Enhancing Rail Life**

Paper IX (ii) **Laboratory-based Wear Life Testing of Sliding and Rolling Contacts – The First Step Towards Technology Transfer**

Paper IX (iii) **Life Prediction for Surface Crack Initiated Contact Fatigue of Silicon Nitride Bearing Balls**

Paper IX (iv) **Experimental Investigations on Rolling Contact Fatigue of Dented Surfaces Using Artificial Defects: Sub-surface Analyses**

Paper IX (v) **Incorporating a Railway Wheel Wear Model into Multi-Body Simulations of Wheelset Dynamics**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

The Role of Wear in Enhancing Rail Life

A. Kapoor[a], D. I. Fletcher[a], F. J. Franklin[a]

[a]Department of mechanical Engineering,
The University of Sheffield,
Mappin Street,
Sheffield,
S1 3JD,
UK

For many engineering components the primary problems to be prevented during operation are wear and fatigue crack growth. However, in many cases one of these mechanisms dominates to the extent that the other may be neglected. In cases where both processes of almost equal importance in determining component life their interaction or coupling becomes critical. This paper presents a generic introduction to the concept of interaction between wear and fatigue in the failure of an engineering component. To illustrate the topics discussed a detailed example is presented of wear-fatigue interaction in the failure of a railway rail.

1. INTRODUCTION

This paper presents a generic introduction to the concept of interaction between wear and fatigue in the failure of an engineering component. To illustrate the topics discussed a detailed example is presented of wear-fatigue interaction in the failure of a railway rail. The combination of the wear and fatigue processes, which are often treated independently, highlights critical issues for components which can suffer from both processes simultaneously.

2. WEAR - FATIGUE INTERACTION

For many engineering components the primary problems to be prevented during operation are wear and fatigue crack growth. However, in many cases one of these mechanisms dominates to the extent that the other may be neglected. In cases where both processes of almost equal importance in determining component life their interaction or coupling becomes critical. Firstly, actions taken to alleviate one of the problems may influence the other. For example, oil may be applied to prevent wear, but oil will enter surface breaking cracks and influence crack growth. Secondly, the failure processes themselves interact. For example, a severely worn surface can have many surface flakes which are beginning to break away, each of which is effectively a small crack. More critical, although largely overlooked, is the ability of wear to shorten or remove existing surface breaking cracks. Therefore, if a degree of wear can be tolerated there exists a means to remove small cracks before they become a serious problem.

2.1. Crack truncation

Figure 1 shows schematically a cross section of a surface which contains a crack (length a) but which is also wearing. Crack tip advance rate $\frac{da}{dN}$ may depend on crack length, friction between the crack faces, and the magnitude and direction of externally applied loads. Wear rate (w) similarly depends on a variety of factors, but its dependence is unlikely to match that of the crack tip advance rate.

Taking into account both the surface wear and crack tip advance rate (which may be defined using the Paris-Erdogan equation [1]) a net crack growth rate can be defined, Equation (1). Constants C and m are the Paris-Erdogan constants, and ΔK is the stress intensity factor for the crack.

$$\frac{da}{dN}_{net} = \frac{da}{dN} - \frac{w}{\sin(\theta)} = C\left(\Delta K\right)^m - \frac{w}{\sin(\theta)} \quad (1)$$

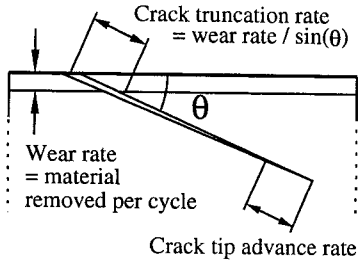


Figure 1. Truncation of a surface breaking crack through wear of the surface. Net crack growth rate equals the crack tip advance rate minus the rate of crack mouth truncation.

2.2. Life determination

Application of the very simple relationship shown in Equation (1) requires knowledge of the dependence of stress intensity factor ΔK and wear rate on external factors, and this is component specific. These relationships are discussed for a railway rail application in Section 3. However, one of the critical determinants of crack growth rate is crack length, while wear is completely independent of crack length. This provides a simple means to investigate the implications of wear-fatigue interaction.

Taking a very simple example, the stress intensity factor for a single edge crack growing normal to the surface ($\theta = 90°$) of a plate (which is much wider than the crack) in uniform tension is given by Equation (2) where σ is the remotely applied stress and a is the crack length [2].

$$\Delta K = 1.12\sigma\sqrt{\pi a} \qquad (2)$$

Component life due to fatigue is typically determined by the development of a crack of a critical length (a_{crit}) and the number of cycles for this length to be reached (based on some initial crack length, a_o) can be determined by integration of the combination of Equations (1) and (2). Taking examples of a plate with an initial edge crack length of 0.1mm and 1.0mm, and a critical crack length of 5mm, the life due to fatigue can be plotted against wear rate, Figure 2. It is assumed that the plates are loaded in tension to 400MPa, and that this load is applied so as to simultaneously produce both wear of the surface and fatigue loading of the crack. In practice wear and fatigue are most likely to occur simultaneously with contact loading, for which the crack growth rate has a more complex relationship to crack length, but an edge crack in tension is considered here to illustrate the essence of wear-fatigue interaction. Constants C and m are taken to be $0.24 \cdot 10^{-11}$ and 3.3 [3], giving crack growth rate in m/cycle for ΔK in MN $m^{-3/2}$. Also shown in Figure 2 is a wear life curve, although since the details of the component have not been specified this is purely schematic, based on an allowable wear depth which is consumed more rapidly as the surface wear rate increases.

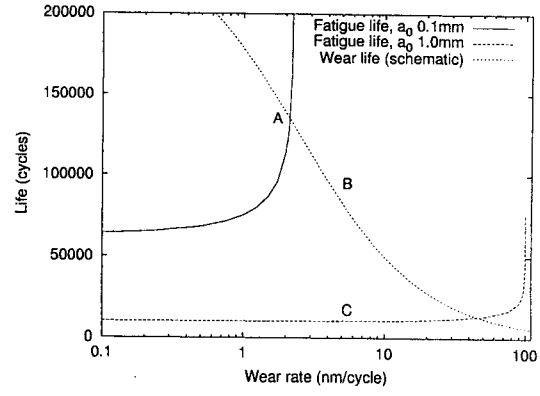


Figure 2. Fatigue life variation with surface wear rate for two different sizes of single edge crack in a plate. Wear life is plotted schematically, based on an allowable wear depth which is consumed more rapidly as the surface wear rate increases.

From Figure 2 it can be seen that the life of the component is significantly influenced by the size of the initial defect considered (as would be expected) but for both initial defect sizes the life is also significantly dependent on the wear rate, i.e. the rate of crack truncation, through its influence on the net crack growth rate. The fatigue life of the component rises as surface wear rate is increased, reaching an asymptote at the wear

rate which is just high enough to prevent a crack of the initial size from growing. At wear rates above this point the initial cracks will actually be removed by further load applications.

Although consideration of fatigue life curves in Figure 2 suggests that with a sufficiently high wear rate the component life can be made infinite, this is not the case because life will become limited by wear. Considering the schematic wear curve shown in Figure 2, the maximum life for each initial defect size is at the crossing of the fatigue and wear life curves. Particularly for the smaller initial defect, this maximum life position (point A) would be difficult to maintain unless wear rate could be closely controlled, because life drops away rapidly on either side of the maximum. Movement onto the fatigue controlled life curve (typically representing a dangerous failure) would therefore be likely if attempting to operate at this maximum position. However, by operating at an increased wear rate (point B) a good overall component life could be obtained while being reasonably certain that failure will occur by wear (typically a slowly occuring and therefore safer failure) and not by fatigue. Unfortunately the presence of defects of different initial sizes complicates the situation because failure can be expected by the mechanism indicating the lowest life at any particular surface wear rate. In the presence of larger initial defects failure may actually occur at point C rather than point B, and it will again be on a fatigue controlled curve.

To illustrate the importance of wear-fatigue interaction a railway engineering example is presented below. This highlights some of the ways that the curves shown in Figure 2 can be defined for a rail-wheel contact, for which it is found that crack growth can be driven by different mechanisms according to crack length.

3. A RAIL-WHEEL CONTACT EXAMPLE

The surface material of wheels and rails accumulates deformation through a process of ratcheting. The deformation continues until the ductility of the material is exhausted and the material fails. This failure manifests itself in the form of wear and/or rolling contact fatigue, Figure 3. Cracks initiating in this way are propagated by the action of the contact loading, fluid pressure inside the cracks, rail bending stresses and residual stresses generated during rail manufacturing or by atmospheric temperature. These factors can act individually or in combination. Various crack growth phases can be identified: crack initiation and growth by ratcheting, crack growth by contact stresses, and crack growth by rail bending stresses.

Each time the contact zone passes over the crack, the crack tip advances, but at the same time, the surface layer of the rail is being worn away, leading to the truncation of crack at its mouth. The net crack length depends on the relative values of the rate of advance of the crack tip and the crack mouth truncation rate. If the rate of crack mouth truncation is greater than the rate of advance of the crack tip then the crack is ultimately eliminated. If the reverse is true, then the crack keeps on advancing.

3.1. Wear and crack initiation

Figure 3 shows a schematic of a rail cross section subjected to many repeated wheel passes. The top surface layer (of thickness tens to hundreds of microns) is heavily sheared in the direction of traction. A portion of this layer can be modelled by dividing the material into elements, Figure 4, subjected to rolling-sliding contact by wheel passes. The material accumulates shear deformation in each wheel pass as shown in the bottom picture in Figure 4. In each cycle the amount of deformation is only small, of the order of plastic strain at yield, but, after the hundreds of thousands of cycles to which the rail and wheel surfaces are subjected, the accumulated deformation can become large. For normal grade rail steels the accumulated shear strain may easily exceed 10! This gradual accumulation of plastic deformation is normally known as ratcheting.

Under atmospheric conditions a steel specimen in a shear test will fail at a strain of about 0.5 to 1.0, however as a result of the high hydrostatic compression, the surface material of the rail accumulates an order of magnitude larger strain without failing. The compressive hydrostatic stress

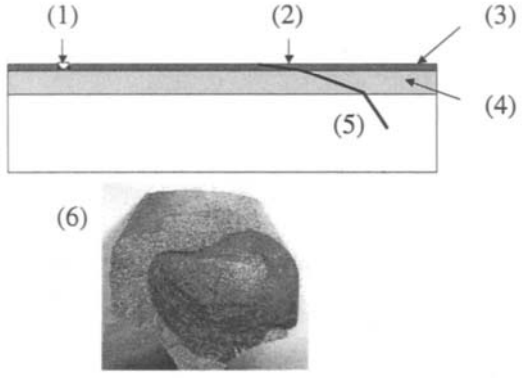


Figure 3. Critical stages in wear and rolling contact fatigue. (1) Material subject to rolling-sliding contact may detach as wear debris, or (2) small cracks may develop. (3) Early propagation of cracks will be in the near surface plastically deformed material. (4) Later propagation will be driven by contact stress and fluid pressurisation of the cracks in deeper, elastic, material. (5) As the crack extends it may branch and be driven by rail bending stress. (6) Eventual rail failure occurs by fast fracture after a critical crack length is reached.

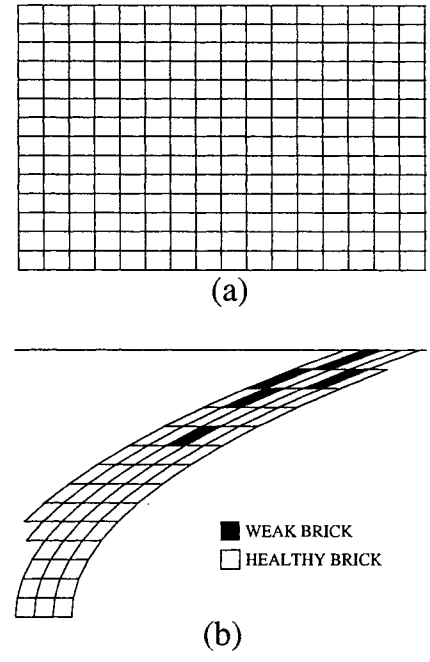


Figure 4. Modelling ratcheting failure of surface material. (a) Division of the material into cells or "bricks". (b) Cells become distorted as they accumulate deformation, which leads to failure of the bricks according to their material properties.

acting on the top surface in a Hertzian contact equals the maximum contact pressure [4] which in rail wheel contact is in the range of 1000 - 3000 MPa. This high hydrostatic compression suppresses void growth and coalescence which is the normal cause of rupture in ductile materials. The material still ruptures but the strain at failure is found to increase many folds [5] making possible the strains typically observed in railway rails.

When subjected to cyclic strains, ductile materials are found to fail either by Low Cycle Fatigue (when the life is given by the well known Coffin-Manson equation) or by Ratcheting Failure (when the life can be estimated by dividing the strain to failure by strain accumulation per cycle). Both LCF and RF act on the same material and so whichever occurs first causes failure. In this competitive mode of failure, the actual number of cycles to failure is the minimum of the cycles to failure by LCF or by RF [6]. It is possible that these failure mechanisms are additive, in which case the actual number of cycle to failure may be estimated by using a Miner's rule type equation. However, the use of Miner's equation demonstrates that the maximum difference in life estimated by the two methods is only about 50% [7]. The life of the material element at the rail surface can be estimated by using either of the methodologies by relating it to operating conditions (maximum contact pressure, friction coefficient) and material properties (shear stress at yield, strain hardening properties, critical strain to failure) [7].

3.1.1. Material properties variation

Material properties are not uniform and have variations of the order of a few percent [8] and this results in material elements failing at different locations at different times. Figure 5 shows how modelling indicates the distribution of failure leads to generation of wear debris and crack-like flaws. When all material elements around a region have failed it cannot remain attached to the surface, and it detaches as wear debris. It is interesting to see that all the material in a wear debris particle need not have failed before it can be detached. Depending on the spatial variation in material properties, wear debris will be detached at different rates in different locations, giving rise to a characteristic roughness in the surface, as shown in Figure 5.

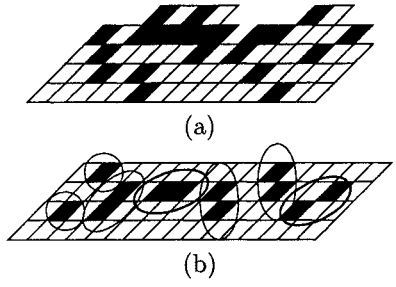


Figure 5. Schematic of ratcheting failure simulation results. (a) Loss of unsupported bricks indicates wear and surface roughness development. (b) Failed but supported bricks represent crack initiation and propagation.

In addition to surface roughness Figure 5 also shows regions which have failed but are still within the body of the rail. Such regions represent flaws which lead to cracks and RCF in rails and wheels. A methodology has been implemented to identify these flaws on a computer which can simulate thousands of cycles of wheel pass in a few minutes using a medium speed desktop PC [9]. Results of one such simulation are shown in Figure 6. As the shear deformation accumulates so does the degree of damage and is characterised in the picture by increasing grey. Black identifies the material which is fully damaged. The right half of the picture is a cross section of a BS11 rail steel disc which has been subjected to 40,000 cycles of rolling-sliding at 1500 MPa with 0.1% slip and water lubrication [10]. The similarity between the real experiment and computer simulation is striking.

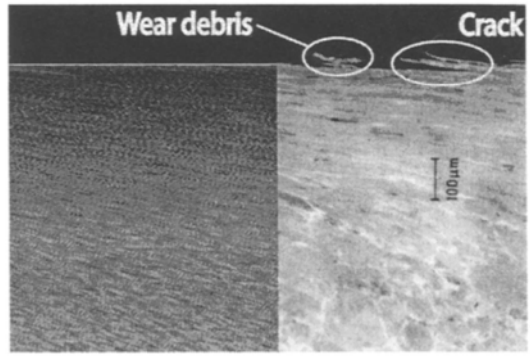


Figure 6. Comparison of ratcheting failure simulation results (left) with a micrograph from a rail steel following wear testing (right). Grey scale in the left picture indicates the degree of strain accumulation.

Figure 7 shows results of the simulation for wear rate [9], based on quantifying loss of material predicted by a simulation such as that shown in Figure 6. The pictures of simulated cross sections have also been analysed by using an image processing technique to determine the development of crack-like flaws. Many flaws of differing length are identified in each cross section, as would be seen in reality. By conducting this analysis on cross sections obtained for increasing number of cycles, it is possible to determine how length of crack-like flaws is increasing. The results have been converted into crack propagation rate (length increase per cycle) and are shown in Figure 8 for a smooth Hertzian contact. The effect of asperities which give contact stress peaks is

also shown schematically in the figure. The crack growth rate is found to decrease with increasing crack length. This is attributable to decreasing effect of asperity stresses which are localised in the top few tens of microns of the surface. Diminishing crack propagation rate for increasing crack lengths is also known for short cracks [11].

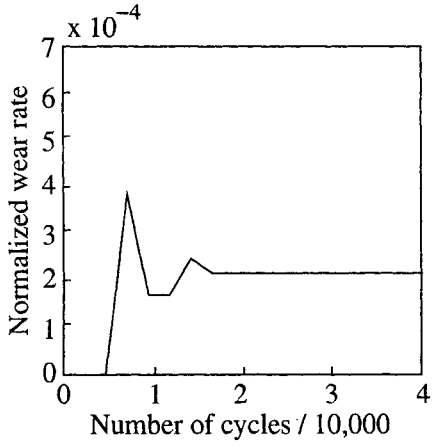


Figure 7. Wear rate predictions from ratcheting simulation. Normalised wear rate (wear depth divided by line contact width) increases with contact pressure and friction coefficient, but decreases with material hardness and increased ability to strain harden.

3.2. Larger cracks: propagation by contact stress and fluid pressure

Cracks initiate after failure of the material in a thin surface layer as discussed above in Section 3.1, and early propagation takes place in a region which is undergoing plastic deformation, but after a certain crack length/depth the crack moves into a region where stresses are elastic. If the loads are sufficiently high then the crack keeps on propagating, and its growth may be analysed using Linear Elastic Fracture Mechanics.

There are three modes of crack propagation to consider: Mode I refers to crack propagation

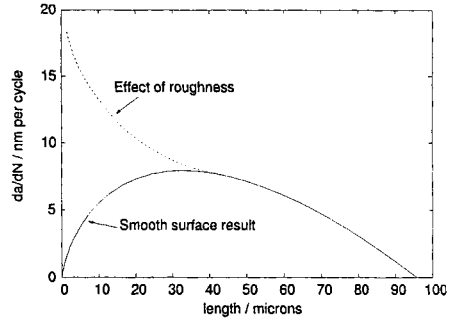


Figure 8. Predictions of crack growth rates during growth driven by ratcheting

when the applied loading pulls open the crack faces, whereas Modes II and III refer to situations when the faces are sheared backwards and forwards, and sideways respectively [8]. In most engineering situations crack faces are pulled open and the crack propagates in Mode I. However, for rail-wheel contacts the large contact loading shuts the crack and Mode I propagation is not possible without fluid entry into the crack. The fluids typically involved are rainwater or flange lubricants, which make model I propagation possible because they can be pressurised by the passing wheel contacts. The hydraulic pressure gets transmitted to the crack tip and propagates the crack in Mode I. The presence of fluid can also reduce friction between the crack faces and increase the Mode II stress intensity factor, and hence the crack propagation rate. These mechanisms are illustrated in Figure 9.

Roughly speaking the stress intensity is linked to the product of stress in the region surrounding the crack (especially around the tip) and the square root of crack length. As the crack propagates it lengthens resulting in an increase in stress intensity and crack growth rate. For many components under a remotely applied stress a crack will reduce cross-sectional area, locally increasing the stress, but for the contact loading considered here the stress is highly localised (for a point contact it decays as the inverse of distance

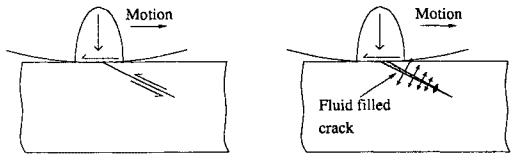


Figure 9. Crack propagation mechanisms driven by the pressure and traction from a wheel contact. (a) Shearing action driven primarily by surface friction. (b) Opening action driven by fluid pressure.

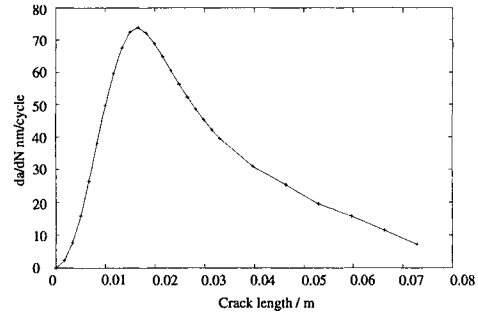


Figure 10. A sample of crack growth rate data for shear dominated crack growth. Conditions: 1255MPa maximum Hertzian contact pressure, line contact width 6.63mm, surface friction coefficient 0.111, crack face friction coefficient 0.15, crack at 30 degrees below the surface.

squared) and as the crack lengthens and deepens the tip moves into less highly stressed material. After some critical crack length and depth is reached any further increase in length therefore brings about a drop in stress intensity and crack growth rate. Crack growth therefore begins to slow down, even though the crack is extending. Such behaviour, even though known for cracks in notches, is quite different from the behaviour expected under remotely applied loads typical of many engineering components.

A Green's function based approach, developed by Fletcher and Beynon [12,13] has been used to estimate the changes in stress intensity as the contact passes over the crack. The cycle of Mode I and Mode II stress intensity thus obtained, is used in the growth rate equation developed by Bold and Brown [14] to determine the crack propagation rate. A sample output is shown in Figure 10.

It is not uncommon to see a network of cracks in rail cross sections, and multiple cracks may shield each other and reduce crack propagation rate. The cracks may also branch. If the crack branches upwards and reaches the surface then a piece may be separated from the rail and lost as wear debris. The resulting spall or pit leads to surface roughness and deterioration in ride quality. If however, the crack branches downwards and continues to propagate then potentially a more dangerous situation develops, and the rail may break. Since, the propagation due to contact stresses is expected to slow down as the crack extends there must be some other mechanism driving this stage of crack growth. This turns out to be rail bending, and is discussed next.

3.3. Crack propagation by rail bending

The rail is supported on the sleepers which in turn are supported on ballast. The wheel load deflects the rail as a non-uniformly supported beam and introduces bending stresses [15]. The beam sags just below the contact and stresses are compressive on the rail top surface. Some distance away from the contact the beam hogs and the stresses are tensile in the top of railhead. Even though not as intense as the contact stresses, bending stresses can effectively propagate long cracks especially the ones which have branched down into the rail. The crack propagates and accelerates as it lengthens. Finally, at a critical length, the crack propagates in fast mode and fracture takes place. Boundary element models of crack growth, and common solutions for cracks in bending, predict that the crack growth rate increases rapidly with crack length [16].

3.4. A combined model

All the phases of crack growth in a rail can be combined into a single plot to obtain the full life

cycle of a crack, Figure 11. Curve A shows crack initiation and early propagation, Curve B the period when the rate of growth increases with increasing crack length, Curve C the stage when the crack is moving away from the heavily stressed region near the contact and growth rate decreases with increasing crack length. The crack then branches and is driven by beam bending stresses, Curve D, and finally at a critical crack length fast fracture takes place.

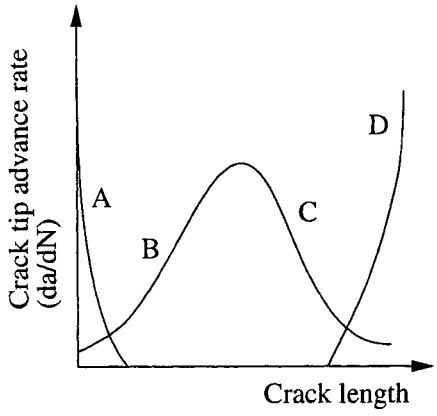


Figure 11. Crack tip advance rate plotted against crack length for all the mechanisms considered. A: Crack initiation and early propagation by ratcheting. B: Contact stress driven propagation. C: As B, but crack is moving away from the contact. D: Crack is driven by rail bending, and influenced by CWR and residual stresses. note that these curves will shift with operating conditions.

Superimposing the crack length reduction by wear on Figure 11 gives Figure 12 which depicts the effect of crack mouth truncation on life cycle of a crack. If the crack mouth truncation rate is higher than the maximum crack tip advance rate (high wear rate line) then the crack will be worn away and the surface will be left without any crack. However, if the wear rate and the corresponding crack mouth truncation rate is low then the crack passes through points where the net crack growth rate (Equation 1) becomes zero. At these points the crack should become stable and not grow further, a situation sometimes seen in practice when cracks develop in rails and do not grow with time. In reality the crack is propagating at the tip but is truncating at the mouth at the same rate giving zero net crack growth rate. These points are shown as 1, 2, 3 and 4. Point 1, where the cracks are very short, represents a desirable situation. However, due to changes in operating conditions such as friction coefficient and loading with each contact pass the curves A and B may shift relative to one another, and relative to the wear rate. If this happens then the crack may move from curve A to B. It can then continue to grow till it reaches point 2. Theoretically it should stabilise there, but this is an unstable equilibrium and any slight increase in crack length or movement of the curves with operating conditions can give a positive net crack growth rate. The crack will therefore almost certainly continue to grow to point 3, where the negative slope of the crack tip advance rate curve ensures stability. However, this is a dangerous situation as cracks are fairly long at this stage and may branch down and propagate by rail bending stresses (curve D). Point 4 is an unstable equilibrium and so unable to stop the crack, which continues to grow and finally fails by fast fracture. A low wear rate represents an unsafe situation as the crack can move without stopping from curve A to D. This situation can be remedied by increasing the wear rate artificially by grinding.

3.5. Application

From the above discussion it is clear that the rail-wheel contact case is much more complicated to analyse that the single edge crack discussed in Section 2.2. Full analysis of how cracks may stabilise at points 1 or 3 in Figure 12, or may be prevented completely, is still underway. However, some useful guidance for rail maintenance planning can be given.

A rail surface fails by either wear or RCF. Based on the permitted amount of material that can be lost, as the wear rate increases the life due to wear decreases. This is shown schematically

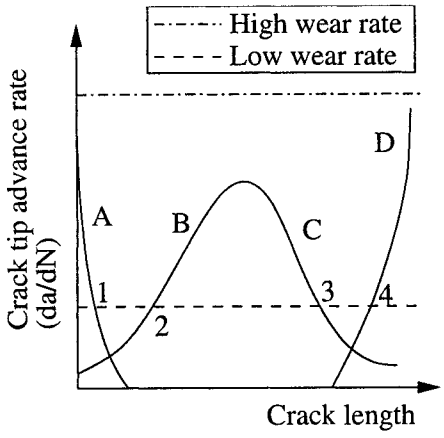


Figure 12. Combining wear rate with crack tip advance data indicates points at which crack may stabilise. For a low wear rate the crack may become stable at point 1 or 3, while points 2 and 4 represent unstable equilibrium points. Dashed lines indicate surface wear rates at high and low rates, which are crack length independent.

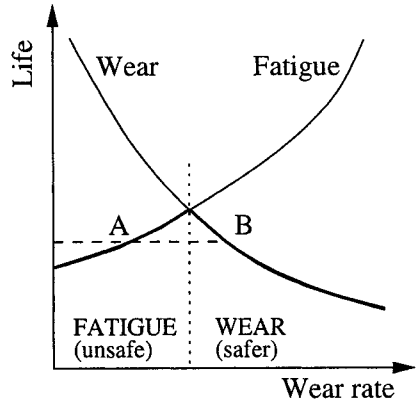


Figure 13. Control of rail life by wear and fatigue. Actual life is the minimum of wear and fatigue life for a given wear rate (thick lines). Operation at the maximum life point (dotted line) is difficult because of variability in operating conditions. For a reduced life (dashed line) operation could be at points A or B, but B is inherently safer.

in Figure 13 [17] in which the fatigue life curve is based on the crack growth rate variation with crack length shown by curves B and C in Figure 12. The Figure is essentially similar to Figure 2, but the decrease in crack growth rate at longer crack lengths (curve C Figure 12) changes the shape of the fatigue life curve. However, as in the simple edge crack case, as the wear rate increases, the crack mouth truncation rate increases and fatigue life therefore increases. The actual life is the lower of the fatigue and wear life, and is shown by the heavy curve. The peak on this curve represents maximum life which can be achieved if the corresponding material removal rate can be obtained by natural wear and grinding.

Railway operation involves variable conditions: the friction coefficient varies due to rain water, leaves and snow on the track and lubricant products which may migrate from flange contact to the railhead. The loading varies with passenger traffic and vehicles. Accurate prediction is very difficult and a good engineering solution involves using an appropriate factor of safety (or ignorance!) and operating below this optimum life, the horizontal line in Figure 13. This line intersects the life curve at two points, A and B. Point A represents failure by RCF and Point B failure by wear. It is interesting to see that for the same life a higher material removal rate leads to a safer operation. This higher material removal rate can be achieved through use of a softer rail material or through grinding.

4. CONCLUSIONS

Wear-fatigue interaction is important in determining the life of components which undergo both processes simultaneously, but for which neither process is guaranteed to dominate. Wear and fatigue are coupled primarily through crack truncation by wear, but also through "the cross actions" of solutions to one or other of the problems, e.g. oil applied to reduce wear may accelerate crack growth through internal fluid pressurisation of the crack. Controlled increases in wear offer the pos-

sibility of extending the fatigue life of a component through truncation of small surface breaking cracks. Although in subsequent operation it will not be possible to operate at the theoretical maximum life of the component, it is possible to ensure that failure will eventually be by wear (an easily detected, and therefore relatively safe failure) rather than by fatigue failure crack growth, which is often undetected until rapid final failure.

REFERENCES

1. P. C. Paris, F. Erdogan, A critical analysis of crack propagation laws, Journal of Basic Engineering 85 (1963) 528–534.
2. Y. Murakami, Stress Intensity Factors Handbook, Pergamon Press, Oxford, UK, 1987.
3. B. P. P, C. R. J, Mechanics of Engineering Materials, Longman Scientific and Technical, 1987.
4. K. L. Johnson, Contact Mechanics, Cambridge University Press, 1985.
5. M. Brandes, Mechanical properties of materials under hydrostatic pressure, in: H. Pugh (Ed.), Mechanical behaviour of materials under pressure, Elsevier, 1970, pp. 236–298.
6. A. Kapoor, A re-evaluation of the life to rupture of ductile metals by cyclic plastic strain, Fatigue and Fracture of Engineering Materials and Structures 17 (1994) 201–219.
7. J. H. Beynon, M. W. Brown, A. Kapoor, Initiation, grwoth and branching of cracks in railway track, in: Beynon (Ed.), Engineering against fatigue, Balkema Publishers, 1999, pp. 461–472.
8. N. E. Dowling, Mechanical Behaviour of Materials, Prentice Hall, 1999, pp. 800-804.
9. F. J. Franklin, I. Widiyarta, A. Kapoor, Computer simulation of wear and rolling contact fatigue, Wear 251 (2001) 949–955.
10. W. R. Tyfour, J. H. Beynon, A. Kapoor, Deterioration of the rolling contact fatigue life of pearlitic rail steel due to dry-wet rolling-sliding line contact, Wear 197 (1996) 255–265.
11. S. Suresh, Fatigue of Materials, 2nd Edition, Cambridge University Press, 1998.
12. D. I. Fletcher, J. H. Beynon, A simple method of stress intensity factor calculation for inclined fluid-filled surface-breaking cracks under contact loading, Proceedings of the Institution of Mechanical Engineering Part J, Journal of Engineering Tribology 213 (1999) 299–304.
13. D. I. Fletcher, J. H. Beynon, A simple method of stress intensity factor calculation for inclined surface breaking cracks with crack face friction under contact loading, Proceedings of the Institution of Mechanical Engineering Part J, Journal of Engineering Tribology 213 (1999) 481–486.
14. P. E. Bold, M. W. Brown, R. J. Allan, Shear mode crack growth and rolling contact fatigue, Wear 144 (1991) 307–317.
15. Winkler, Die Lehre von der Elasticitat und Festigkeit, Prague, 1867.
16. F. C. Loh, Boundary element analysis of a crack in a rail, B.eng thesis, The University of Sheffield (2002).
17. A. Kapoor, F. Schmid, D. I. Fletcher, Managing the critical wheel/rail interface, Railway Gazette International 158 (1) (2002) 25–28.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Laboratory-based wear life testing of sliding and rolling contacts – the first step towards technology transfer

T.P. Sperring and B.J. Roylance

Tribology and Condition Monitoring Research Group, Department of Mechanical Engineering, University of Wales Swansea, Singleton Park, Swansea, UK, SA2 8PP

A transition from mild to severe wear in a sliding or rolling contact denotes surface failure. It also registers the initial stages of the deterioration in a tribological component, such as a bearing or a gear, to machine failure and the need ultimately to perform a maintenance activity.

This paper will report on the results of tests performed on a number of machines in the authors' laboratory – four-ball (sliding and rolling modes), pin-on disc and ball bearing fatigue life testing machines. The principal measurements and analyses carried out, during and after testing include friction, temperature and wear, together with surface features and associated wear debris characterisation. The main purpose of the test programme was to simulate wear life conditions that are, by their very nature, either short term, (mild to severe wear transitions - sliding) or, longer term, (rolling contact fatigue) phenomena. In the context of these tests, wear debris analysis provides an important additional source of information with regard to how the wear develops and progresses in each instance. However, it also has an important part to play in linking the behaviour in the laboratory to the conditions that occur in the field that lead to component failure. This latter issue is addressed in the paper by comparing wear debris obtained from laboratory-based tests to that captured during industrial machine operations. The prospects for adopting a designed-out maintenance strategy are discussed.

1. INTRODUCTION

Prediction of the remaining useful life of a component has been a goal of tribologists for many years and wear debris analysis has been used to assess machine condition since the early 1950's [1]. By using a series of laboratory-based experiments, a history of the nature of wear throughout a component's life can be gathered under controlled load and speed conditions and associated lubrication regime and contact form. By testing components to failure in various modes, the changes in wear debris are noted and compared to that found in military engines and gearboxes.

2. TEST MACHINERY

Four test devices are used in the laboratory to simulate the various wear situations encountered in actual military engines and gearboxes. These are: Four-ball lubrication tester which can operate in two modes – sliding or rolling pitting fatigue; pin-on-disc machine, which provides for a wide variety of load/speed conditions as well as testing dissimilar materials; a rolling element bearing test stand, which tests a pair of angular contact bearings. Various wear mechanisms are recreated to simulate the load/speed conditions experienced in military engines and gearboxes [2].

2.1 FOUR-BALL MACHINE

Originally designed as a lubricant tester, the four-ball machine was used to generate wear debris under known conditions (load, speed, lubrication regime, etc.). It was operated in two modes of motion: sliding, as detailed in IP Standard 239/97 [3] and rolling pitting fatigue, in accordance with IP

Standard 300/75 [4]. For the former operation the cup was set up as shown in Figure 1. The lower three balls were clamped in a triangular configuration while the top ball was clamped in a chuck and rotated in contact with the lower three balls. Load was provided via a cantilever from below the cup and the contact was lubricated by filling the assembly with the selected oil to a prescribed level, approximately 7cc by volume. Tests were normally conducted for 60 seconds after which the assembly was taken apart and the oil decanted into a sample bottle. To ensure no wear elements remained, the parts in contact with the oil were also washed with solvent, which was collected in the same bottle.

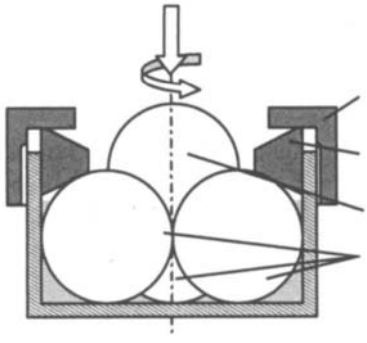


Figure 1 Sliding four-ball test

Rolling pitting fatigue tests using the same machine were conducted differently in that the lower balls were spaced apart from one another and were free to rotate in a race (Figure 2).

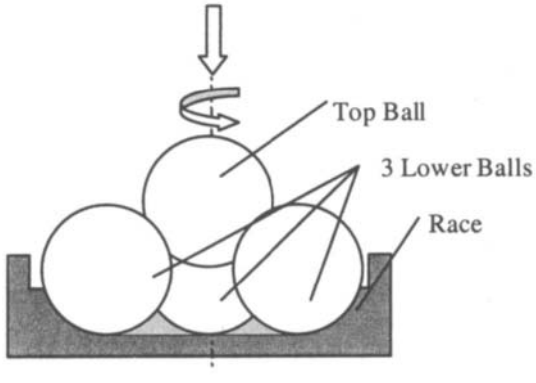


Figure 2 Rolling four-ball test

These experiments were deemed to have failed when a pit or spall greater than 1mm^2 occurred on the upper ball, and which was also characterised by a large increase in the noise and vibration. The time to failure was not predictable before the test was started but would typically last between ten minutes and three hours, with an average test length of approximately thirty minutes.

On completion of the test, which included tests, which had purposely been stopped prior to failure, the oil was collected from the cup and the wear products were washed into a sample bottle.

2.2 PIN-ON-DISC MACHINE

The pin-on-disc machine is shown in Figure 3. Its basic configuration comprises a steel disc (typically mild steel, EN 8 or gear steel, EN 34) and a 12.7mm diameter ball bearing (52100). Load is applied using dead weights on a cantilever and the motion is provided through a motor that spins the disc up to 1000 rpm giving a maximum linear speed of approximately 8 m/s [5].

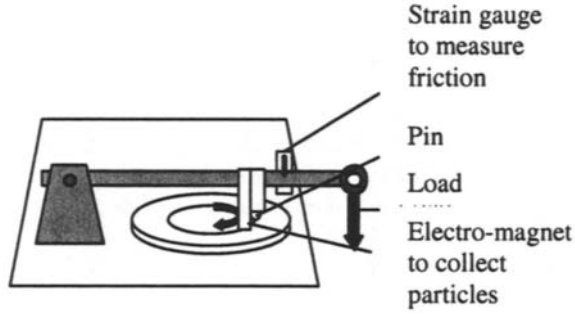


Figure 3 Pin-on-disc machine

In a similar manner to the four-ball test, the oil used to lubricate the contact was collected and the various parts washed with solvent to ensure maximum capture efficiency of the wear particles.

2.3 BEARING TEST STAND

A pair of magneto bearings is loaded via a cantilever as shown in Figure 4. Magneto bearings are used because the inner and outer races readily come apart. This is essential for fitting to the test head, thereby eliminating any damage that might occur if the outer race was partly machined or the bearing forced apart.

The test load applied is a maximum of 1040 kg at a speed of 1000 rpm applied through a belt from a 1.5 kW three-phase electric motor. Lubrication is supplied through the top of the test head (not shown) and output at either end of the test head; an oil bath level is established coincident with the centre of the lowest ball. With an oil sump of 11 litres and a flow

rate of 100ml/min there is ample time for the oil to cool to room temperature before being re-circulated.

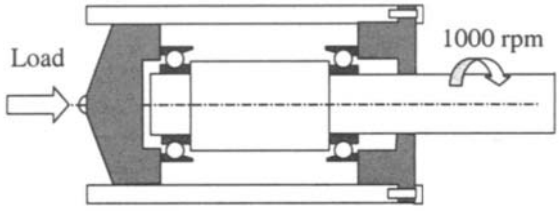


Figure 4 Bearing test head

2.4 WEAR DEBRIS ANALYSIS

Once the various samples of oil and debris have been collected it is desirable to view them with a metallurgical microscope. To present the particles in an acceptable manner, the sample bottle is drawn through a 0.8 mm Millipore Isopore® filter using a standard filter assembly with vacuum pump. This particular filter is used because it is transparent when examining the particles in transmitted lighting conditions that can yield silhouette images useful for computer-based analyses.

Images of the particles are captured using a JVC CCD camera mounted on top of either an Olympus BX60 metallurgical microscope or an Olympus SZX 9 stereo zoom microscope. The latter is used for objects greater than 1 mm chord length. This in turn is connected to a computer which captures several images in different focal planes and collaged them together to create a single in-focus picture [6].

3. TEST RESULTS

Results for the four-ball machine fall into two categories, which correspond with the two principal modes of operation.

Sliding Four-Ball Test Results

The wear achieved from a sliding wear test is in the form of an annulus on the top ball and a worn circle on each of the lower three balls. A standard method of assessing the severity of wear is to measure the wear scar diameters of the lower balls and examine how this varies with increasing load. It can be seen from Figure 5 there was a sudden transformation in the wear regime at the 110 kg load. This is known as the transition region, in which the wear behaviour below is known as mild wear and above as severe wear. When the contact experiences mild wear the oil film between the contacts is intact and mainly elasto-hydrodynamic lubrication is present. When severe wear occurs, the oil film has broken down thereby creating severe mixed or boundary lubrication conditions and associated metal-on-metal contact.

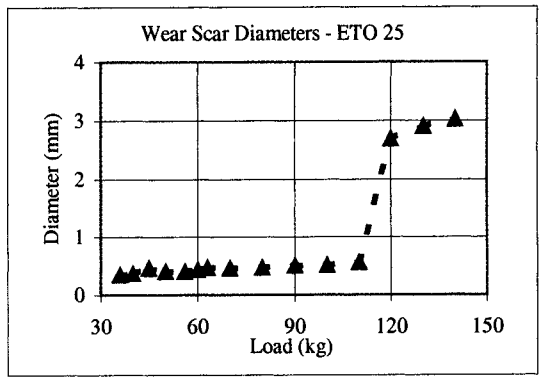


Figure 5 Wear scar diameter

When the two modes are compared in terms of the friction forces generated, the seizure period is manifest through a large increase in the coefficient of friction.

Typical bottom ball severe wear scar and upper wear annulus are shown in Figure 6 and Figure 7. The direction of rotation of the top ball is clear as are the irregularities at the trailing edge of the scar that are akin to the wear particles found in oil bath.

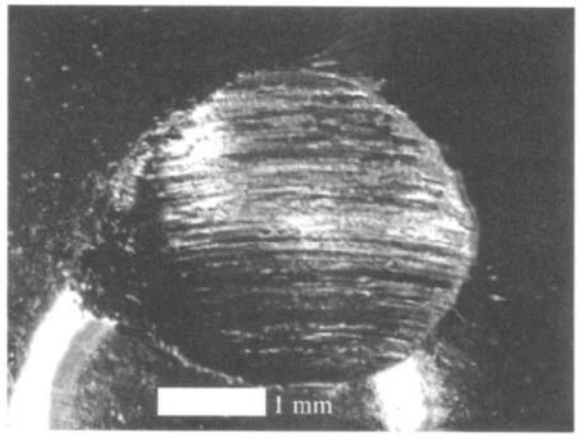


Figure 6 Wear scar - sliding test

When examining the wear particles under a microscope, two distinct types are observed – mild and severe sliding, which correspond with the two regimes of lubrication. Figure 8 shows a representative example of a mild wear particle. It is typically less than 50 μm in chord length, exhibiting few surface features with a low aspect ratio. These

particles are the result of the imperfections on the ball surface being removed in the contact by abrasion.

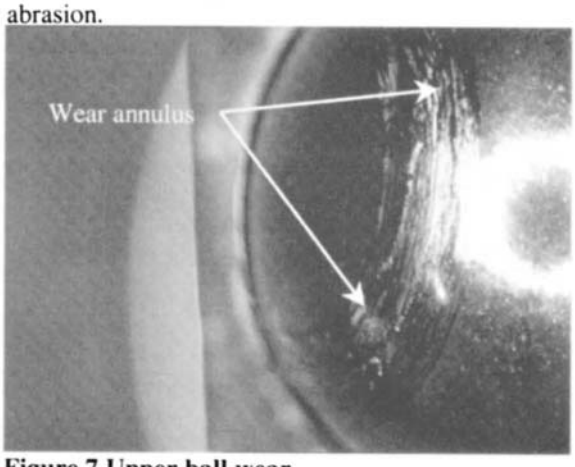


Figure 7 Upper ball wear

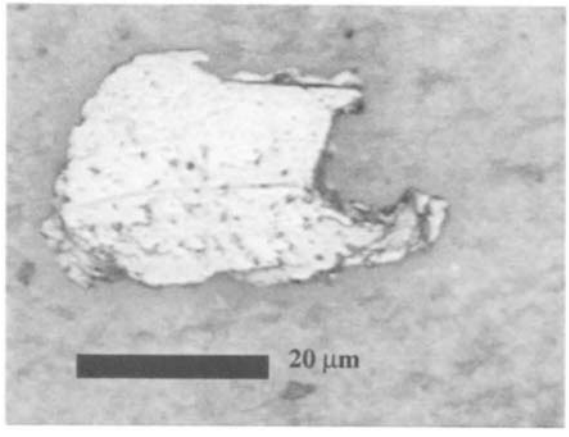


Figure 8 Mild wear particle

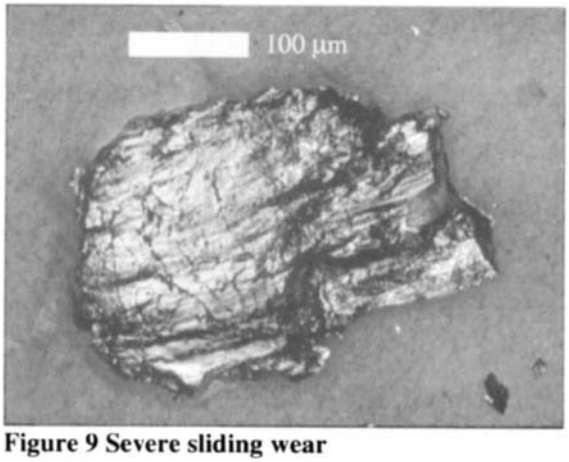


Figure 9 Severe sliding wear

Figure 9 shows a typical severe sliding particle captured from the severe wear regime. Much larger than the mild wear particles, the surface displays striations, corresponding with those found on the wear scar, and is 'chunky', i.e. a thicker particle rather than the plate-like mild wear. Another feature that may occur is temper colouring, which indicates the heightened temperatures occurring in the contact region due to flash temperature behaviour.

Severe sliding debris does not occur in isolation. On average, tests lasted for 5 – 10 seconds before the breakdown of the lubricant film and the onset of seizure. During this period mild wear particles are generated. Also, after seizure, they continue to occur at the edge of the wear scar, where lower contact pressures are experienced.

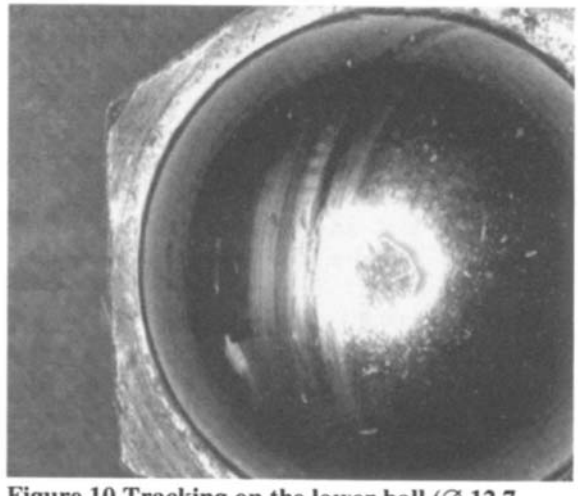


Figure 10 Tracking on the lower ball (∅ 12.7 mm)

Rolling Four-ball Results

Rolling pitting fatigue may be bought as a result of the lower balls 'tracking' which is the action of the balls spinning about a single axis point instead of rotating randomly, which in turn leads to cyclical loading. This can result in a dark track on the lower ball as shown in Figure 10.

Figure 11 shows a spall on the upper of ball. When the experiment was allowed to continue beyond the first signs of failure, the spall increased in size forming, from the initiation point, around the annulus and eventually resulted in a large portion of the upper ball parting company at which point the test was stopped to protect the test apparatus.

Subsurface cracking and surface cracking are the consequence of fatigue loading and lead, eventually, to the formation of a spall. Figure 12 is a prepared cross-section that shows cracks that occurred from a failed upper ball (i.e. a ball that had formed a spall).

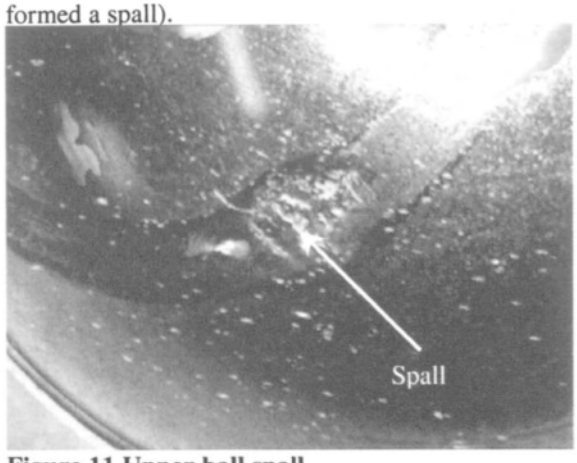


Figure 11 Upper ball spall

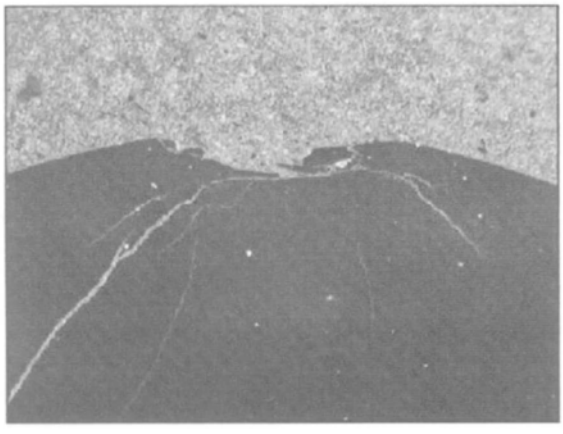


Figure 12 Cracking near fatigue spall

Fatigue debris exhibits some very distinct characteristics, as can be seen in Figure 13. Cracking radiating from the edge inwards to give a petal-like appearance, and surface pitting, are classic features indicating a pitting failure. They are generally greater than 100 μm in chord length and, as in the case of the severe sliding wear particles, they occur together with a population of mild wear debris.

Pin-on-Disc Results

Results from the pin-on-disc machine can be compared with the four-ball sliding test due to the similarity in the nature of the contact. Again, the two distinct phases are seen, mild sliding, which is primarily abrasive wear, and a severe wear phase that is dominated by adhesive wear. The mild wear regime can be further broken down, into wear that causes plastic deformation of the disc and wear that does not. The following series of figures (Figure 14 – Figure 16) shows a series of three wear tracks under the same lubrication, material and geometric conditions but at different loads.

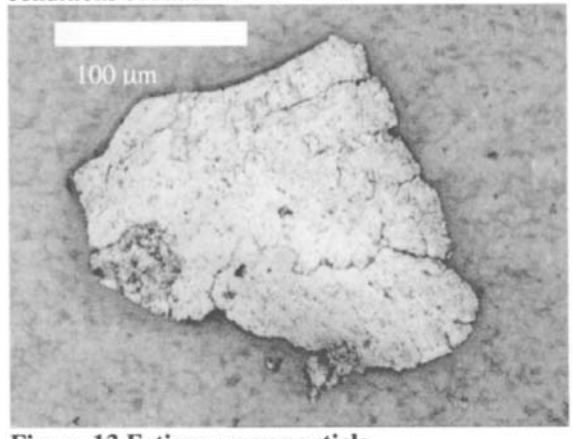


Figure 13 Fatigue wear particle

The transfer of material to and from the pin (in this case a ball contact) can be seen by the additional material being deposited on the disc, and also by examining the ball, which shows very similar features. This differs from the mild regimes, which exhibits mainly wear of the disc.

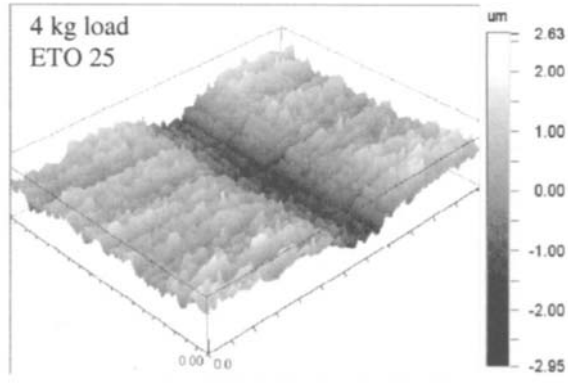


Figure 14 Mild wear – no plastic deformation

Previous work on machines of this type has shown a marked increase in flash temperature rise corresponding with the regime change.

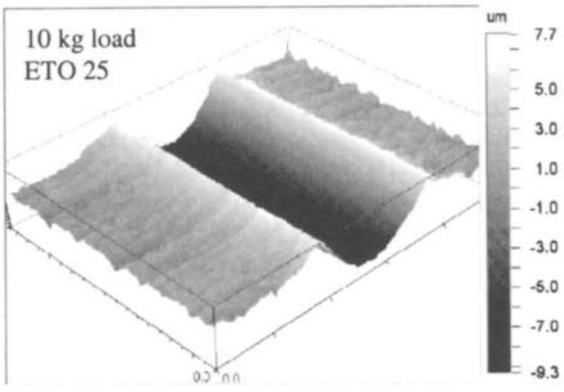


Figure 15 Mild wear with plastic deformation of the disk

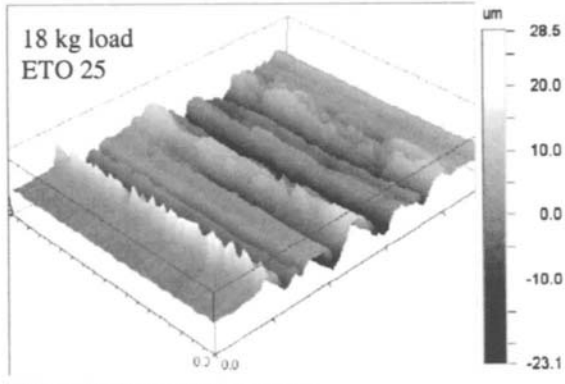


Figure 16 Severe disk wear

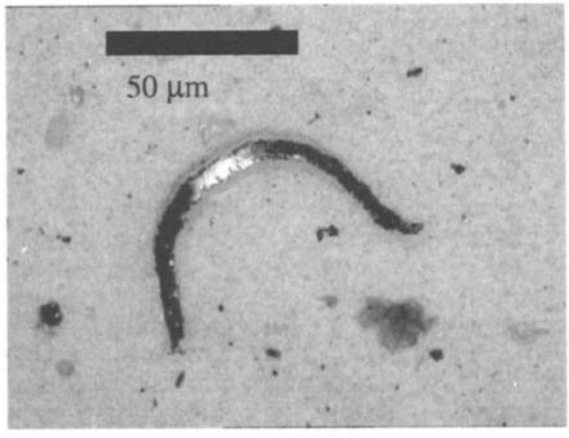


Figure 17 Cutting wear particle

Three types of debris were generated; mild and severe sliding, as previously discussed, and also cutting wear. Figure 17 shows a cutting wear particle, which takes the appearance of swarf from a lathe machine. It can be generated in one of two ways. Two body cutting wear is created when one surface is harder than the other with the harder surface ploughing the surface, as might occur with a hard steel on soft steel, but also from the work hardening of one of a pair of like contacting metals. Cutting wear can also occur from three body abrasion by the entrainment of hard material into the contact, such as sand or work-hardened wear particles being re-entrained.

Bearing Test Stand Results

More akin to the bearing contacts found in industrial situations, the bearing test stand takes longer to fail than the previous apparatus, typically, seven hours. Also the failure is not as defined as in four-ball pitting fatigue tests, although the mode of failure is very similar.

Figure 18 shows the inner race of the failed bearing. The outer race, and the inner and outer race of the other test bearing, showed no visible signs of wear.

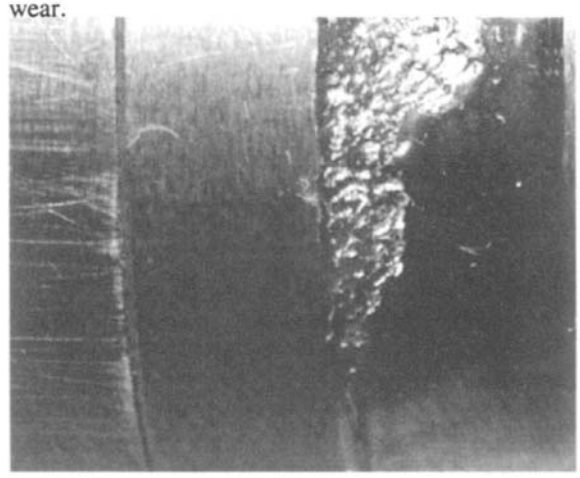


Figure 18 Initiation site of bearing failure

To monitor the wear situation over time it is necessary to take samples at various intervals throughout the test period.

4. DISCUSSION

From the results of the various tests discussed in this paper, a number of common properties emerge which may have an affect on the prediction of remaining useful life of a machine. Where a contact is primarily experiencing a simple sliding condition the wear life is more easily predicted provided that a full lubrication film is maintained. A relaxation of

this regime will result in a sudden change to a severe, adhesive wear mode with metal transfer and excess wear. Rolling contacts, however, are harder to predict by examination of the lubrication regime alone.

In addition to selecting load, speed and lubricant properties comparable to those found in the military engines and gearboxes, it was necessary to compare the wear and wear debris with examples from actual aircraft. Debris taken from various aircraft of the Royal Air Force was examined in the same manner as the test debris and was found to consist of the four main particle groups discussed in this paper. Using a combination of both laboratory and flight data a picture of the wear history of a particular platform can be established – the former giving more examples of failure whilst the latter contains information on the engines' life history.

Exceptions were found, for example, Rolls Royce RB 199 engines from the Tornado contained many particles from the abradable nickel seals. In instances such as these, it is imperative that the exceptions are included in future analyses.

Future work in this subject will include gathering data on damage tolerance of bearings which can often perform satisfactorily despite some initial indication of failure from either wear debris, vibration or acoustic emission sensors. This will include continuously monitoring a system using a variety of sensors and fusing the data to give a more accurate picture of the component life.

5. CONCLUSIONS

Continuous monitoring of the various wear related parameters will lead to a more accurate prediction of remaining useful life of a component once the wear has been thoroughly examined and trends identified. This can be achieved in part by utilising laboratory tests, using similar load, speed and lubricant conditions to those found in the machinery under investigation. By using tests that fail in different modes, a wear debris history of the life of a component can be determined which can lead to this information being used in real machinery failure situations.

The four main types of wear debris seen in bearings and gearboxes are: mild sliding, severe sliding, rolling pitting fatigue and cutting wear.

The ultimate aim of this research is to yield a method for predicting the remaining useful life of any given component, engine or gearbox.

ACKNOWLEDGEMENTS

The authors would like to thank the US Office of Naval Research, the Royal Air Force and QinetiQ for sponsoring this work and also for the assistance of Dan Price for Figure 12.

REFERENCES

1 Dowson D, 'History of Tribology', Professional Engineering Publishing, 1998

2 Sperring TP, Tucker J, Reintjes J, Shultz A, Lu C and Roylance BJ, 'Wear Particle Imaging and Analysis – a contribution towards monitoring the health of military ships and aircraft.' International Conference on Condition Monitoring, Swansea, 1999, pp 539-545.

3 IP Standards for Petroleum and Its Products, Part 1 'Extreme Pressure and Wear - Four-ball', Vol. 1, pp 239.1-239.16, Institute of Petroleum, London, John Wiley and Sons, 1986.

4 IP Standards for Petroleum and Its Products, Part 1, 'Rolling Contact Fatigue Test for Fluids in a Modified Four-Ball Machine', Vol. 2, pp 300.1-300.11, Institute of Petroleum, London, John Wiley and Sons, 1986.

5 Roylance BJ, Sperring TP and Barraclough TB, 'Bench test determination of wear modes to classify the morphological attributes of wear debris', ASTM Conference on Bench Testing in Seattle, USA, 2000

6 Sperring TP, Roylance BJ, Nowell T, and Hodges D, 'From research to application – The development of the wear debris classification system SYCLOPS', International Conference on Condition Monitoring, Oxford, UK, 2001, pp 363 – 373.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Life prediction for surface crack initiated contact fatigue of silicon nitride bearing balls

Y. Wang and M. Hadfield

School of Design, Engineering and Computing, Bournemouth University
12 Christchurch Road, Bournemouth BH1 3NA, UK

Technical ceramics as materials for rolling contact bearing components show some practical advantages over traditional bearing steels. The properties of ceramics, specifically low density and high stiffness, are of most interest to gas turbine and machine tool manufacturers. High hardness, low coefficient of thermal expansion and high temperature capability are properties also suited to rolling element materials. Silicon nitride has been found to have the optimum combination of properties suitable for this application. Much research over the past two decades on its structure, quality control and manufacturing techniques has produced material which can seriously be considered for rolling contact bearing design. However, the difficulties of both sintering and machining the material may result in surface crack defects. It is difficult to detect surface cracks during high volume production processes and, hence there is an important need to understand their influence and fundamental mechanism of the failures they cause. In the present study, a quantitative three-dimensional boundary element model has been developed which can be used to determine the geometry of acceptable surface defects and to predict rolling contact fatigue life. The concentrated contact analysis focuses on a hybrid ceramic/steel combination. A modified four-ball machine is employed to validate the predictions. Results from the experimental test are in good agreement with the results from the numerical analysis.

1. INTRODUCTION

Technical ceramics as materials for rolling contact bearing components show considerable advantages over traditional bearing steels. The properties of ceramics, specifically low density and high stiffness, are of most interest to gas turbine [1] and machine tool manufacturers [2]. High hardness, low coefficient of thermal expansion and high temperature capability are properties also suited to rolling element materials. Much research over the past two decades on structure, quality control and manufacturing techniques has produced material which can seriously be considered for rolling contact bearing design. This is especially true for hybrid ball bearings, i.e. precision angular contact ball bearings using ceramic rolling elements are now offered as standard components within the ball bearing manufacturing industry.

One major factor limiting widespread use of ceramics in advanced industrial application is the presence of defects on the surface of the ceramics. Defects include cracks and inclusions, and can be introduced in each stage of the manufacturing process. Each type of defect contributes to the reduction of the ceramic's structural integrity and leads to variability in its material properties, both of which can be unacceptable in industrial applications. Cracks that exist on the surface are particularly critical to rolling contact elements since operational stresses in this region of the ceramic component are greatest. Therefore, flaws in the vicinity of the surface can compromise the strength of the ceramic component, reduce performance, and shorten

lifetime. These flaws can be introduced in any stages of fabrication. Surface cracks formed prior to, or during, the final lapping or polishing stage are particularly difficult to detect [3]. It is now well recognised that existing cracks decrease the fatigue life performance. Numerous experimental studies have confirmed that the balls typically fail by crack propagation from the existing surface cracks. How surface cracks affect the rolling contact fatigue of ceramics is still not fully understood. Hence there is an important need to assess their influence on fatigue failure.

Although the life prediction of rolling elements has been developed extensively, most have been applied to bearing steels [4-6]. The fundamental premise is that failure originates from subsurface defects within the material subjected to alternating shear stress. For ceramic bearing elements, the ceramic balls typically fail by crack propagation from existing surface cracks [7-14]. Failure modes of ring, radial and lateral pre-cracked ceramic elements have been studied using a modified four-ball machine [7-9, 15-16]. Experimental results revealed that the artificially produced cracks failed in a non-catastrophic way and that the failure mode was strongly influenced by the lubricating fluid. Recent studies [11, 12] on 'natural' ring cracks showed that the 'natural' ring crack initiated failure depends on location/orientation of a crack within the contact path. The spalling fatigue failure only occurs at a few crack positions.

Many efforts have been completed to relate crack propagation rate with wear performance of materials [17-21]. A fracture mechanics approach can offer a significant potential advantage over the statistical approach in accurately predicting bearing element life. It is critical to develop a more widely applicable understanding of fatigue failure from surface crack defects and to determine the crack geometry for a given set of design conditions using the fracture mechanics criteria. To provide improvement over empirical methods used in engineering design the present paper will use boundary element analysis to predict the fatigue life of rolling elements which contain surface defects. In the present study, a quantitative three-dimensional boundary element model has been developed which can be used to determine the geometry of acceptable surface cracks and to predict rolling contact fatigue life. A modified four-ball machine is employed to validate the predictions.

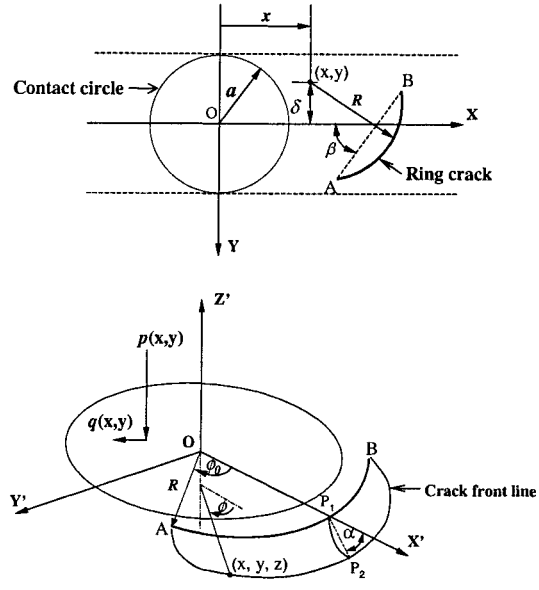


Figure 1. Co-ordinate system of a surface crack subject to rolling contact.

2. ANALYTICAL MODEL DESCRIPTION

2.1. Loading configuration

A surface loading traverses the surface of the silicon nitride ball. This cyclic load moves across the surface crack, repeatedly, as shown in figure 1. For the purpose of the calculation, the surface loading is specified in terms of an assumed contact stress distribution. Assume that normal pressure $p(x,y)$ and tangential traction $q(x,y)$ within the contact region are expressed by the following equations:

$$p(x,y) = p_0\left(1-\frac{x^2+y^2}{a^2}\right)^{1/2} \tag{1}$$

$$q(x,y) = fp(x,y) \tag{2}$$

where p_0 is maximum normal pressure, a is the radius of the contact circle and f is the friction coefficient of Hertzian contact. It is assumed that $f < 0$ when the contact circle passes over the ring crack (arc AB) from right to left and $f > 0$ when passing from left to right. The co-ordinate x measures the distance from the centre of the contact circle to the centre of the ring crack circle and the stress intensity factors will be changed as the distance x changes.

As the contact circle passes over the ball surface, the crack front will pass from compressive to tensile stress continuously. When the crack lies in the compressive region ($-2a < x < 0$), the crack is suppressed ($K_I \leq 0$). Consequently, the crack front will not propagate along the main growth path in the compressive region. However, subsequent surface cracks may occur due to the pre-existence of ring cracks, and the branch cracks on the crack faces may also occur due to the crack face contact. The co-ordinate x affects the subsurface behaviour, which may be a rather complex process due to Hertzian loading in a cyclic fashion.

2.2. Crack geometry

The geometry of a crack is very important to fracture mechanics analysis and will be considered carefully. Based on the experimental study, the surface crack is assumed to have a cone shaped crack face with a curved line as the crack front. The position of an element on the crack surface in figure 1 may be written as

$$x = (R + l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha)\cos\phi \quad (3)$$

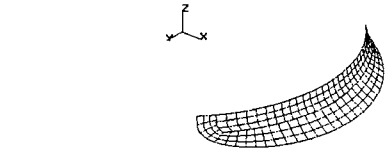
$$y = (R + l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha)\sin\phi \quad (4)$$

$$z = -0.224\,(l\cos\alpha\sqrt{1 - (\phi / \phi_0)^2})^{1/2} \quad (5)$$

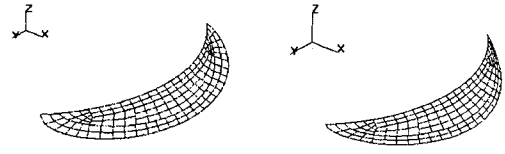
where l denotes the crack depth, $0 < l \leq$ maximum crack depth (line P_1P_2), R is the radius of the ring crack, α is the angle of line P_1P_2 to the ball surface, $0 < \alpha < \pi/2$, ϕ represents the angle of an element position on the crack face, $-\phi_0 \leq \phi \leq \phi_0$ and ϕ_0 denotes the half angle of the ring crack arc AB, $0 < \phi_0 < \pi/2$.

According to the experimental observations, the dimension of an original ring crack was defined as crack radius R = 0.21 mm, crack angle $\alpha = 50^\circ$, crack depth l = 0.05 mm, and half angle $\phi_0 = \pi/4$.

The influence of the crack geometry change on crack propagation is also considered and three extra crack geometrical shapes are adopted to study their effects. Figure 2 presents a detailed view of the selected crack geometry.



(a) Geometry I



(b) Geometry II (c) Geometry III

Figure 2. Detailed views of various crack shapes.

For the geometry I, the co-ordinate system has been shown in figure 1 and the position of an element on the crack face has been presented in the equations (3) to (5). For the geometry II, the position of an element on the crack surface can be described by the following equations:

$$x = (R + l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha)\cos\phi \quad (6)$$

$$y = (R + l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha)\sin\phi \quad (7)$$

$$z = -l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha\tan\alpha \quad (8)$$

For the geometry III, the position of an element on the crack surface can be described by the following equations:

$$x = (R + l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha)\cos\phi \quad (9)$$

$$y = R\sin\phi \quad (10)$$

$$z = -l\sqrt{1 - (\phi / \phi_0)^2}\cos\alpha\tan\alpha \quad (11)$$

2.3. Stress intensity factor calculations

2.3.1. Boundary element model

In general, the geometry and loading encountered in three-dimensional crack problems is too complex for the stress intensity factors (SIF) to be solved analytically. The SIF calculation is further complicated because it is a function of the position along the crack front, the crack size and shape, type of loading and geometry of the structure. Therefore, in this study the current available numerical analysis tools are used to perform a fracture mechanics analysis.

A boundary element method is utilised in the present study. There are several ways to model a crack in boundary element modelling. The dual boundary element method (DBEM) is used for the crack growth analysis. The dual boundary element method [22, 23] considers two independent equations: the displacement and traction boundary integral equations, with the same integration path for each pair of coincident source points. The crack is represented by two elements occupying the same physical location, each element representing a face of the crack. The displacement and traction boundary integral equations are on the crack surface. The use of dual elements for 3-D crack growth means that only one surface of the crack needs to be defined.

Because of the continuity requirements of the displacements and tractions for the existence of the traction boundary integral equations and the coplanar characteristic of the crack surface, special consideration has to be given to modelling discretization. Discontinuous quadratic elements are used for the crack modelling.

A discontinuous element plays an important role in the boundary element method because the problem variables are not forced to be continuous across the elements. The major benefit of discontinuous elements is their ability to model discontinuous stress results. They are therefore very useful in fracture mechanics analysis for modelling the stress behaviour at a crack front.

2.3.2. Stress intensity factors (SIF)

Since the importance of the stress intensity factors was demonstrated in determining crack-tip stress fields in two-dimensional problems [24], many different methods have been devised for obtaining SIF [25].

In this study, the stress intensity factors are computed using the crack opening displacement method. When one point formulae are employed, the Mode I, II and III stress intensity factors are evaluated using the following equations [25]:

$$K_I^{P'} = \frac{E}{4(1-\nu^2)}\sqrt{\frac{\pi}{2r}}\left(u_b^P\,|_{\theta=\pi} - u_b^P\,|_{\theta=-\pi}\right) \qquad (12)$$

$$K_{II}^{P'} = \frac{E}{4(1-\nu^2)}\sqrt{\frac{\pi}{2r}}\left(u_n^P\,|_{\theta=\pi} - u_n^P\,|_{\theta=-\pi}\right) \qquad (13)$$

$$K_{III}^{P'} = \frac{E}{4(1-\nu^2)}\sqrt{\frac{\pi}{2r}}\left(u_t^P\,|_{\theta=\pi} - u_t^P\,|_{\theta=-\pi}\right) \qquad (14)$$

where the displacement u^P is evaluated at point P as shown in figure 3; u_b^P u_n^P and u_t^P are projections of u^P on the co-ordinate directions of the local crack front co-ordinate system and $\theta = \pi$ and $\theta = -\pi$ denote upper and lower crack surfaces respectively. $K_I^{P'}$, $K_{II}^{P'}$ *and* $K_{III}^{P'}$ are approximations of stress intensity factors corresponding to the point P' along the crack front and on a normal line to the front.

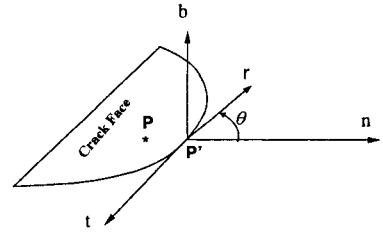


Figure 3. Local crack front co-ordinate.

3. CALCULATION RESULTS

3.1. Influence of crack geometry

Three types of crack surfaces (figure 2) are adopted to study the influence of the crack surface shape on the stress intensity factors. For the purpose of the calculation comparison, assume that these cracks have the same crack angle ($\alpha = 50^o$, crack length (one forth of the ring crack circle) and crack depth (l = 0.05mm). The crack location for the

calculation is $\beta = 90°$, $\delta = 0$, and $x = 0$. The maximum contact stress is assigned to a value of 5.58GPa, and friction coefficient f = -0.05.

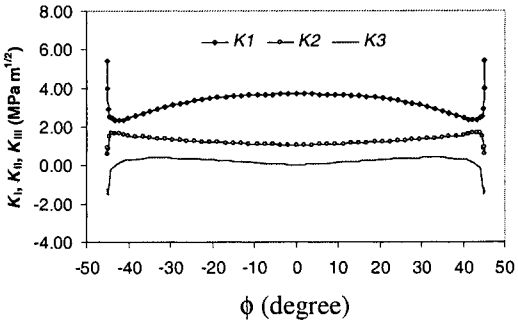


Figure 4. Stress intensity factors plot (geometry I).

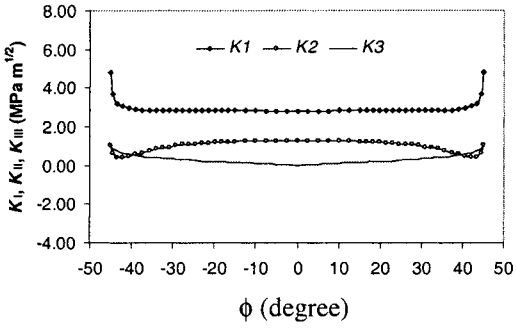


Figure 5. Stress intensity factors plot (geometry II).

Figure 4 shows the calculated results of the stress intensity factors. As can be seen from the plots, the SIF solutions are changed as the crack geometry changes. Mode I stress intensity K_I is always higher than mode II and III stress intensities. It is reasonable to assume that the mode I stress intensity will dominate the crack growth behaviour. For certain geometry, the K value is a function of ϕ. For the geometry I, K_I value at $\phi = 0°$ (lower point) is about 3.5 MPa m $^{1/2}$ and gradually decreases as ϕ changes until nearly $\phi = \pm 44°$, and rises up 5.5 MPa m $^{1/2}$ quickly as ϕ approaches to $\pm 45°$ (surface point). K_{II} and K_{III} vary slightly ϕ except $\phi = \pm 45°$ (surface point), in which the values of K_{II} and K_{III} drop down sharply.

The calculation result for geometry II is shown in figure 5. At $\phi = 0°$(lower point) the mode I stress intensity K_I is about 3.0 MPa m $^{1/2}$ and stays in the same value until nearly $\phi = \pm 40°$, and rises quickly to 5.0 MPa m $^{1/2}$ as ϕ approaches $\pm 45°$ (the surface point). There are slight variations in K_{II} and K_{III} value.

The calculation result from geometry III is shown in figure 6. The mode I stress intensity K_I is about 3.0 MPa m $^{1/2}$ at $\phi = 0°$ and nearly no change until nearly $\phi = \pm 30°$, and rises gradually to 4.7 MPa m $^{1/2}$ as ϕ approaches $\pm 45°$. The K_{II} and K_{III} vary slightly with ϕ.

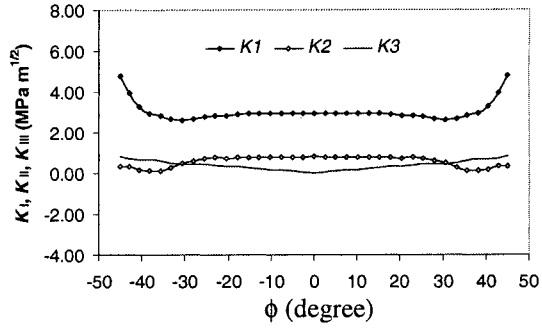


Figure 6. Stress intensity factors plot (geometry III).

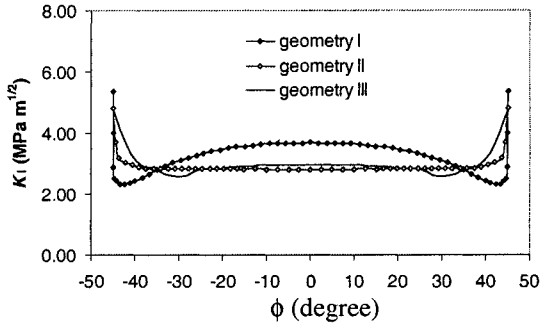


Figure 7. Influence of crack geometry on K_I.

Figure 7 shows the plots of K_I for the geometry I, II and III. The K_I for the geometry I is higher than that of geometry II and III. There is nearly no difference between geometry I and II. The results indicate that the mode I stress intensity K_I plays an important role in surface ring crack propagation. Consequently, K_I may be a main concern in determining the crack propagation behaviour. For HIP silicon nitride bearing material, the fracture

toughness K_{IC} is around 6.0 MPa $m^{1/2}$. The threshold ΔK_{th} of crack propagation ranges from 2 to 3 MPa $m^{1/2}$. The calculation results indicate that there is no possibility of unstable fracture. However, the crack front will propagate in the calculated geometry. The crack growth behaviour will be changed slightly due to the crack geometry effect.

3.2. Influence of normal pressure

Figure 8 shows the plots of the K_I variations caused by normal contact pressure. The geometry III is employed. Phenomena of contact radius varying with contact pressure bring some difficulties into the calculation. To compare precisely, the ring crack is placed at the same distance (0.03mm) away from the contact circle. As can be seen from the plots, K_I increases with increasing contact stress, and reaches a maximum value at $\phi = \pm 45°$. The result reveals that there is no possibility of unstable fracture due to the calculated K_I less than K_{IC} (6.0 MPa $m^{1/2}$). Crack propagation may happen if contact stress is more than 5 GPa due to the threshold value ΔK_{th} = 2-3 MPa $m^{1/2}$. If contact stress is less than 4.0 GPa, the spalling fatigue will not happen. These predictions are consistent with the experimental observations.

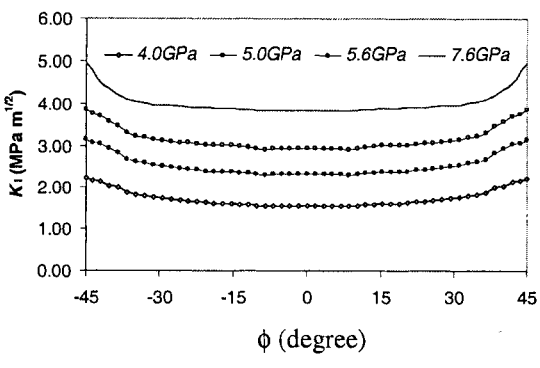


Figure 8. Influence of contact stress on K_I.

3.3. Influence of tangential traction

The influence of surface tangential traction on crack growth is calculated. The different friction coefficients are used to describe the magnitude and direction of the tangential traction. The calculation result is shown in figure 9. As can be seen from the plots, friction coefficients have a significant influence on the stress intensities. Increasing traction force increases the magnitude of the stress intensity factors. It can be concluded that the fatigue life performance decreases as the tangential traction increases.

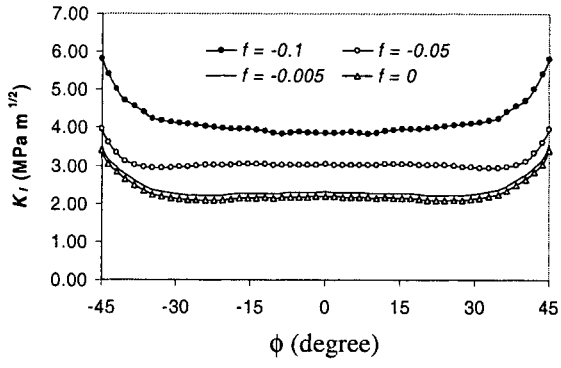


Figure 9. Influence of friction coefficients on K_I.

3.4. Life prediction of fatigue failure

Figure 10(a) is a plot of the stress intensity factor K_I as a function of crack extension. The results indicate that crack propagation will be arrested when a certain extension size is reached. This may explain why the rolling contact failure mode of silicon nitride is non-catastrophic. The computations are based on an assumed crack path derived from the equations (3) to (5). The contact stress is 5.58GPa, and friction coefficient f = -0.05. The aim is to relate the stress intensity factors to crack propagation. The calculation of stress intensity factors presented here, while approximately in the plane of $\phi = 0°$, may be useful in fatigue life prediction.

To determine what spall size may be reached at the fatigue failure, we will use the Paris equation

$$da/dn = c(\Delta K)^m \qquad (15)$$

by substituting $K_I = -19.181x + 4.5475$ from figure 10(a) to solve fatigue cycles requested for a certain crack size. Fatigue cycles can be obtained by integrating equation (15), which is expressed as follows:

$$n = \int_{a_0}^{a_c} \frac{da}{c(\Delta K)^m} = \int_{a_0}^{a_c} \frac{da}{c(4.5475 - 19181a)^m} \quad (16)$$

Integration of equation (16), then

$$n = \frac{1}{19181c(m-1)}(4.5475 - 19181a_c)^{-m+1} - \frac{1}{19181c(m-1)}(4.5475 - 19181a_0)^{-m+1} \quad (17)$$

The constants in the Paris law of equation (16) are obtained from the fatigue crack propagation studies. The value for c is 1.01×10^{-21} and the exponent m for HIP material is 18. Fatigue cycles for the different crack extensions are shown in figure 10(b). The calculated fatigue cycles indicate that the original ring crack propagation is able to reach the dimension normally observed (0.3 to 0.4 mm in diameter).

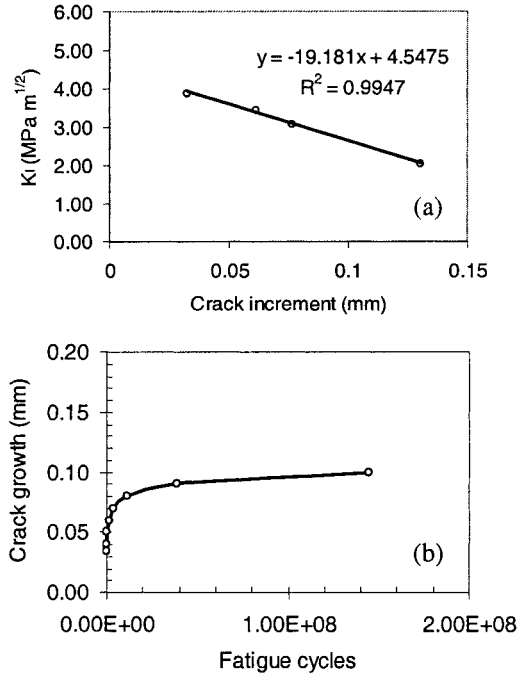


Figure 10. (a) K_I versus crack increment,
(b) Crack growth versus fatigue cycles

4. EXPERIMENT VERIFICATION

4.1. Experiment methods

The rolling contact fatigue tests were performed using a modified four-ball machine. This machine was employed as it simulates rolling elements and precisely defines the contact load. It consists of an assembly that simulates an angular contact rolling element bearing. The stationary steel cup represents a bearing outer-race, three lower balls represent the rolling elements within a bearing-race and the upper ball represents the inner-race. The upper-ball was assembled to a drive shaft via a collet and contacts three lower-balls when the machine is stationary. The contacting positions between the upper ball and lower balls were immersed in lubricating oil. The tests were terminated automatically at a set number of drive-shaft revolutions measured by a tachometer. A vibration sensor automatically stopped tests at a pre-determined potentiometer sensitivity and light emitting diode duration.

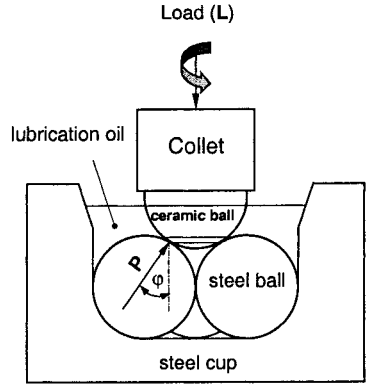


Figure 11. Loading configuration of the test machine.

The loading configuration of the test machine is shown in figure 11. The contact load is given by

$$P = \frac{L}{3\cos\varphi} \quad (18)$$

where P is contact load, L is applied load (shaft load), φ is contact angle (φ = 35.3°). The maximum contact pressure and contact radius are calculated using the following expressions [26]:

$$p_0 = \left(\frac{6PE^{*2}}{\pi^3 R^{*2}} \right)^{1/3} \tag{19}$$

$$a = \left(\frac{3PR^*}{4E^*} \right)^{1/3} \tag{20}$$

$$E^* = \left(\frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \right)^{-1} \tag{21}$$

$$R^* = \left(\frac{1}{R_1} + \frac{1}{R_2} \right)^{-1} \tag{22}$$

where p_0 is maximum contact pressure, P is contact load, a is contact radius, $E_{1,2}$ and $\nu_{1,2}$ are the Young's modulus and Poisson's ratios of the ceramic ball and steel ball, and $R_{1,2}$ is the radius of the ceramic ball and steel ball respectively.

In this study, the upper ball was a ceramic ball with a surface ring crack. Since fatigue performance considerably depends on the crack orientation within the contact path, the crack position was taken into full account and the crack was positioned in terms of the experimental procedure [11]. All of the tests were conducted at various Hertzian contact pressures. The upper ball speed was set at 5000 rpm. The test will continue until failure or expiry of the suspending time (180 hours). Two lubrication oils are used in the present study. One is standard turbine oil and another is traction oil. The use of the traction oil is to increase the tangential traction in order to verify the prediction. Fatigue cycles are recorded from the ceramic upper ball and tests were stopped at high vibration levels. The test balls are selected carefully to ensure the crack sizes are as the same as possible. In general, crack geometry (crack depth and crack angle) is very close if the radius and length are the same. The radius of the natural ring cracks, commonly found on the commercial ceramic ball surface, is in a range of 0.2 to 0.25, and the depth is about 0.05 to 0.06 mm.

Silicon nitride balls tested in this study were procured from specialist manufacturers. They were manufactured by a hot isostatically pressed (HIP) method. Ball blanks were ground and polished to 12.7 mm diameter in this case by a manufacturer; standardised procedures were adopted to ensure consistent quality of material and geometry. The elastic modulus and Poisson ratio were 310 GPa and 0.28 respectively. Average roughness (R_a) of the silicon nitride ball surfaces was 0.01μm and ball roundness was within ball bearing tolerances. The steel ball was carbon chromium steel with a 12.7mm diameter and a surface roughness (R_a) of 0.02 μm. The hardness of the silicon nitride balls was on average 1637 Hv. The hardness of the steel ball was 839 Hv.

4.2. Experiment results

Figure 12 shows the rolling contact fatigue (RCF) life at different test conditions. It is clear that the RCF life decreases with the increase of the maximum contact pressure. Also, the rolling contact fatigue performance is considerably influenced by the lubricant properties. For the traction oil, the RCF life is much less than that in the turbine oil. The main reason for this huge difference may be the increase in tangential force. As indicated in the previous calculation (figure 9), The SIF values increase significantly with the increase of the traction coefficients. The high SIF values may result in a poor RCF life performance. It should be noted that, for the turbine oil, there is no failure at the maximum contact pressure of 5.0 GPa. The fatigue test is suspended after 120 million stress cycles as indicated in figure 12.

The calculated results indicate that if contact stress is less than 4.0 GPa, it is difficult to predict fatigue failure under rolling contact. The experimental observations show a good agreement with the model.

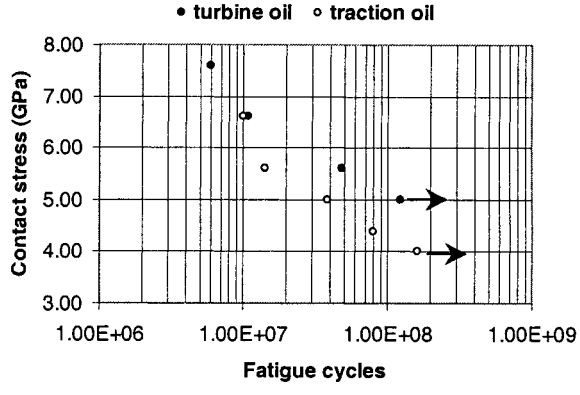


Figure 12. Plot of RCF life performance.

5. CONCLUSION

A quantitative three-dimensional boundary element model has been developed which can be used to determine the geometry of acceptable surface crack defects. Results indicate that the crack growth behaviour will be changed slightly due to the crack geometry variation. The stress intensities are the function of the normal loading and are sensitive to the tangential traction.

The failure map for the natural surface cracks was created. The results indicate that it is hard to get failure for ceramics balls when the contact stress is less than a certain value. For example, when the maximum contact stress is less than 5.0 GPa for the turbine oil, and 4.0 GPa for the traction oil, there will be no failure for the 'natural' surface ring cracks in 10×10^7 stress cycles.

The SIF values along the crack front can provide the necessary information to determine whether the ceramic ball would survive or fail. Rolling contact fatigue failure of silicon nitride bearing elements with surface cracks can be predicted using the numerical fracture mechanics analysis.

REFERENCES

1. G. Hamburg, P. Cowley, and R. Valori, Operation of an all-ceramic mainshaft roller bearing in J-402 gas turbine engine, J. ASME, Lubrication Engineering, 37 (1981) 407-415.
2. H. Aramaki, Y. Shoda, Y. Morishita, and T. Sawamoto, The performance of ball bearings with silicon-nitride ceramic balls in high speed spindles for machine tools, J. of Tribology, 110 (1988) 693-698.
3. R. T. Cundill, Rolling element bearings into the 21st century, IMechE Seminar, London, 1990.
4. G. Lundberg and A. Palmgren, Dynamic capacity of rolling bearings, Journal of Applied Mechanics, (1949) 165-172.
5. E. Ioannides and T. A. Harris, A new fatigue life model for rolling bearings, ASME Journal of Tribology, 107 (1985) 367-378.
6. T. Harris, Rolling Bearing Analysis, John Wiley and Sons, Inc, 1991.
7. M. Hadfield, T. A. Stolarski, and R. T. Cundill, Failure modes of ceramics in rolling contact, Proc R. Soc. Lond. A, 443 (1993) 607-621.
8. M. Hadfield, T. A. Stolarski, R. T. Cundill, and S. Horton, Failure modes of pre-cracked ceramic elements under rolling-contact, Wear, 169 (1993) 69-75.
9. M. Hadfield, S. Tobe, and T. A. Stolarski, Subsurface crack investigation on delaminated ceramic elements, Tribology International, 27 (1994) 359-367.
10. M. Hadfield, Failure of silicon nitride rolling elements with ring crack defects, Ceramics International, 24 (1998) 379-386.
11. Y. Wang and M. Hadfield, The influence of ring crack location on the rolling contact fatigue failure of lubricated silicon nitride: experimental studies, Wear, 243 (2000) 167-174.
12. Y. Wang and M. Hadfield, The influence of ring crack location on the rolling contact fatigue failure of lubricated silicon nitride: fracture mechanics analysis, Wear, 243 (2000) 157-166.
13. Y. Wang and M. Hadfield, Ring crack propagation in silicon nitride under rolling contact, Wear, 250 (2001) 282-292.
14. Y. Wang and M. Hadfield, Rolling contact fatigue of ceramics, in ASM Handbook, Vol. 11(2002) 957-964.
15. M. Hadfield, T. A. Stolarski, R. T. Cundill, and S. Horton, Failure modes of ceramic elements with ring-crack defects, Tribology International, 26 (1993) 157-164.
16. M. Hadfield and T. A. Stolarski, Observations of delamination fatigue on pre-cracked ceramic elements in rolling-contact, Ceramics International, 21 (1995) 125-130.
17. J. R. Fleming and N. P. Suh, Mechanics of crack propagation in delamination wear, Wear, 44 (1977) 39-56.
18. N. P. Suh, An overview of the delamination theory of wear, Wear, 44 (1977) 1-16.
19. L. M. Keer and M. D. Bryant, A pitting model for rolling-contact fatigue, Journal of Lubrication Technology-Transactions of the ASME, 105 (1983) 198-205.
20. A. F. Bower, The influence of crack face friction and trapped fluid on surface initiated rolling-contact fatigue cracks, Journal of Tribology-Transactions of the ASME, 110 (1988) 704-711.

21. A. F. Bower and N. A. Fleck, Brittle fracture under a sliding line contact, J. Mech. Phys. Solids, 42 (1994) 1375-1396.
22. A. Portela, M. H. Aliabadi, and D. P. Rooke, The dual boundary element method: effective implementation for crack problems, Int. J. Num. Meth. in Eng., 23 (1992) 1269-1287.
23. C. A. Brebbia, Boundary element Techniques, Springer-Verlag, Berlin Heideberg New York Tokyo, 1984.
24. G. R. Irwin, Analysis of stresses and strains near the end of a crack traversing a plate, Trans. ASME J. Appl. Mech., 24 (1957) 361-364.
25. M. H. Aliabadi, Numerical Fracture Mechanics, Computational Mechanics Publications, Southampton UK and D. P. Rooke, R. A. E. Hampshire UK, 1991.
26. K. L. Johnson, Contact Mechanics, Cambridge University Press, Cambridge, 1985.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Experimental investigations on Rolling Contact Fatigue of dented surfaces using artificial defects: subsurface analyses

Y. DIAB, S. COULON, F. VILLE, L. FLAMAND

Laboratoire de Mécanique des Contacts, UMR CNRS/INSA n°5514
INSA Bât. J. D'Alembert, 20 av. A. Einstein, 69621 Villeurbanne Cedex, France.
e-mail : fabrice.ville@insa-lyon.fr

It is no longer necessary to demonstrate the influence of oil contamination on Rolling Contact Fatigue (RCF). Previous studies have shown the consequences induced by the travel of solid particles inside the contacts. In fact, when particles enter the contact, they indent the surfaces. The resulting surface defects, called dents, become stress raisers and increase the risk of spalling and consequently of RCF. In order to compare numerical and experimental results, some tests were performed on a two disk machine with artificially dented surfaces. Cross-sections were performed to complete the surface observations, and to analyse the crack initiation sites and the crack propagation directions. This paper proposes an analysis of these cross-sections to identify the cracks that lead to micro and macro spalling.

1. INTRODUCTION

It is no longer necessary to demonstrate the influence of oil contamination on Rolling Contact Fatigue (RCF) [1, 2, 3, 4, 5, 6].

Previous studies have shown the consequences induced by the travel of solid particles inside the contacts [7, 8, 9, 10, 11, 12, 13, 14]. In fact, when particles enter the contact, they indent the surfaces. The resulting surface defects, called dents, become stress raisers and increase the risk of spalling and consequently of RCF.

It has been demonstrated previously, experimentally and numerically [15, 16, 17], that specific sites of dent initiated RCF exist. On the one hand, they are due to the dent geometry. In fact, as the dents induce pressure field variations, they produce an increase of the shear stresses near the surface. On the other hand, under sliding conditions, it has been shown that, even if friction is low, hydrodynamic effects induce an increase of the shear stresses behind the dent according to the sliding direction. These two phenomena explain the crack initiation and propagation. Nevertheless all these results are based on surface observations.

Other authors [18, 19, 20, 21] try to predict the fatigue life of dented surfaces. Generally they use statistical modelling, for example the Lundberg-Palmgren one [22], to predict surface fatigue. These models do not take into account the physical aspects and only calculate stress contribution to fatigue. But, if we want to build a model based on physical aspects, as those developed to predict subsurface damage around inclusions [23], it is necessary before to understand the RCF mechanisms induced.

Studying RCF of dented surfaces using natural dents (caused by the particle travel through the contact) is quite complicated. In fact, all the dents are different in term of size. So to simplify this study, we choose to perform some tests on a two disk machine with artificially dented surfaces.

In our study, cross-sections were performed to complete the surface observations, and to analyse the crack initiation sites and the crack propagation directions. This paper proposes an analysis of these cross-sections according to different slide-to-roll ratios (pure rolling, positive sliding and negative one) to identify the cracks that lead to micro and macro spalling. It is the first step to build a model of RCF due to dented surfaces.

2. EXPERIMENTAL SET-UP AND CONDITIONS

The rolling contact fatigue experiments consist in testing an artificially dented disk against a smooth one. In order to do that, the two disk machine was used [24] (Figure 1).

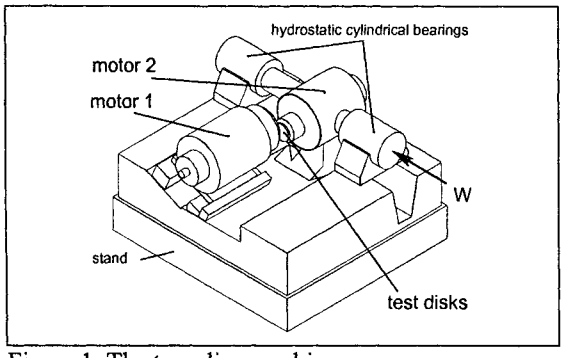


Figure 1. The two-disc machine

Prior to the tests, the disks were indented using a hardness machine with a Rockwell penetrator. The load applied on this one was 50 kg which induces a diameter of the dents of 250 µm and a depth of 38 µm. The artificial dent was chosen as its geometry is not too far from the natural dent produced by ductile particle, as shown in Figure 2 and Figure 3. As the goal of this work is to understand the fatigue crack initiation and propagation, it is important to control the dent geometry parameters and its position in the contact area.

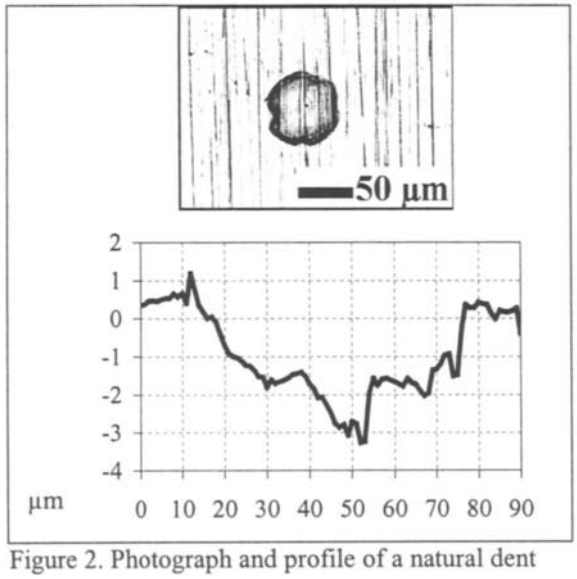


Figure 2. Photograph and profile of a natural dent

Thus, 12 artificial dents were performed on the top of the disk radius of curvature each 30 degrees. Each dent was photographed and measured using a profilometer before and after the tests (Figure 3).

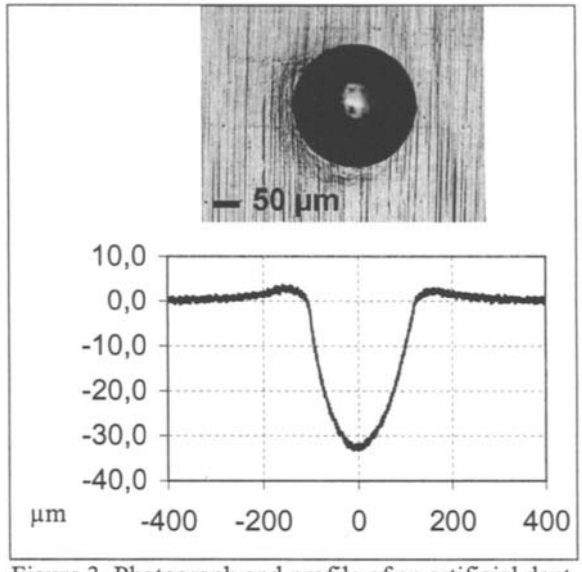


Figure 3. Photograph and profile of an artificial dent

The operating conditions (Table 1) were chosen in order to show the influence of the sliding direction on surface fatigue initiation.

Contact parameters	
Radius	40 mm
Radius of track curvature	40 mm
disc widths	10 mm
disc material	100Cr6 (AISI52100)
Rolling speed $(U_1+U_2)/2$	40 m/s
Slide to roll ratio$(U_1-U_2)/(U_1+U_2)$	0% -1.5% +1.5%
Maximal Hertz pressure P_{hmax}	3.5 GPa
Lubricant	ISO-L-MHA
Oil jet temperature	80 °C

Table 1. Operating conditions.

The crack initiation sites and the crack propagation directions are observed doing cross-sections of the disks. In the rest of the paper, the observations are presented in tables. For each cross-section, the left side corresponds to a photograph of this cross-section and the right side to a photograph of the surface (Figure 4).

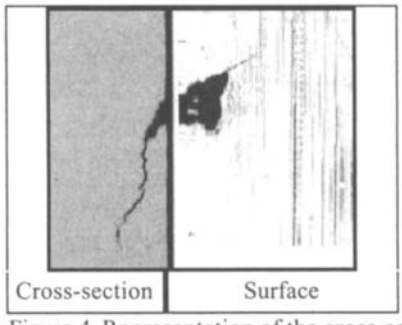


Figure 4. Representation of the cross-sections

3. SUBSURFACE DAMAGE UNDER PURE ROLLING CONDITIONS

The test conducted under pure rolling conditions was stopped after 11.1 millions of cycles. After surface observations, it appears that the smooth disk is damaged (Figure 5-a.1) while only plastic flow of the dent shoulders is observed on the dented one (Figure 5-b.1).

Concerning the dented disk, the cross-sections (Figure 5-b.2 and b.3) show that the dent shoulder has subsided into the dent. The crack initiated below this shoulder (crack noted A) is certainly due to the high shear stresses induced by the shoulder subsidence. The micro cracks observed around the dent are superficial and due to the plastic flow of the shoulders. So it could be concluded that no damage will occur on the dented surface.

Concerning the smooth disk, the photograph a.2 (Figure 5) has two interesting areas (C1 and C2). The cut of the micro spall (area C1) shows that its bottom is parallel to the surface at 11 μm depth and goes up to the surface. The crack situated at the top of the photograph (area C1) is also parallel to the surface and its growth, according to the load direction, branches off into the surface. This crack corresponds certainly to the micro spall initiation.

The photograph a.3 (Figure 5) shows a crack orientated in the direction of the load passage. Its bottom is situated at around 40 μm. This crack presents a stage parallel to the surface at 15 μm depth and branches off into the material. Some dry contact calculations were performed [19] and the stress analyses shows that between 10 and 15 μm depth, only the shear stress σxz is not equal to zero. this stress corresponds to a shear stress parallel to the surface.

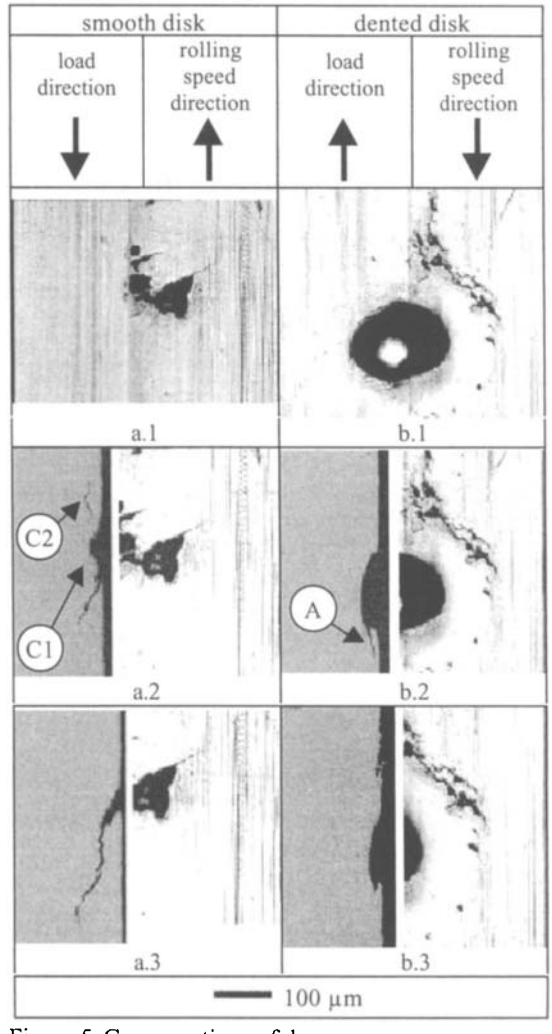


Figure 5. Cross-sections of damages under pure rolling conditions

The Figure 6 schematically shows the crack initiation and propagation for the smooth and dented surfaces.

Owing to the cross sections observations, it appears that no crack growth is possible for the dented surface. This could be explained by the fact that the crack growth direction is opposite to the load direction. In that case it was demonstrated [25] that the crack can not propagate and is going to be closed. It could be also mentioned that compressive residual stresses induced by the indentation process also protect the surface around the dent and so reduce the shear stress influence. For the smooth surface, as the pressure profile is modified by the

dent presence, some high shear stresses appear parallel to the surface. This will induce crack nucleation and as the direction of crack growth (Figure 5-a.3) is similar to the load direction, propagation is possible and will induce micro spalling and so spalling.

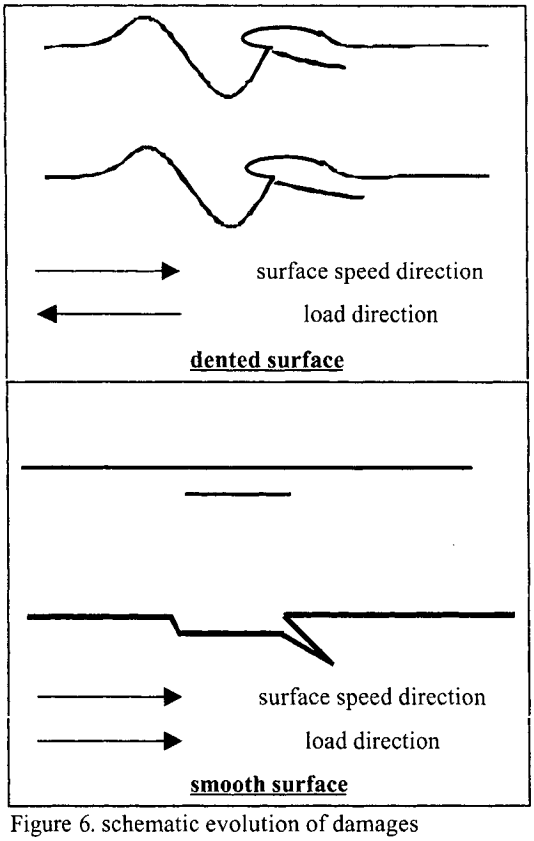


Figure 6. schematic evolution of damages under pure rolling conditions

4. SUBSURFACE DAMAGE UNDER ROLLING-SLIDING CONDITIONS

Two different experiments are conducted under rolling-sliding conditions. All the operating conditions keep unchanged and only the sliding direction was changed (+1.5 % and – 1.5 %). It is necessary to note that all the dents on each disk present damages more or less propagated.

4.1. slide-to-roll ratio +1.5 %

The test conducted under rolling-sliding conditions and with a positive slide-to-roll ratio is stopped after 2.9 millions of cycles. Damages appear both on the smooth disk and the dented one (Figure 7-a.1 and b.1).

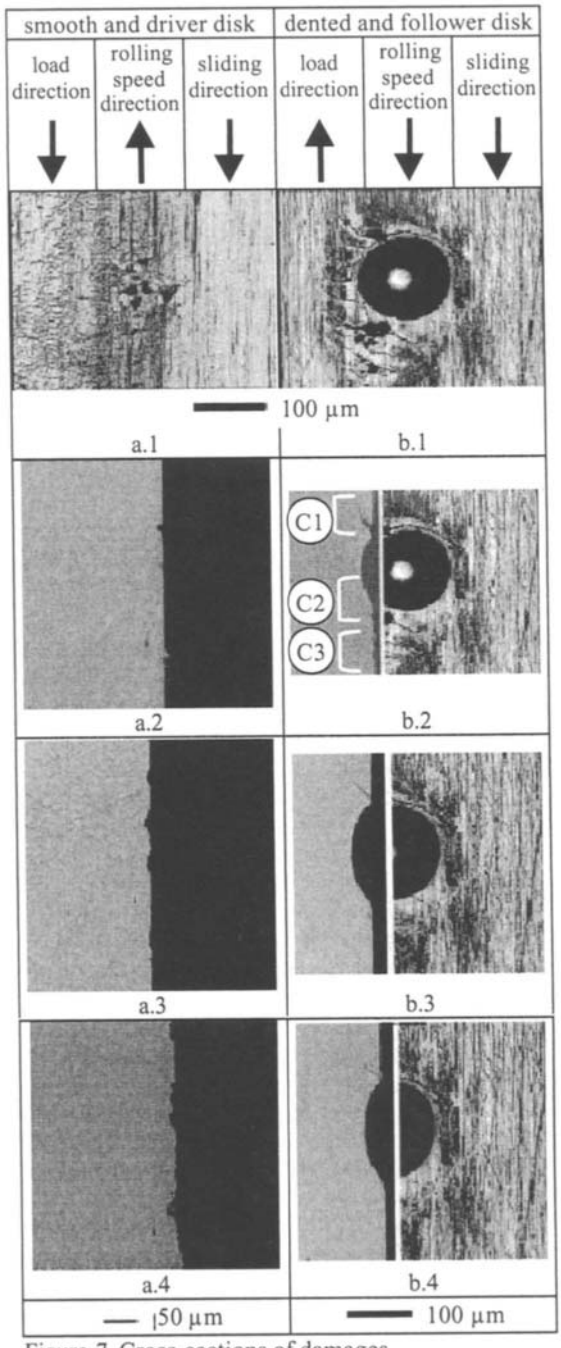


Figure 7. Cross-sections of damages under rolling-sliding conditions, g = +1.5 %

For the smooth disk, it was quite difficult to situate the cross-sections in relation with the surface photographs due to the great amount of micro spalls. The micro cracks observed on the surface are less propagated. They induce micro spalling which not exceed 3 up to 4 µm (Figure 7-a.2, a.3 and a.4). As it was mentioned before, the load direction tends to close the cracks and stop the crack growth.

For the dented surface, three different areas could be observed.

Ahead the dent according to the sliding direction ((Figure 7-b.2; C1), a crack stretches to 40 µm depth. The geometrical effects induced by the shoulders initiate this crack. The crack growth direction being similar to the load direction, this crack will certainly propagate until the hertzian zone situated to 450 µm depth leading to a deep spall. It could be noticed that this crack growths in an area free of microstructural modifications due to plastic deformations.

Behind the dent according to the sliding directions, near the shoulder ((Figure 7-b.2; C2), the cracks growth up to 4 µm depth. Both the geometrical effects and the hydrodynamic ones [16] are responsible for this crack nucleation. These cracks have the same orientation as the crack C1. Nevertheless, it seems that propagation is not possible. It could be explained by the fact that this area was highly plastically deformed after the indentation process and during the running in process. As a consequence, this area presents high compressive residual stresses which stop the crack growth.

Finally, behind the dent according to the sliding directions, far from the shoulder ((Figure 7-b.2; C3), the cracks growth up to 4 µm depth. They are due to the hydrodynamic effects induced by the micro pits of the area C2.

The Figure 8 presents schematically the initiation sites of the cracks and their growth. It is shown that the crack due the geometrical effects (left part of the scheme) could propagate because it is in the same direction as the load. The cracks on the right side, due to both geometrical and hydrodynamic effects could also propagate for the same reason. But, as they propagate in the direction of the high compressive residual stresses, they are blocked.

For the smooth surface, the cracks growth slowly as they are on the driver surface [25] because the crack growth is opposite to the load direction.

4.2. slide-to-roll ratio -1.5 %

The test conducted under rolling-sliding conditions and with a negative slide-to-roll ratio is stopped after 4.4 millions of cycles. The smooth disk presents micro cracks and micro pits (Figure 9-a.1). On the dented one (Figure 9- b.1), the damages around the dent are not really similar.

On the smooth disk, the cracks observed on the surface are very propagated in subsurface and their direction is similar to the load direction. On the photograph a.2 (Figure 9), the crack growths up to 70 µm depth and on the photograph a.3, up to 90 µm depth branching off a lot. The cut of the micro spall noted C1, shows that its bottom is situated at around 15 µm depth. A network of parallel cracks (Figure 9-a.3; C2) growth up to 10 to 15 µm depth. It could be noticed that at 10-15 µm depth, the cracks are parallel to the surface. These cracks lead to micro spall initiation.

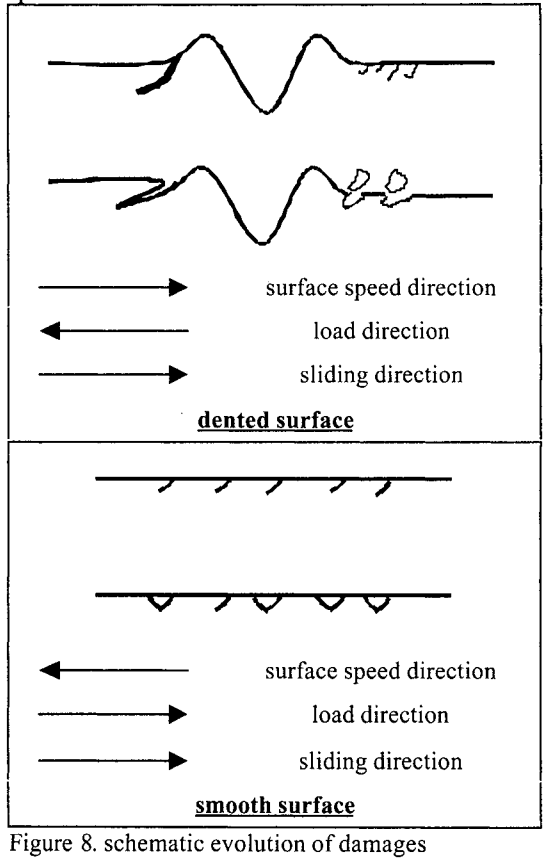


Figure 8. schematic evolution of damages under rolling-sliding conditions, g = +1.5 %

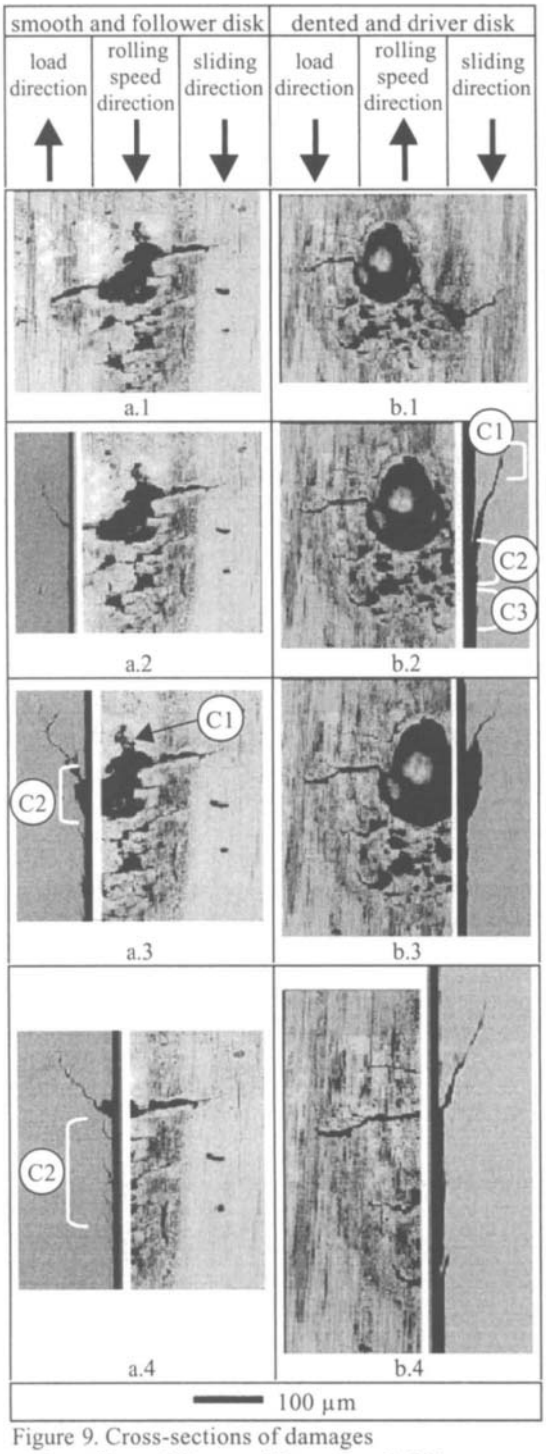


Figure 9. Cross-sections of damages under rolling-sliding conditions, g = -1.5 %

On the dented disk, although the load direction can not explain the crack propagation, one of these is very developed (Figure 9-b.2, b.3, b.4). Three different areas are observed.

In the C1 area situated ahead the dent according to the sliding direction, the crack stretches to 60 µm depth (Figure 9-b.2, b.3, b.4). This crack is due to the geometrical effects and also to the high shear stresses induced by the subsidence of the shoulder.

The C2 and C3 areas, behind the dent according to the sliding direction, present the same type of cracks as those observed for the negative slide-to-roll ratio. They growth up to 4 µm (Figure 9-b.2, b.3) and are in the opposite direction to the sliding direction. In the C2 area, the non propagation of cracks is due to the load direction and to the important microstructure modifications. In the C3 area, only the load direction explains the non propagation.

The Figure 10 shows schematically the initiation and propagation of cracks.

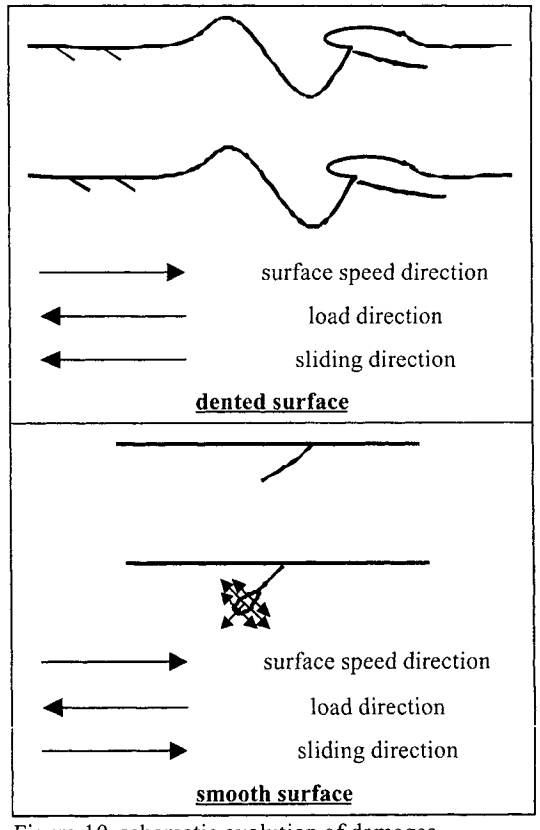


Figure 10. schematic evolution of damages under rolling-sliding conditions, g = -1.5 %

As mentioned before, the longest crack for the dented surface is certainly due to the subsidence of the shoulder and to the induced high shear stress. All the cracks on this surface can not propagate as they are on the opposite direction to the load passage.

For the smooth surface, the crack growth is explained by the load direction but also to the inflow of oil as the crack mouth is opened during the load passage. This phenomena was clearly explain by Kaneta (1).

5. CONCLUSION

The cross-sections performed to study the crack initiation and propagation on the smooth and dented surfaces for different slide-to-roll ratios show that:

1/ under pure rolling conditions, the dented surface presents few micro cracks which can not propagate as they are on the opposite direction to the load passage. The smooth surface presents a lot of cracks which seem to be initiated below the surface and parallel to it where there is a maximal shear stress. Their growth leads to micro spall initiation.

2/ under positive slide-to-roll ratio, the smooth surface presents a lot of micro cracks which induced micro spalls which not exceed 3 up to 4 μm depth. For the dented surface, geometrical and hydrodynamic effects initiate cracks. Only the crack ahead the dent according to the sliding direction could propagate until the hertzian zone because for the other ones, the microstructure modifications induced by the indentation process lead to high compressive residual stresses.

3/ under negative slide-to-roll ratio, the shoulder subsidence and the induced high shear stresses lead to the crack nucleation on the dented surface. On the smooth one, the crack growth is explained by the load direction and the oil inflow.

These cross-sections give information concerning the combined effects of load passage, sliding direction, lubricant effect and microstructure modifications due to the indentation process on the crack propagation. These observations correspond to the first step which allows a physical approach of the crack initiation and propagation under rolling-sliding conditions for dented surfaces.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge Professor Ton Lubrecht for all scientific discussions concerning fatigue life. They also wish to thank C. Godeau for his experience in performing cross-sections.

Finally, they want to pay homage to G. Tourlonias, today departed, for his assistance for the experiments and the analyses.

REFERENCES

1. Fitzsimmons, B., Clevenger, H.D., "Contaminated Lubricants and Tapered Roller Bearing Wear". Tribology Transactions, 1977, vol 20, p 97-107.
2. Ioannides, E., Beghini, B., Jacobson, B., Bergling, G., Goodall Wuttkowski, J., "Cleanliness and its Importance to Bearing Performance". Lubrication Engineering, 1993, vol 49, p 657-663.
3. Loewenthal, S.H., and Moyer, D.W., "Filtration Effects on Ball Bearing Life and Condition in a Contaminated Lubricant". Journal of Lubrication Technology, 1979, Vol.101, pp. 171-179.
4. Lorösch, H.K., "Research on Longer Life for Rolling-Element Bearings". Lubrication Engineering, 1985, vol 41, p 37-43.
5. Lösche, T., "New Aspects in the Realistic Prediction of the Fatigue Life of Rolling Bearings". Wear, 1989, vol 134, p 357-375.
6. Nixon, H. P. and Zantopulos, H., "Fatigue Life Performance Comparisons of Tapered Roller Bearings with Debris-Damaged Raceways". Lubrication Engineering, 1995, Vol. 51, n° 9, pp. 732-736.
7. Cann, P. M. E., Hamer, J. C., Sayles, R. S., Spikes, H. A. and Ioannides, E., "Direct Observation of Particle Entry and Deformation in Rolling EHD Contact". Proceedings of 22nd Leeds-Lyon Symposium on Tribology, D. Dowson, C. M. Taylor and M. Godet, Elsevier, Amsterdam, 1996, pp. 127-134.
8. Cheng, W., Cheng, H.S., Keer, L.M., "Experimental Investigation on Rolling/Sliding Contact Fatigue Crack Initiation with Artificial Defects". Tribology Transactions, 1994, vol 37, p 1-12.
9. Coulon S., Ville F., Lubrecht A.A., "Effect of a dent on the Pressure Distribution in Dry Point

Contacts", ASME Journal of Tribology, to be published.
10. Dwyer-Joyce, R. S. and Heymer, J., "The Entrainment of Solid Particles into Rolling Elastohydrodynamics Contacts". Proceedings of 22nd Leeds-Lyon Symposium on Tribology, D. Dowson, C. M. Taylor and M. Godet, Elsevier, Amsterdam, 1996, pp. 135-140.
11. Girodin, D., Ville, F., Guers, R., Dudragne, G., "Experimental investigation of surface initiated rolling contact fatigue using dented surfaces". Proceeding of the 27th Leeds Lyon Symposium on Tribology, Lyon Sept. 2000, G. Dalmaz et al. Editors, Amsterdam : Elsevier, 2001, p 449-458.
12. Nikas, G.K., Sayles, I., Ioannides, E., "Effect of Debris Particles in Sliding/Rolling Elastohydrodynamic Contacts". Proc. Instn. Mech. Engrs., 1998, vol 212, p 333-343.
13. Ville, F., Nelias, D., "An experimental Study on the Concentration and Shape Caused by Spherical Metallic Particles in EHL Contacts". Tribology Transactions, 1999, vol 42, p 231-240.
14. Wedeven, L.D., "Influence of Debris Dent on EHD Lubrication". Tribology Transactions, 1977, vol 21, p 41-52.
15. Coulon, S., Ville, F., Nelias, D., "Numerical and Experimental Investigations on Rolling Contact Fatigue for Dented Surfaces". Proceeding of the 27th Leeds Lyon Symposium on Tribology, Lyon Sept. 2000, G. Dalmaz & al. Editors, Amsterdam : Elsevier, 2001, p 459-467.
16. Ville, F., Nélias, D. "Early Fatigue Failure Due to Dents in EHL Contacts". ASLE Transactions, Vol.42, 1999, pp. 795-800.
17. Xu, G., Sadeghi F., "Spall Initiation and Propagation due to Debris Denting". Wear, 1996, vol 201, p 106-116.
18. Ai, X., and Nixon, H. P., "Fatigue Life Reduction of Roller Bearings due to Debris Denting : Part II – Experimental Validation," Tribology Transactions, 2000, Vol. 43, pp. 311-317.
19. Coulon S., Ville F., Lubrecht A.A., "An abacus for predicting the rolling contact fatigue life reduction due to debris dents ", Proceedings of the 28th Leeds-Lyon Symposium on Tribology. D. Dowson, C.M. Taylor and M. Godet, Elsevier, Amsterdam, in Tribology Series, 11 pages, to be published.
20. Lubrecht, A. A., Dwyer-Joyce, R. S. and Ioannides, E., "Analysis of the Influence of Indentations on Contact Life," Proceedings of 19th Leeds-Lyon Symposium on Tribology, D. Dowson, C. M. Taylor and M. Godet, Elsevier, Amsterdam, 1993, pp. 173-181.
21 Sayles, R.S., Ioannides, E., "Debris Damage in Rolling Bearings and its Effects on Fatigue Life", Journal of Tribology, 1988, Vol.110, pp.26-31.
22. Lundberg, G., Palmgren, A., "Dynamic capacity of Roller Bearings". Acta Polytechnica, Mechanical Engineering Series, Royal Swedish Academy of Engineering Sciences, 1952, vol 2, n°4, 96 p.
23. Lamagnere, P., Fougeres, R., Lormand, G., Vincent, A., Girodin, D., Dudragne, G., Vergne, F., "A Physically Based Model for Endurance Limit of Bearing Steels". Journal of Tribology, 1998, vol 120, p 421-426.
24. Ville F., Nélias D., Tourlonias G., Flamand L. & Sainsot P., "On The Two Disc Machine : A Polyvalent and Powerful Tool to Study Fundamental and Industrial Problems Related to Elastohydrodynamic Lubrication ". Tribology series "Tribology Research: From model experiment to industrial problem: A century of efforts in mechanics, materials science and physico-chemistry", D. Dowson et al. Editeurs. Elsevier, Amsterdam, 2001, Vol. 39, pp.393-402
25. Kaneta, M., Yatsuzuka, H., Murakami, Y., "Mechanism of crack growth in Lubricated Rolling/Sliding Contact". Journal of Tribology, 1985, vol 28, p 407-414.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Incorporating a Railway Wheel Wear Model into Multi-Body Simulations of Wheelset Dynamics

A. Ward, R. Lewis, R.S. Dwyer-Joyce

Department of Mechanical Engineering, University of Sheffield, UK

A model has been developed to predict the wear of railway wheels. The wear modelling approach is based on a wear index commonly used in rail wear predictions. This assumes wear is proportional to $T\gamma$, where T is tractive force and γ is slip at the wheel/rail interface. Twin disc testing of rail and wheel materials was carried out to generate wear coefficients for use in the model.

The modelling code is interfaced with ADAMS/Rail, which produces multi-body dynamics simulations of a railway wheelset and contact conditions at the wheel/rail interface. Simplified theory of rolling contact is used to discretise the contact patches produced by ADAMS/Rail and calculate local traction and slip.

The wear model combines the simplified theory of rolling contact, ADAMS/Rail output and the wear coefficients to predict the wear and hence the change of wheel profile as the vehicle passes over given track layouts.

1. INTRODUCTION

New specifications are being imposed on railway wheel wear and reliability to increase the time between wheel reprofiling operations, improve safety and reduce total wheelset lifecycle costs. In parallel with these requirements, changes in railway vehicle missions are also occurring. These have led to the need to operate rolling stock on track with low as well as high radius curves, increase speeds and axle loads and contend with a decrease in track quality due to a reduction in maintenance. These changes are leading to an increase in the severity of the wheel/rail contact conditions [1] which will increase the likelihood of wear occurring.

In order to deal with these demands new design methodologies are required that integrate advanced numerical tools for modelling of railway vehicle dynamics and suitable models to predict damage sustained by wheels under typical operating conditions.

The aim of the work described was to develop a wear model for integration with multi-body dynamics simulations of a railway wheelset to develop an engineering design tool for predicting wheel profile evolution. This will help in designing wheels for minimum wear, optimising railway vehicle suspensions and wheel profiles and in evaluating new wheel materials.

This work was carried out as part of the European Community funded project HIPERWheel. One aspect of this project was the development of an integrated CAE procedure for assessing wheelset durability taking into account damage mechanisms such as metal fatigue, rolling contact fatigue, wear and fretting.

2. MODELLING METHODOLOGY

The modelling methodology for predicting the evolution of a railway wheel profile consists of four main elements:

(i) *Multi-Body Simulation of the Wheelset.* Used to obtain, as a function of time, position, force and slip for the wheel/rail contact as the wheelset passes around a pre-defined track layout.

(ii) *Local Contact Analysis.* Used to derive the pressure, traction and slip distributions within the wheel/rail contact. The inputs for these calculations are taken directly from ADAMS/Rail multi-body simulations.

(iii) *Wear Determination.* Based on wear constants derived from twin disc wear testing and used to relate the local contact conditions to the amount of material removed from the wheel.

(iv) *Wheel Profile Evolution Prediction.* The wheel is discretised and wear caused by each set of contact conditions generated in ADAMS/Rail is summed to determine the change in profile. The worn profile is then fed back to ADAMS/Rail for a new set of input conditions to update the contact analysis.

The complete procedure is schematically represented in Figure 1. This shows how the modelling scheme is integrated with the multi-body dynamics simulations generated by ADAMS/Rail (see Figure 2) to produce a complete CAE tool for predicting wheel profile evolution.

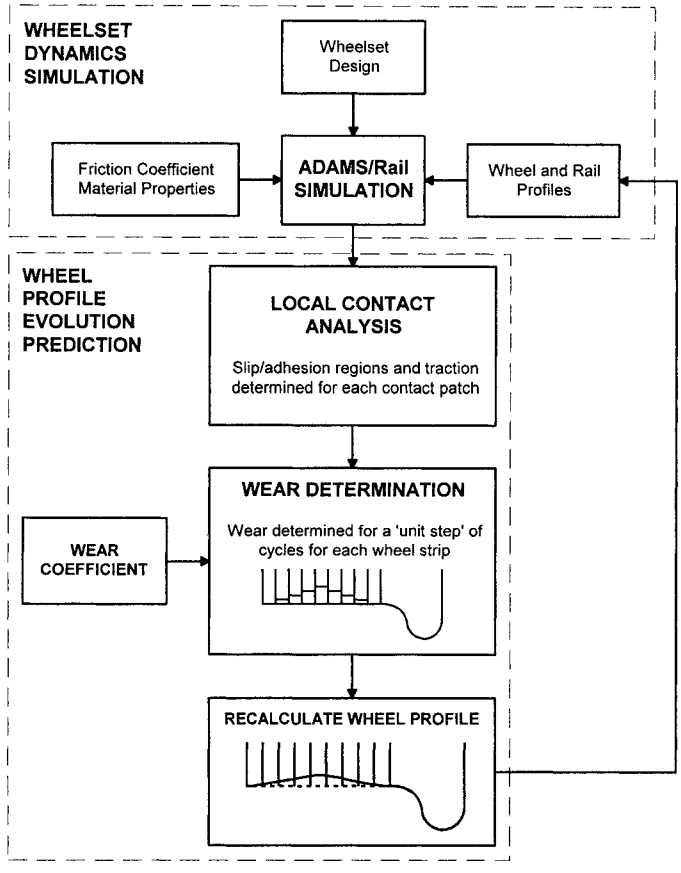


Figure 1. Railway Wheel Wear Modelling Scheme

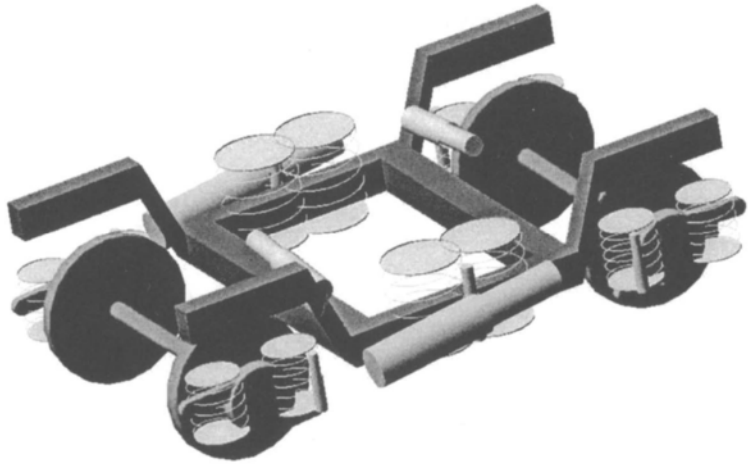


Figure 2. Multi-Body Wheelset Simulation Produced using ADAMS/Rail

3. WEAR TESTING

The approach used in the model for predicting wear is based on a rail wear index [2]. An energy approach is adopted in the analysis of the relationship between wear rate and contact conditions. This assumes that wear rate (μg/m rolled/mm^2 contact area) is related to work done at the wheel/rail contact (wear rate = $KT\gamma/A$, where T is tractive force and γ is slip at the wheel/rail interface, K is a wear coefficient and A is the contact area).

In order to generate wear coefficients required for the model, twin disc testing was carried out. This allowed the close control of slip and load required to study the performance of the wheel material over a range of contact conditions. The original development of the twin disc machine and more recent work carried out to add computer control have been described previously [3, 4].

The test discs are hydraulically loaded together and driven at controlled rotational speed by independent electric motors. Shaft encoders monitor the speeds continuously. A torque transducer is mounted on one of the drive shafts and a load cell is mounted beneath the hydraulic jack. The slip ratio required is achieved by adjustment of the rotational speeds. All data is acquired on a computer which is also used for load and speed control.

Discs used in the testing were machined from R8T wheel rims and UIC60 900A rail sections. Wear tests were carried out using the wheel disc as the driving disc and the rail disc as the braking disc.

A nominal disc rotational speed of 400rpm was used in the tests. An environment chamber enclosed the discs and air cooling was provided to both to prevent a bulk temperature rise and subsequent oxide formation. Wear measurement was determined by mass loss of the discs, measured before and after tests and at intervals during initial tests to determine the number of cycles required to reach steady state wear. Contact stresses and slip were changed in order to vary the $T\gamma$ parameter for

the wear index analysis (for further details on the wear tests see [5]).

The results (see Figure 3) indicate that a number of wear regimes exist. A wear constant has been defined for each regime as shown.

At low values of $T\gamma$ (typical of tread contact conditions) (K_1) the wheel disc surface turned a rusty brown colour indicating oxidative wear was occurring. Closer examination of the wear surface revealed abrasive score marks and evidence of the oxide layer breaking away. As $T\gamma$ was increased (to values typical of flange contacts) (K_2 - K_3), the wheel material appeared to wear by a delamination process. Observation of the subsurface morphology revealed that a larger amount of plastic deformation was occurring below the disc wear surface and crack formation just below the surface was leading to thin slivers of material breaking away.

4. DISCRETISATION OF THE WHEEL PROFILE

To determine the evolution of the wheel profile as it wears from the output of ADAMS/Rail, the wheel profile was discretised into circumferential strips. This enabled the elliptical contact patches to be located on the wheel circumference. The discretisation of the contact patches could then be made to coincide exactly with the strips of the wheel (see Figure 4). By aligning the discretisation strips in this way, wear calculated for each strip could be summed to calculate the wear depth for that region of the wheel profile. This assumes that the profile wears uniformly over the wheel circumference (and so it will not consider out of roundness or wheel flat aspects of profile change).

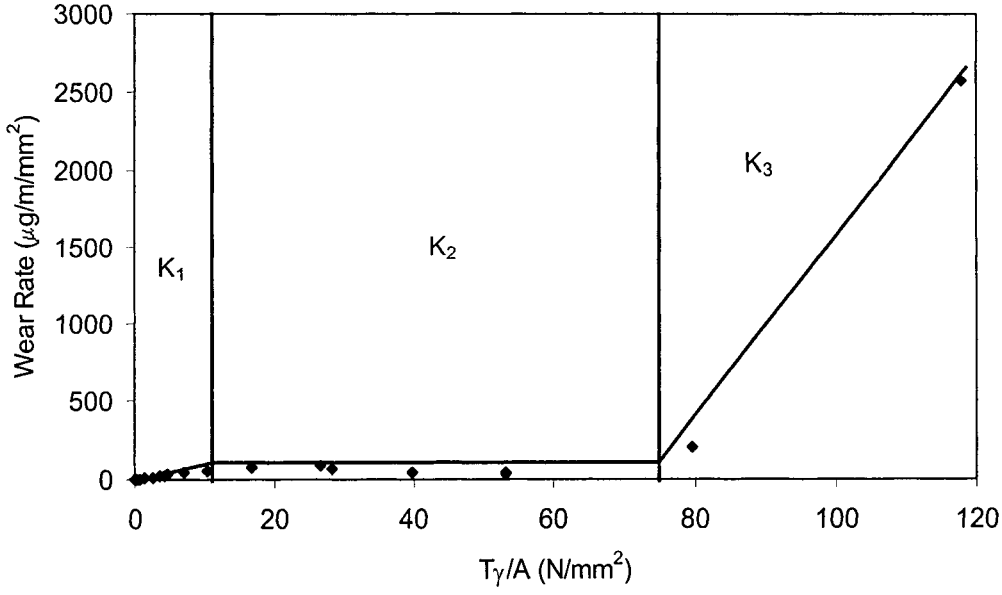


Figure 3. R8T Wheel Material Wear Rates and Regimes

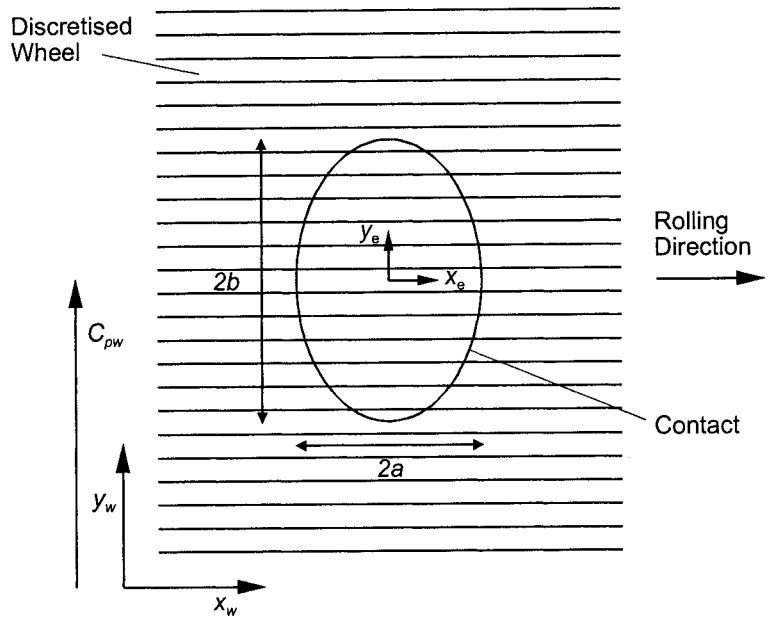


Figure 4. Discretised Wheel and Elliptical Contact Co-ordinate Systems

The positioning of the wheel strips was achieved by setting a co-ordinate system for the wheel and deciding on an interval width for discretisation. The strip width can be altered to improve resolution or increased to reduce computing time. For subsequent pressure, traction and slip calculations the contact is normalised giving the relationship between the wheel and ellipse co-ordinate systems as:

$$y_e = \frac{y_w - C_{pw}}{b} \tag{1}$$

where y_e is the y value in the ellipse co-ordinate system, y_w is the y value in the wheel co-ordinate system, C_{pw} is the contact point on the wheel and b is the ellipse semi-axis.

5. LOCAL CONTACT ANALYSIS

ADAMS/Rail provides information about the load and tractions applied to the whole contact patch and its location. In this model it is assumed that the contact between the wheel and rail can be approximated by an elliptical contact patch. This allows the use of a faster solution method to determine the traction and slip distribution within the contact.

To enable the calculation of wear for each discretised strip of the wheel profile the contact ellipse is discretised into strips that correspond to the wheel strips. Each strip within the ellipse is then divided into equal sized cells.

ADAMS/Rail provides the total normal force, traction applied and the semi-axes a and b for each elliptical contact patch and it is from these values that local contact

analysis is performed. The contact analysis calculates traction, T, pressure, p, and slip, γ, for each cell (see Figure 5) from the leading edge of the contact, sequentially to the rear. The forces determined for each cell are compared to slip/adhesion criteria enabling slip and adhesion regions to be identified and the slip magnitude and direction to be determined. This approach is similar to the method employed in the FASTSIM algorithm [6].

Some examples of output from the local contact analysis are shown in Figure 6. These are in good agreement with those generated using FASTSIM using similar conditions [6, 7].

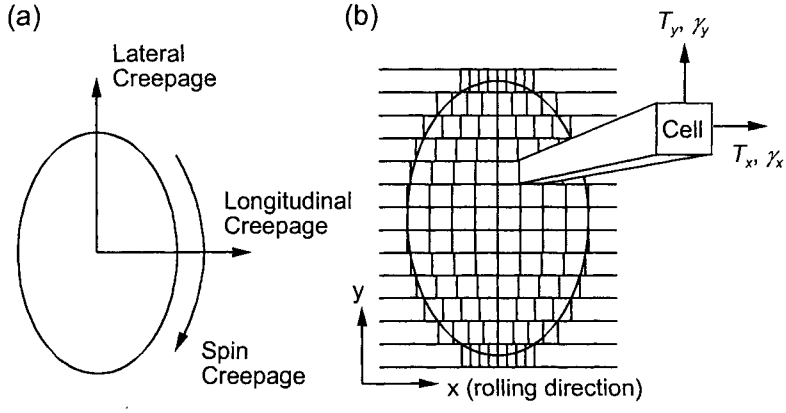


Figure 5. Output from: (a) ADAMS/Rail and (b) Local Contact Analysis

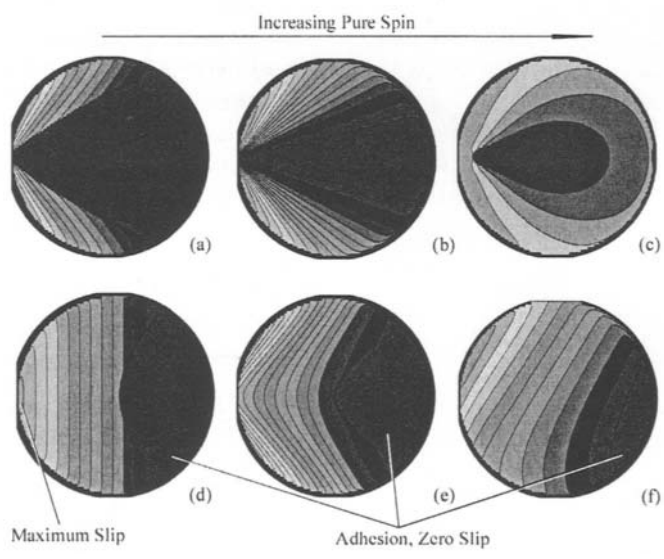


Figure 6. Local Contact Analysis for (a)-(c) Increasing Pure Spin; (d) Pure Creepage; (e) Lateral Creepage with Spin and (f) General Case

6. WEAR PREDICTION

A value for $T\gamma/A$ was calculated for each cell in the local contact analysis and compared to the wear regimes identified from wear test data (as shown in Figure 3 and Table 1). The corresponding wear coefficient, K, for each wear regime was used to calculate the wear volume per cell, WD_{cell}. The depth of wear per cell is then given by:

$$WD_{cell} = \frac{T\gamma}{A} K \frac{Vt}{\rho} \qquad (2)$$

where V is the rolling velocity, ρ is the density of the wheel steel and t is the duration of the contact.

Each cell in the contact ellipse has an associated wear depth. The wear on each strip (in the rolling direction) is summed and that depth removed from the corresponding circumferential strip on the wheel. The contact time interval, t, in the output data from ADAMS/Rail is seldom equal to one revolution of the wheel. It is generally observed that wear occurs uniformly over a wheel circumference. Therefore it is reasonable to proportion the wear caused by a given patch location over the time, t, equally over the whole circumference of the wheel according to the ratio $Vt/2\pi R$, where R is the wheel radius. This is carried out for each strip in the contact patch to determine the wear caused by that contact over the time, t. Each contact patch obtained from the ADAMS/Rail simulation is treated in the same way to obtain the worn wheel profile. This procedure is shown schematically in Figure 7.

7. WORN WHEEL PROFILES

A sample of ADAMS/Rail output data was run through the wear prediction programme. The data consisted of a vehicle running on a sample length of track as detailed below:

- 950 metres of straight track
- 240 metres of transition (straight to curve)
- 1440 metres of constant radius right hand curve
- 260 metres of transition (curve to straight)
- 1110 metres of straight track

The velocity of the vehicle was 220km/h, the radius of the curve was 7000m and the cant was 0.06m. The vehicle takes 65 seconds to cover the length of track.

The upper graph in Figure 8 shows the location of a contact patch on the right wheel of the leading bogie as the vehicle negotiates the curve. The lower graphs show the evolution of the wheel profile determined over 5 second intervals. The wear from the profile is shown both as that occurring just during the 5 second period and as a cumulative total.

Table 1.
Wear Regimes and Coefficients

Regime	$T\gamma/A$ (N/mm^2)	Wear Rate (μg/m/mm^2)
K_1	$T\gamma/A < 10.4$	$5.3T\gamma/A$
K_2	$10.4 < T\gamma/A < 77.2$	55
K_3	$77.2 < T\gamma/A$	$61.9T\gamma/A$

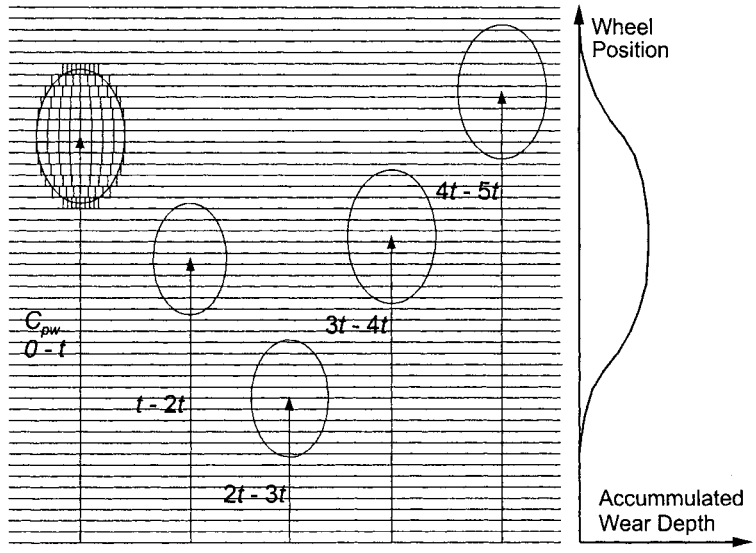


Figure 7. Summation of the Wear per Strip to give Total Wear Depth

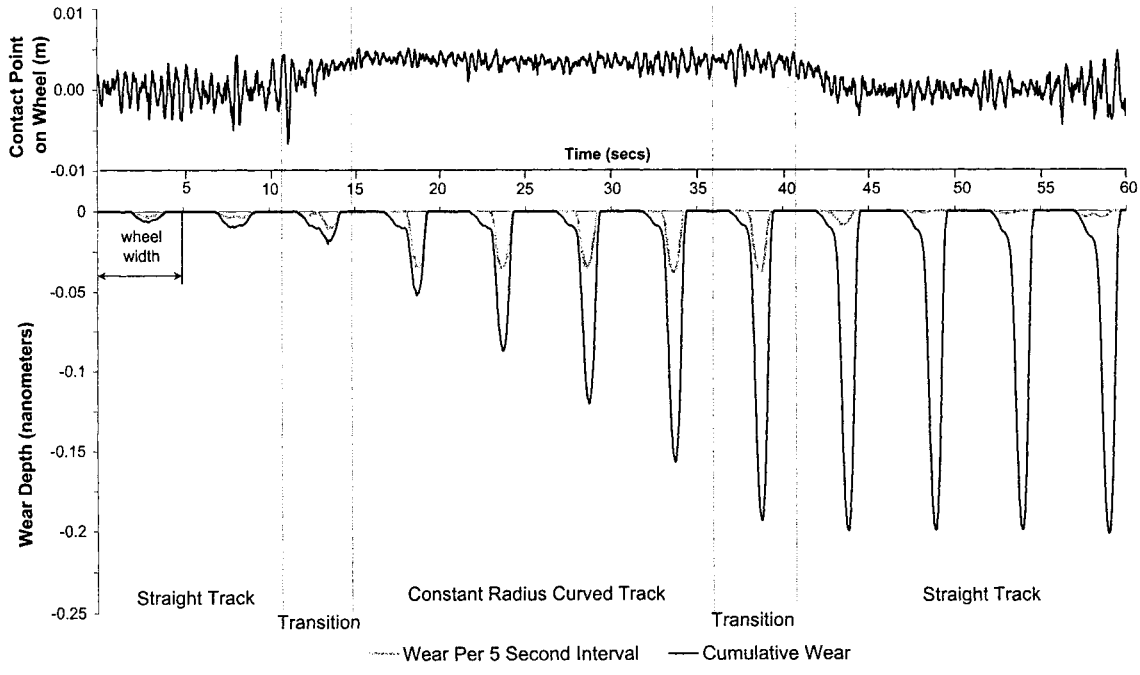


Figure 8. Wheel Wear per 5 Second Interval against Rail Position and Contact Point on the Wheel

The majority of wear on the tread of a wheel occurs around a curve. The straight track conditions produce little wear, although from the contact point on the wheel data it can be seen that the wheel moves most relative to the track in this region.

Figure 9a shows the depth of wear after 33000 wheel revolutions where the vehicle has been repeatedly running over the above sample of track. Figure 9b shows the location of this material removed with respect to the wheel profile. The wear from this short section of track is confined to a small section of the tread far from the flange (this is because the data used is for a right wheel on a right hand curve where the left wheel will be in flange contact). Using more severe curvature of the track, the contact area will significantly increase profile wear.

It has been observed that for a wheel rolling along a straight track for 33000 cycles, approximately 1 micron of wear occurred [8]. Whilst this data is for a different vehicle and track configuration, the amount of wear is of a similar order. Especially considering that for the sample result in the present case the contact patch position does not vary over a large region. The result of this is to concentrate the wear to a small region of the wheel.

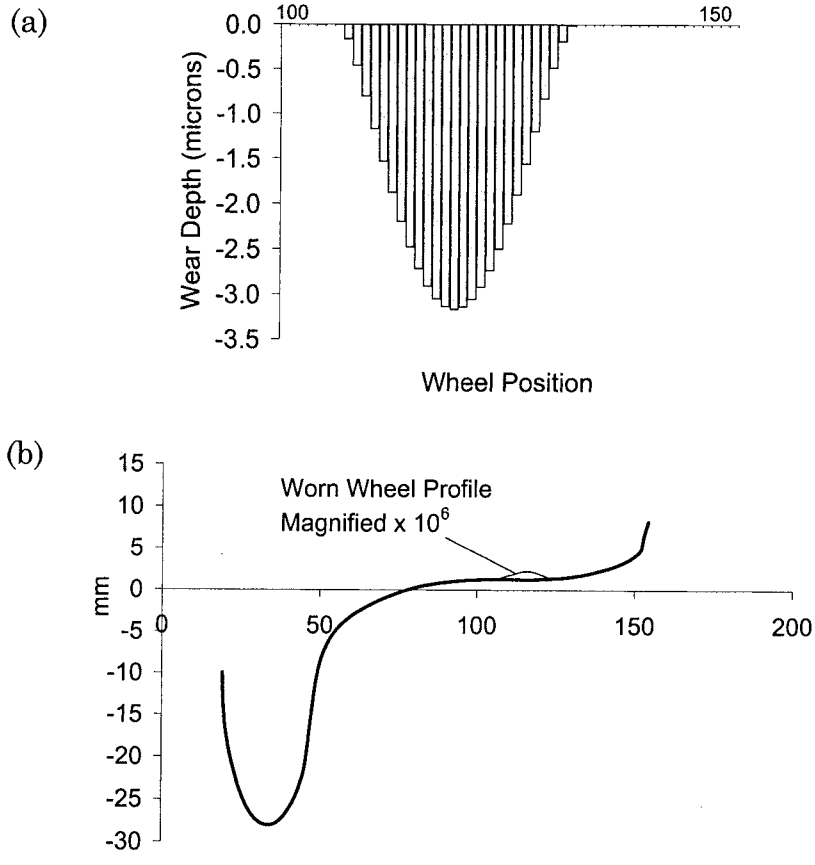


Figure 9. Wear Depth after 33000 Wheel Revolutions (a) per Strip of Wheel and (b) with Respect to the Wheel Profile

8. CONCLUSIONS

A model has been developed to predict the wear of railway wheels. This assumes wear is related to $T\gamma$, where T is tractive force and γ is slip at the wheel/rail interface. Twin disc testing of rail and wheel materials was carried out to generate wear coefficients for use in the model.

The model has been incorporated into ADAMS/Rail, which produces multi-body dynamics simulation of a railway wheelset. The methodology for predicting the evolution of a railway wheel profile is made up of three stages: ADAMS/Rail outputs contact conditions at the wheel/rail interface; simplified theory of rolling contact is then used to discretise the contact patches and calculate traction and slip within each; finally the wear coefficients are used with the tractions and slips to predict the wear and hence the change of wheel profile.

This simulation routine will be used as part of a CAE package for assessing wheelset durability for the design of new wheelsets.

The model has been applied to a multi-body simulation of a passage around a sample track length. The worn profile is quantitatively similar to published data.

9. ACKNOWLEDGEMENTS

The work reported in this paper was carried out as part of the European Community funded project HIPERWheel (contract number: G3RD-CT2000-00244). The authors would like to acknowledge the assistance of M. Cavaletti of MDI.

REFERENCES

1. Stanca, M., Stefanini, A., Gallo, R., 2001, "Development of an Integrated Design Methodology for a New Generation of High Performance Rail Wheelsets", *Proceedings of the 16th European MDI User Conference*, Berchtesgaden, Germany, 14-15 November.
2. Elkins, J.A., Eickhoff, B.M., 1979, "Advances in Non-Linear Wheel/Rail Force Prediction Methods and Their Validation", ASME Winter Annual Meeting.
3. Garnham, J.E., Beynon, J.H., 1991, "The Early Detection of Rolling-Sliding Contact Fatigue Cracks", *Wear*, 1991, Vol. 144, pp103-116.
4. Fletcher, D.I., Beynon, J.H., 2000, "Development of a Machine for Closely Controlled Rolling Contact Fatigue and Wear Testing" *Journal of Testing and Evaluation*, Vol. 28, No. 4, pp267-275.
5. Lewis, R., Dwyer-Joyce, R.S., Kapoor, A., "Wear Mechanisms in Railway Wheel Steels", in press.
6. Kalker, J.J., 1982, "Fast Algorithm for the Simplified Theory of Rolling Contact", *Vehicle System Dynamics*, Vol. 11, pp1-13.
7. Kalker, J.J., 1979, "Survey of Wheel-Rail Rolling Contact Theory", *Vehicle System Dynamics*, Vol. 5, pp317-358.
8. Kalker, J.J., "1991, "Simulation of the Development of a Railway Wheel Profile through Wear", *Wear*, Vol. 150, pp355-365.

SESSION X
COMPONENT TRIBOLOGY (1) - PLAIN BEARINGS

Chairman : Dr P M Cann

Paper X (i) **Tilting Pad Thrust Bearings**

Paper X (ii) **Study on the Development of a Thin-film Sensor for Measuring Oil-film Pressure in Engine Bearing and Piston Pin-boss**

Paper X (iii) **Optimisation of Static Behaviour and Stability Level of a Multilobe Journal Gas Bearing**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tilting pad thrust bearings

S.B.Glavatskih

Sirius Laboratory, Luleå University of Technology,
SE-971 87 Luleå, Sweden

This paper is an extensive review of the published contributions on experimental and design research conducted on tilting pad thrust bearings. Past and recent bearing designs are reviewed with a special reference to large and PTFE-faced bearings. It also includes a brief discussion of bearing lubrication and monitoring techniques. Likely future trends in bearing development are highlighted as well.

1 INTRODUCTION

The remarkable story of tilting pad bearing invention and industrial introduction is thoroughly documented [1]. As highlighted in [1] tilting pad bearings "were the direct results of engineering applications of Osborne Reynolds' theory of fluid film lubrication" and "problems with hydraulic machinery promoted the inventions".

The first tilting pad thrust bearing designed by A.G.M.Michell was brought into service 95 years ago. The bearing was installed in a centrifugal pump on the Murray River in Australia and carried a load of 13.3 kN at 200 rpm [2]. Over the years bearing loads and operating speeds have increased tremendously. Modern thrust bearings are required to carry loads of up to 60 MN [3]. This certainly requires a better understanding of bearing behaviour to provide adequate design modifications.

The development of numerical methods and computational power has allowed an accurate three-dimensional analysis of bearing operation. Such factors as thermal and mechanical deformations, turbulence, thermal effects and others can now be accounted for in bearing models. A comprehensive review on this subject is given in [4,5].

Design studies and the translation of theory into practical applications are just as important. Modern design tools backed by computers are very efficient, yet heavily loaded bearings such as those used in hydroelectric power stations still present a challenge to bearing designers. There is then a question of how the design studies have affected bearing operation and reliability. And what are the major technological advances and innovations in bearing design?

This paper is an attempt to answer these questions through an extensive survey of published contributions on experimental and design research conducted on tilting pad thrust bearings. The paper also includes a discussion of existing bearing designs and likely future trends in their development.

2 PAD AND PIVOT DISIGNS

For optimum bearing operation the pad should crown slightly. As shown in the fundamental work by [6] a slight crowning is beneficial for offset pads and essential for centrally pivoted pads. The crowning is achieved by a combination of initial machining, deflection due to load and thermal effects. According to [6] the crowning ratio introduced to describe the relative degree of pad surface deformation must be maintained within reasonable limits to obtain maximum load carrying capacity. Control of thermal distortion becomes more difficult with an increase in pad dimensions. Thus, pad and pivot design depends on bearing size.

2.1 Small and Medium Size pads

As in the early days there are two common support configurations for small and medium size thrust bearings. The first configuration, a line contact support, is a radial rib on the back face of the thrust pad, generally with a cylindrical surface.

The other one, a spherical support, is formed by a spherically faced surface of the support disc, mounted in the back face of the thrust pad. Spherical support configuration is often preferred due to the ability of the pad to tilt not only in the direction of rotation but also in the radial direction. The latter provides some compensation for misalignment between the thrust bearing face and the collar. On the other hand, the radial rib support reduces radial elastic pad distortion and does not require hardening. Spherical pivots must be hardened because of high contact stresses at the pad pivot [7].

Other pad pivot designs have also been reported but are not yet gained wide acceptance or in the development stage such as a rubber-in-shear mount in swing-pad bearings [8]. The objective was to create a tilting pad bearing in a simpler and less costly configuration than a conventional one. A rubber-in-shear mount employed in the swing-pad bearing as a means to support the pads was composed of multiple, spherical laminations of rubber and metal. Desired tilt was obtained when the rubber underwent shear displacements in response to surface traction imposed on pads. Test results indicated that the frictional behaviour of the bearing with centrally pivoted pads was poor but pads with 65% offset showed more promising performance. It was concluded that additional research was required to define the optimum geometry of the swing pad bearing to be comparable with a tilting pad bearing.

Another more successful design is described in [9,10]. In flexible pad bearings a flexible pad and the support structure were made integral. This eliminated variation in pad heights that is present in conventional multicomponent tilting pad bearings. Pad movement was provided by deflection of the pad on its support structure. Due to the flexible support the pad could support a moment. This capability, depending on the load vector position relative the pivot, helped to deflect the structure to form a converging edge or resisted the formation of the diverging edge. Another advantage of one-piece design was claimed to be the elimination of pivot wear and fretting due to the absence of relative part movement. In the tests the circular pad geometry showed lower temperatures and power loss compared to the sector shaped pads. It was mentioned that the pivot was offset though no exact values were given.

Centrally pivoted pads are used in applications where the bearing must accept rotation in either direction. Such pads are also practical for other applications as the risk of incorrect assembly for whatever reason is eliminated. In high-speed bearings where the main limiting factor is that of temperature it is of advantage to use offset pads [7, 11,12]. Moving the pivot towards the trailing edge forces the leading edge film thickness to be greater, thereby bringing more cold oil into the film, resulting in lower operating temperature, thicker oil film and higher load carrying capacity.

Various isoviscous analyses, see e.g. [13], indicated that a flat pad had maximum load carrying capacity for a pivot location of 60 per cent. This can probably explain why 60 per cent offset is common today. However, experiments and modern THD analyses show that the metal temperature minimum occurs at pivot offsets between 70 and 80 per cent [14-16]. The range accounts for different designs. From the standpoint of minimum film thickness the best pivot location is between 70 and 75 per cent. Maximum oil film pressure must also be taken into account, as it is a combination of temperature and pressure that affects strength of pad surface material. Thus, optimum pivot position is at about 65 per cent. Application of the novel surface materials that can withstand higher pressures will allow 70 per cent pivot to be used with a subsequent significant reduction in oil film temperature. Lower film temperatures are also beneficial from the lubrication point of view. Degradation process in conventional mineral oils goes faster at elevated temperatures.

An important question of whether offset pads can operate with the collar rotating in reverse was answered by [17,18]. A substantial reverse running load carrying capacity of offset pads was observed. The difference in temperature between backwards running offset pads and centre-pivoted pads was small compared to the difference between the normal operation of offset- and centre-pivoted pads [18].

2.2 Large size pads

Large bearings mostly above 1 meter diameter constitute the most important tribological element of hydroelectric power stations. This specific field of application presents a challenge to bearing designers due to the unique operating conditions: very high loads in combination with low speeds. The situation is aggravated by the existence of a harmful size effect [19] requiring special attention to the control of pad thermal deformations. A complex, often

custom, design is thus required to avoid excessive pad distortion. This resulted in a variety of pad and support structure designs.

Various methods were proposed to control thermal distortion. Basically, all the methods can be divided into three groups. The first group deals with the hot oil carryover. Pad temperature and consequently thermal distortion are reduced if greater amount of cold oil is introduced into the film. This is achieved by spray jets [20-22] or scrapers [23]. Both arrangements are intended to scour off the hot oil carryover that adheres to the rotating collar. Scrapers are more efficient. As shown in [23] pad temperatures can be decreased by 20°C and more but it enacts a penalty in a form of additional power loss.

Methods related to pad design comprise the second group. During normal operation pad surface becomes convex due to the differential thermal expansion. This thermal deformation can be cancelled by mechanical deformation if pad or support structure are designed in such a way that oil film pressure makes the pad concave [20,23,24].

Another way of reducing thermal deformation is to use the double layer design, in which a thin babbitted plate is mounted on a thick pad that in turn rests on a pivot [22,23,25,26]. A system of channels is provided between the plate and the pad to induce an oil flow by the collar motion. This simultaneously increases thermal gradient in the plate and reduces thermal distortion of the lower thick pad. At the same time the flexible thin plate is pressed flat against the thick pad by the action of the oil film pressure. The deformation and, in addition, the film temperature are consequently reduced. A further reduction of thermal distortion and oil film temperatures is provided by the direct water cooling as reported in [22,27,28].

Lower pad convex deformation can be achieved by using pad surface and pad backing materials of high thermal conductivity. Insulation of the pad face is also possible. These methods are considered in Section 4. Insulation of pad bottom and its sides by an oil resistant rubber to minimise thermal gradient was reported in [22].

Other solutions include double row construction and arrangements with rotating pads [23]. In order to accommodate high thrust loads (higher than 10 MN) while keeping unit loading under 4 MPa bearing dimensions need to be greatly increased, which in turn give rise to thermal deformation. A solution is to arrange pads in two rows. Each pair of pads, one from the outer row and one from the inner row, is supported by an equaliser, a lever of the first order. The lever rests on a cylindrical pivot. Location of the equaliser fulcrum affects load distribution between the pads. Double row bearings carry thrust loads of up to 34 MN [23].

An unusual design of a thrust unit was reported in [23]. It was designed in 1920's for thrust loads of up to 21 MN and is still in use. The bearing with two rows of pads was rotating whereas collar was stationary. The collar was located on a toroidal surface for self-aligning in the radial direction. Pads with radial ribs were installed in two coaxial ring shape recesses and hold in place by pins from falling down on the collar when the shaft was lifted. In each recess there were flat steel rings with separators in between, forming flexible elements. This flexibility helped in load equalisation. Despite bulky construction these bearings showed good reliability, which was probably due to the efficient heat transfer from the pads rotating in the bath with oil.

Large bearings are at risk during starting and stopping periods or in reversing when the velocity changes direction. Boundary lubrication that occurs during these periods can lead to severe wear and bearing wiping. A common solution is the introduction of oil jacking system that delivers high pressure oil between the bearing surface and the collar. An important design consideration is that the jacking system oil inlets should have the least possible derating effect on the normal hydrodynamic lubrication. This problem was considered by [23,29, 30]. It was found that a circular oil recess had a negative effect on bearing operation affecting hydrodynamic pressure in such a way that in some conditions pad could be forced to tilt backward. A ring shape (see figure 1) or an "eyebrow" [30] configuration was found to be a better solution. Utilisation of a magnetic bearing instead of the oil jacking system was described in [31].

The third group of methods for reduction of thermal distortion is related to pad support structure design. A common method is to vary pad support extent. A single concentrated support yields too high peak pressure in large pads and very thick pads are therefore required to avoid bearing damage [24]. However, thin pads are desirable from many standpoints, such as handling, babbitting and the ability to conform to variations in the collar surface. High peak pressures or the use of thick pads can be

avoided by supporting the pads on a large number of points to obtain an optimum pressure and oil film profile. Disc-type support is widely used with good results [23,32]. Better load distribution between bearing pads is achieved if there is an elastic element in the support. Pivot-spring type support combines radial pivot with such a flexible element [22]. For very large pads a multiple support is recommended. The pad can be supported on beams [24,33], discs [33] or areas [20].

Another common support structure is based on springs. The spring supported thrust bearing was invented around 1915 and was initially a thin steel, washer type, ring mounted on many stiff precompressed springs [34]. The ring had one radial gap and radial surface grooves. Collar had also radial grooves for better oil circulation. Despite costly maintenance these original thrust bearings are still in operation carrying loads of up to 3.5 MN (specific load 3 MPa) [23]. Later the ring was replaced by segmented pads. In the early designs the pad was made thin with springs spaced uniformly over the whole back face of the pad. The bearing operated at high temperature so that thermal crowning could compensate the mechanical concave deformation caused by load. In modern designs, such as shown in figure 1, fairly thick pads are used.



Figure 1. A modern spring supported thrust bearing (Courtesy of ALSTOM Power Sweden AB).

Due to the springs load distribution between the pads presents no difficulty but pad tilting and oil film formation are less than satisfactory. To achieve a better film formation springs under the pad leading edge can be removed. Performance of spring supported bearings at loads less than spring preload are analysed in [35]. Measured film thickness, pressure and temperature were reported in [23,36] whereas the limit of hydrodynamic lubrication at low collar speeds was sought in [37].

Recent development is a bearing with pivoted pads of the double layer design with elastic pins [38]. The upper thin plate is supported above the lower thick pad by means of a number of vertical pins of various diameters and thus different elasticities. The resulting deformation of the thin plate is determined mainly by the shortening of the pins under load so that the plate moves down parallel to it original position. This design comprises advantages of spring-supported and pivoted pads in regard to oil film formation and load distribution between the pads.

3 ALIGNMENT

Misalignment that always occurs to some extent in any rotating machinery due to manufacturing tolerances and thermomechanical distortions results in unequal pad loading. The most highly loaded pad may be damaged causing subsequent failure of the remaining pads. To cope with misalignment a conventional bearing can be mounted on a spherical seat. However, under load the friction at the seat surface may be high enough to make this method ineffective. A better compensation for misalignment as well as for minor random differences in height of the pads is provided by an equalising system.

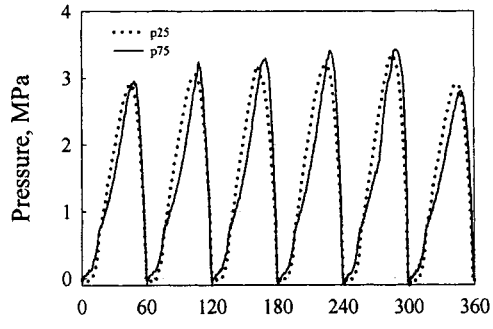


Figure 2. Pressure profiles along circumferential lines located at 25% and 75% of pad width counted from the pad inner radius. Specific load 1.5 MPa.

Each pad in a bearing with the mechanical equalising system has its own upper levelling plate

and shares two lower levelling plates. Equalisation is achieved when the levelling links lower the overloaded pad and raise the underloaded pad. As the levelling links intermesh, the load on adjacent pads is equalised. However, friction at the contacts of the links hinders a complete load equalisation. Typical pressure profiles across a six-pad bearing reveal that there is some unequal pad loading left despite the equalising system (figure 2). Friction at the contacts can also lead to redistribution in pad loading in transient operating conditions as shown in [39]. Friction between the links can be reduced by changing the geometry of the levelling mechanism. Tests of such an improved system in a static rig were reported in [40]. If the collar is not perpendicular to the shaft a "swash" effect will occur. In this case even the most sensitive mechanical equalising system won't do the work. The effect of a "swashing" collar on bearing operation was investigated in [41].

In large bearings alignment is achieved by manual adjustment of the pad position. Further alignment is provided by deformation of the various flexible elements in the pad support structure. In some units there can be a hydraulic equalising system [23,42]. Such system is the most efficient. As shown in [43] a collar run-out of 100 μm caused pad load variation of 100 kN for a single row bearing with a disk-type support, 50 kN for a double row bearing and 15 kN for a bearing with the hydraulic system.

4 BEARING MATERIALS

Depending on operating conditions, a wide range of metal alloys can be used on the working surface of bearing pads that are normally made of steel. Babbitt or white metal invented by Isaac Babbitt in 1839 [1] remains the most commonly used bearing surface material. It has exceptional anti-seizure characteristics, good conformability, embeddability and corrosion resistance. Tin and lead base alloys are in use. A drawback with tin base babbitt is thermal ratcheting that occurs when the material is thermally cycled [44, 45]. Lead base babbitt does not have this negative feature but is more prone to corrosion. Babbitt main weakness is poor fatigue strength especially when it is applied in the thickness of 0.25 mm and over [46]. The strength also falls with rising temperature [47] so that in practice it is not recommended to allow pad surface temperature of more than 130°C [7]. To improve the strength various alloying elements were tried such as cadmium, nickel and chromium. In 50's and 60's tin-aluminium alloys were introduced. Their strength is not temperature dependent which permits the operating temperature to be raised slightly. They also have very good corrosion resistance. As shown in [45] these alloys can withstand an attack of EP additives that affect babbitt.

A high thermal conductivity pad material can be used to dissipate the heat developed in the oil film, and therefore maintain maximum strength of the babbitt. Influence of various copper based pad materials on bearing thermal performance was studied in [48]. Pads of high strength copper were also investigated in [11] where in addition test results for pads of solid aluminium were reported. Lower temperatures, compared with steel pads, observed in the tests of copper alloys and aluminium pads resulted in a reduced amount of thermal crowning, higher oil viscosity and thicker oil films.

In the same reference [11] interesting information was given on the performance of a bearing with offset pads lined with a graphite fiber composite. Temperature measured by a thermo-couple mounted flush with the pad face was found to be 3 to 10°C lower compared to the similar babbitted pads. This was attributed to the higher thermal insulation of the composite. No data were given on temperature gradients in the pad with the composite lining though it was mentioned that a very low temperature difference existed across the pad backing.

Apart from controlling thermal crowning bearing performance can be improved by using a non-rigid working surface. A compliant-surface three-pad bearing operating in an oil bath was studied in [49]. Pads were faced with a 0.368 cm elastomer layer. Friction, pressure and film thickness profiles were obtained for a range of pivot locations. It was concluded that such bearings had a number of advantages over conventional rigid surface bearings. These are higher tolerance to dirt, misalignment and distortion. In addition, the elastic nature of the facing material helped in reducing oil escape from the loaded zone. Bearing safety was thus increased especially in the conditions of reduced oil feed.

Thus, laboratory tests indicate that it is of advantage to have a pad facing made of a material with good elastic properties and low thermal conductivity. There is a unique bearing material that

is already widely used in dry sliding bearings. It is polytetrafluoroethylene (PTFE). This material has a low coefficient of friction, the broad temperature range, excellent anti-seizure properties, and the superior resistance to chemical attack and moisture. PTFE will not chemically bond to anything else so the main problem is how to make this material adhere to the steel surface of the bearing pad.

Various bonding technique have been proposed over the years. Some researches tried adhesives [50] other used mechanical bonds [51-53]. A porous bronze layer can be bonded to a steel backing and impregnated with PTFE as shown in [51]. In this composite PTFE thickness is limited to about 100 μm but it can be used with various fillers or as a filler itself depending on application. No published data is available on the effect of this facing on thrust bearing operating characteristics.

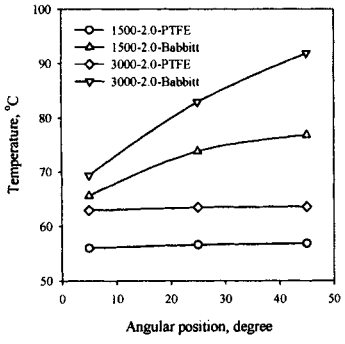


Figure 3. Temperatures along the pad central line at 2.0 MPa specific load and 1500/3000 rpm shaft speeds [52].

Another approach was reported in [52]. Operating characteristics of equalising tilting pad thrust bearings with babbitt and PTFE composite facings were compared and analysed. Each bearing arrangement included six pads with 60% offset. Thickness of the PTFE composite facing was approximately 1.5 mm. Pad temperature distributions, collar temperatures and bearing friction torque were continuously measured. Test results (figure 3) showed the outstanding feature of the PTFE bearing: almost no temperature rise across the pad from the leading to the trailing edge. The PTFE composite provided excellent thermal insulation (figure 4) so that pad thermal crowning was eliminated. In the tests PTFE-faced bearings operated with lower power loss and slightly higher collar temperatures compared to similar babbitted bearings.

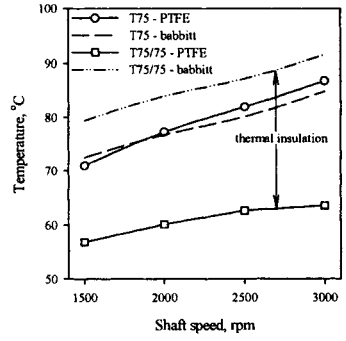


Figure 4. Pad T75/75 and collar T75 temperatures versus shaft speed at 2.0 MPa specific bearing load [52].

A further insight into thermal performance of PTFE faced bearings can be gained through a TEHD model. Figure 5 shows temperature distributions in the 2 mm facing and in the pad of a spring-supported bearing. The facing material has the same thermal properties as PTFE and ten times lower modulus of elasticity compared to babbitt. It can be seen that there is only a slight temperature gradient in the pad whilst temperature difference in the facing reaches 45°C. If babbitt is used instead of PTFE maximum temperature difference in the facing will be 3°C. Thus, thermal distortion of the PTFE faced pad is minimised, which is beneficial for large bearings.

Performance of large bearings can be further improved if PTFE is applied in combination with metal wire mesh as proposed in [53]. Sheet PTFE was applied under high pressure to a cushion of metal wire mesh so that the PTFE flowed into gaps in the wire mesh to a depth of 1 to 2 mm. In this way a firm bond was established between the PTFE and the cushion. This composite was then soldered to the steel pads and machined to a total liner thickness of 10 mm that included 2 mm of pure PTFE, 1 to 2 mm of combined PTFE-wire mesh and 6 to 7 mm of wire mesh (figure 6). Wire can be made of steel, copper or bronze. The bronze is preferable since it will serve as a good bearing material in the event PTFE is lost due to excessive wear.

This facing has proved to be very efficient for large thrust bearings due to its resilient nature and low friction working surface. Bearings with this

facing have been in successful use in hydroelectric power stations of the former Soviet Union since 1974 [54] and allowed for specific bearing loads of up to 11 MPa in continuous operation [55]. Such loads are much higher than those for babbitted bearings. This is a significant advantage as smaller bearings can be produced. Deformations are lower in smaller bearings, which makes their design easier. In addition, power loss will be lower.

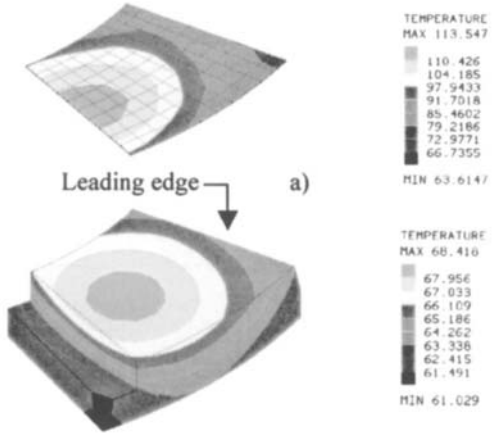


Figure 5. Temperature distributions in the pad facing (a) and in the pad backing (b) of a spring supported bearing at specific load of 3.8 MPa and speed of 600 rpm.

Resilient nature of the PTFE and wire mesh contributes to better load equalisation between bearing pads and enables higher degree of misalignment to be accepted. More relaxed tolerances are also allowed [56]. If a pivot is used for pad support then this design has advantages of both spring-supported and pivoted-pad bearings.

Utilisation of the PTFE composite as a bearing facing provides essential advantages over conventional babbitted bearings that ultimately reduce maintenance and prolong bearing service life. PTFE has low coefficient of friction so that oil jacking system can be dispensed with. It has unique chemical stability and high temperature capability, which is important for bearings operating in high temperature and corrosive environments and applications where thermal cycling, resulting from regular changes in operating conditions, is prevalent. This material is ideal for bearings used in electrical machinery due to the excellent dielectric properties. In addition, a risk for bearing damage caused by electrostatic voltage or electromagnetic currents is eliminated.

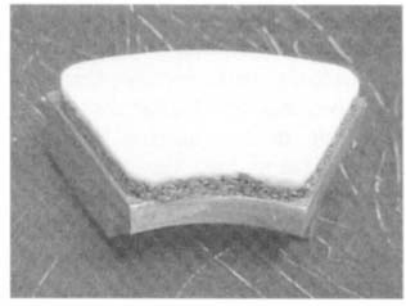


Figure 6. Pad with PTFE wire mesh facing.

Given the numerous advantages of this design it is not surprising that this bonding technology has been adopted by bearing manufacturers in China [3], UK [57] and Japan [58]. They have further developed it, mainly by applying PTFE with reinforcing fillers [58,59]. Various fillers have been tried to improve wear resistance and compressive strength of pure PTFE. The glass filled PTFE gave the best performance in terms of wear [58,59] and friction [58]. Glass filled PTFE was found to scratch the counter-surface [59] whereas according to [58] wear was negligible. Consequently, carbon filled PTFE was chosen by [59], and glass filled PTFE that showed lower friction than pure PTFE was preferred by [58].

Although PTFE faced pads can sustain a load of up to 20 MPa without failure [59], care is required in their design. Oil film pressure produces a concavity in the soft facing. The concavity is useful in the radial direction as the escape of pressurized oil from the film is retarded. In the circumferential direction this deflection can affect the crown that is essential for centrally pivoted pads. Tapers at the leading and trailing edges are thus required to allow the necessary crown to form [60]. Their dimensions should be optimised for a required range of loads taking into account elasticity of the PTFE and wire mesh that also affects formation of the concavity.

There exist some other concerns to bearing designers and end users. What happens to the bearing and collar surface in case of a failure? According to [56] damage of one pad did not lead to

subsequent failure of the rest of the pads or to collar scoring. This was established for pure PTFE but further tests are necessary for PTFE composites. A possibility of slip at the oil-PTFE interface due to low adhesion of oil to PTFE is another concern. It may be useful to comment here that the rotating bearing described in Section 2 shows no signs of distress after being fitted with PTFE facing [55].

Some operating data, mostly temperature, for PTFE faced bearings are available in the literature [61-63] though to the authors' knowledge no detailed investigation has been published so far. Further theoretical and experimental work is therefore required for improved design of the PTFE-faced bearings and better understanding of their behaviour.

5 SURFACE FINISH AND TEXTURE

In early days the surface of the bearing pads was usually scraped to a surface plate. This procedure of achieving flatness is still common today in manufacturing babbitted pads of large bearings. As an added benefit scraping provides subtle indentations in the babbitt surface where oil can be trapped. This oil will lubricate bearing during start-up reducing the breakaway torque required to establish hydrodynamic operation.

Surface finish of both collar and pads is known to be especially important for bearings that must accept starting under load. The effect of collar surface finish on friction coefficient during start-ups under load was investigated in [64]. It was concluded that a smooth surface was necessary during the starting period. Indeed, if the collar is too rough then the wear process of babbitted pads cannot occur gradually enough to avoid significant damage and seizure is likely to occur.

Smooth collar is also of importance in the hydrodynamic regime. It was shown in [11] that an improvement in collar finish from Ra0.4 to Ra0.15 significantly increased the ultimate bearing load carrying capacity. However, the surface of the scraped pads wears smooth over time. As soon as it is worn flat high adhesion between too smooth collar and babbitt surface can occur resulting in bearing seizure during start-ups [65]. One remedy is to introduce an externally pressurized oil lift system. However, this solution raises another technical problem of increased maintenance work.

It follows then that the collar and bearing surface texture must be arranged to suit the conditions of boundary and hydrodynamic regions. For boundary lubrication small pockets are preferable on the collar surface, as it does not wear out, which prolongs bearing life. Information available in the scientific literature on the effect of a dimpled surface on bearing operation in the hydrodynamic regime is limited so that even data related to journal bearings are relevant. It was reported [66] that dimples machined on the shaft surface allowed a significant increase in load carrying capacity of a journal bearing operating at high speeds (figure 7). Better heat removal from the loaded zone due to more efficient oil circulation was thought to be the explanation. At low speeds no pressure generation was found over the dimples, which explained a reduction in load carrying capacity.

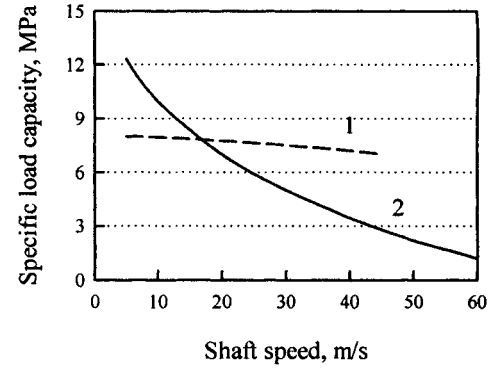


Figure 7. Specific ultimate load carrying capacity of a journal bearing with (1) and without (2) dimples on the shaft surface. Dimple dimensions: depth 12 μm, width 0.6 mm, relative area 40% [66].

Somewhat contrary results were obtained when a dimpled-surface thrust bearing was tested at low speeds (below 1 m/s) and loads of up to 3.45 MPa [67]. The bearing was of a bending-pad type with three pads rotated in an oil bath with a stationary collar. The measured pressure profiles had bumps in the converging part of each dimple. It was concluded that they could improve load capacity. The study also revealed that depending on Sommerfeld number the friction coefficients of the dimpled bearing were up to 7.3% lower than those associated with a plain surface. Unfortunately, dimple dimensions were not given.

Nevertheless, it is evident that there are surface textures, which improve bearing performance. Further investigations will help in optimisation of the pocket dimensions, their distribution and shape. A good example of how an optimised texture improves bearing operation was reported in [68]. A continuous spiral microgroove made on the liner surface of a journal bearing was found to improve its running-in and anti-seizure performance. Lower bearing temperature was also observed.

6 BEARING LUBRICATION

6.1 Lubricant supply methods

There are two common types of oil feed arrangements. In the first case, the bearing is submerge in an oil bath and is free to ingest as much oil as it requires. Oil bath temperature is kept constant by a cooling system and oil circulation is provided by a separate pump or by a "self-acting" centrifugal pump arranged in the thrust block [35]. This arrangement is usually used in large bearings. The effect of oil viscosity on power loss of a bearing with spring-supported pads was studied in [69].

The second type uses a controlled oil flow to the bearing in an amount deemed suitable for operation. A standard oil flow is determined by the condition of maintaining a maximum temperature rise of 15-20°C from supply to discharge [7]. Recommended supply oil temperature is in a range of 40-50°C [7].

There are three different configurations of the controlled flow arrangement represented by outlet orifice, throttled inlet and directed lubrication. Comparative studies of the effect of these lubrication systems on bearing performance were reported in [70,71]. A common way is to use a restricted discharge or control orifice at the collar outer diameter. In this mode of lubrication the entire bearing chamber is filled with oil and churning of the oil by the collar introduces additional power losses and bearing heating. The effect of supply oil temperature on bearing operating temperatures and power loss was investigated by [72]. Design factors affecting power loss of a flooded bearing were studied in [73]. At high shaft speeds churning losses are a significant part of the total power loss [74]. They can be decreased by using lubrication with throttled inlet. The oil flow to the bearing is restricted and a free discharge is provided. Whilst this method was found to reduce power losses it is too sensitive to oil flow and below a critical flow the surface temperature rises rapidly [45].

The most efficient system is directed lubrication introduced in the end of 1960's [75]. The oil was sprayed at high speed on to the collar surface between the pads of the freely drained bearing. The power and pad temperatures were considerably reduced compared to the flooded design. Thicker oil films were also observed [70]. Various distribution systems for the spray jets were tested. Umbrella nozzle system was studied in [70] while the spray bars were tested in [71].

Directed lubrication with a leading edge groove system was introduced in 1980's. The system performance was described in a series of papers [71,76,17]. The oil was supplied directly into the oil film from a groove made in the pad near the leading edge. This system was found to be more efficient than the spray-feed arrangement [71]. A possible explanation by the thermal boundary layer theory was given in [77].

At low shaft speeds directed lubrication is less efficient in terms of decreasing power loss but it allows a significant reduction in pad temperature [76] and may be useful in large bearings to decrease pad thermal crowning.

6.2 Lubricant choice

The most commonly utilised and versatile lubricants are mineral oils. Viscosity is the main oil characteristic. The effect of oil viscosity grade on bearing operating limits was considered in [7]. From the point of view of minimal power loss and pad temperature thinner oils are recommended at high speeds. Thicker ones are usually used at low speeds to avoid too thin oil films.

Viscosity is a strong function of temperature. It is expected that the flatter the viscosity-temperature relationship, the higher the load carrying capacity. The simplest way to flatten this relationship is to enhance mineral oils with viscosity index improvers. No published experimental data is available on the effect of viscosity index improves on thrust bearing operation.

Another way of improving the viscosity-temperature relationship is to use synthetic oils that usually have higher viscosity indices. In addition, some synthetic oils are environmentally adapted which is very important in applications where substantial amounts of oils are used such as in

hydroelectric power stations. The oil may be released, whether through accident or system failure, into a sensitive river ecosystem causing adverse effects.

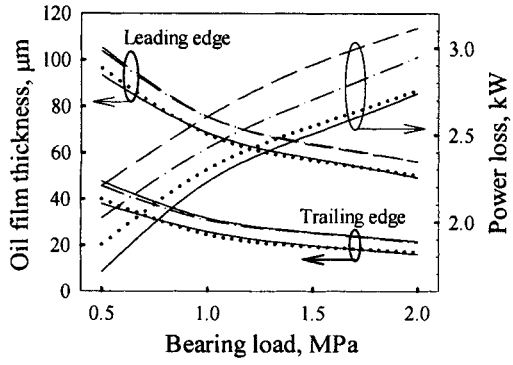


Figure 8. Effect of oil type on film thickness and power loss [80]: poly-α-olefin VG46 (solid), mineral VG46 (dotted), ester VG46 (dash-dot), mineral VG46 (dashed).

The laboratory evaluation of adipate diester synthetic oils in a centrally-pivoted tilting-pad thrust bearing was reported in [78]. The bearing was tested in flooded and directed lubrication modes. In another comparative study of ester base and poly-α-olefin oils [79] an equalizing bearing with offset pads was used. Results of these investigations indicate that it is possible to increase bearing efficiency without sacrificing its safety by substituting a mineral oil with an ester base synthetic oil of lower viscosity grade. Alternatively, bearing safety can be improved if an ester base synthetic oil is used instead of a mineral oil of the same viscosity grade. Results from the tests [80] comparing the effects of different oil types are shown in figure 8.

7 MONITORING TECHNIQUES

There are three basic operating characteristics that need to be considered in the bearing design and limited in order to insure satisfactory operation. These are: maximum pad temperature, minimum film thickness, and peak oil film pressure [7]. The last two of these are difficult to measure due to the lack of simple and robust sensors and problems with their mounting. Test data available on film thickness and pressure is limited. Film pressures were reported in [36,81,82] while oil film information was given in [26,75,78,80,83-85].

Since temperatures may be measured relatively easy, temperature monitoring is now a well-established technique to monitor bearing conditions. Choice of the temperature sensor, its position in the pad, methods of installation and alarm settings were discussed in [86]. Additional information can also be found in [87-88]. Several authors presented test data on bearing temperatures. References [81-86] are representative. Generalisation of some published and new results was made by [18] to see if there was a consistent overall pattern.

Utilisation of PTFE based bearing surface materials may require modification of the existing and development of new monitoring techniques. PTFE is a powerful insulator so that method of sensor installation must be changed to obtain adequate sensitivity to temperature in the film. A possible solution was presented in [94]. New techniques are also required for film profile measurements as PTFE is a non-conductive material. Application of an optic lever technique to film thickness measurements in PTFE faced bearings was reported in [95].

As for conventional babbitted bearings, temperature and oil film sensors can be combined to provide a very convenient detection system as shown in [96].

CONCLUDING REMARKS

The existing designs and current developments of tilting pad thrust bearings have been reviewed based on published experimental papers and design studies. Design modifications made to improve bearing operation and reliability depended on a range of operating speeds and loads. In high speed bearings the major developments were direct lubrication system and new pad surface and backing materials. The result was significantly reduced pad temperatures and power loss. In heavily loaded bearings the most important innovation was utilisation of PTFE as a pad facing. This gave a significant increase in load carrying capacity without sacrificing bearing safety. Other numerous advantages were also attained due to the outstanding properties of PTFE.

It is anticipated that future trends in bearing research and development are as follows:

- adoption of PTFE or its composites as bearing facings,
- utilisation of microdimples on bearing or collar surfaces,
- wider use of environmentally adapted synthetic oils,
- industrial introduction of more versatile bearing monitoring systems.

ACKNOWLEDGEMENT

The financial support of the Swedish National Board for Industrial and Technical Development (NUTEK), ALSTOM Power Sweden AB, Elforsk AB, Kingsbury Inc., Luleå University of Technology, Mobil Oil AB, Statoil Lubricants and Svenska Bearing AB is gratefully acknowledged.

REFERENCES

1. D. Dowson, History of tribology, London, 1998.
2. J.E.L.Simmons and S.D.Advani, Proc. Of the 16th Leed-Lyon Symp. on Tribology, (1989) 49.
3. D.Schafer et al, Proc. Int. Conf. On Large High Voltage Electric Systems CIGRE, (2000).
4. M.M. Khonsari, ASLE Trans., 30, No.1, (1987) 19.
5. O. Pinkus, Thermal Aspects of Fluid Film Tribology, ASME Press, NY, 1990.
6. A.Raimondi, Trans. ASLE, 3 (1960) 265.
7. A.J.Leopard, ASLE Lubr. Eng., 32 (1976) 637.
8. G.K.Rightmire, ASME J. Tribology, 106 (1984) 456.
9. P.D.Choudhury, M.R.Hill, D.J.Paquette, Proc. 21st Turbomachinery Symp., (1992) 578.
10. D.Jain, F.Zeidan, M.Wittmeyer, ASME Paper 94-GT-190.
11. W.W.Gardner, J. Lubr. Tech., 97 (1975) 430.
12. D.R.Garner, N.H.New, Proc. AE Symp., 1982, Paper N 13.
13. A.Raimondi, J.Boyd, Trans. ASME, 77 (1955) 287.
14. W.W.Gardner Proc. 13th Leeds-Lyon Symp. on Tribology, (1986) 113.
15. W.W. Gardner, ASME J. Tribology, 110 (1988) 609.
16. J.K.Whalen, Proc. 25th Turbomachinery Symp., (1996) 17.
17. A.M.Mikula, ASME Journal of Tribology, Vol. 110, 1988, pp. 174-180.
18. D.Horner, J.E.L.Simmons, S.D.Advani, STLE trans., 31 (1988) 44.
19. C.M.M. Ettles, Wear, 79 (1982) 53.
20. W.S.Chambers, A.M.Mikula, Trans. STLE, 31 (1988) 61.
21. Thrust bearing for hydraulic turbine generators, Fuji Electric, 1977.
22. Features of modern hydro generator design, Hitachi Ltd., 1982.
23. A.E.Alexandrov, Thrust bearings of hydro-generators, Moscow, 1975.
24. R.A. Baudry, E.C. Kuhn and W.W. Wise, ASME Trans., 80 (1958) 807.
25. H.C.Bahr, ASME J. Eng. Power, 83 (1961) 130.
26. K.Kawaike, K.Okano, Y.Furukawa, ASLE Trans., 22 (1979) 125.
27. E.C.Kuhn, Power, (1971) 100.
28. P.S.Johrde, Proc. Int. Conf. On Large High Voltage Electric Systems CIGRE (1976).
29. D.V.Nelson, M.C.Plummer, R.C.McCulloch, Proc. Americam Power Conf, 46 (1984).
30. M.Hemmi, T.Inoue, Tribology Trans., 42 (1999) 907.
31. Magnetic thrust bearing for generator-motor, Fuji Electric, 1981.
32. K.Tazuke, A.Ishii, A.Yamamoto, Mitsubishi Denki Engineer, December, 1969.
33. C.M.Ettles, H.G.Anderson, J. Tribology, 113 (1991) 405.
34. S.Abramovitz, Discussion to C.M.Ettles, ASME J. Tribology, 113 (1991) 632.
35. C.M.Ettles, J. Tribology, 113 (1991) 626.
36. J.H.Yuan, J.B.Medley, J.H.Ferguson, Tribology Trans., 42 (1999) 126.
37. A.L.Brown, J.B.Medley, J.H.Ferguson, Proc. 26th Leeds-Lyon Symp. on Tribology, (1999).
38. K.Fiesenig, M.Starcevic, A.Fust, Brown Boveri Rev., 70 (1983) 258.
39. S.B.Glavatskih, Proc. 26th Leeds-Lyon Symp. on Tribology, (1999) 229.
40. W.W.Gardner, Proc. JSLE international conf., Tokyo, Japan, (1985) 61.
41. A.K.Tieu, J. Tribology, 113 (1991) 633.
42. Neyrpic hydraulic turbomachines, Water Power, November (1972).
43. A.E.Aleksandrov, Gidrotekhnicheskoe stroitel-stvo, 3 (1980) 39.
44. J.M.Conway-Jones, A.J.Leopard, Proc. 4th Turbomachinery Symp., (1975) 55.
45. A.Hill, Trans. IMarE, 88 (1976) 204.

46. F.Y.Zeidan, B.S.Herbage, Proc. 20th Turbomachinery Symp., (1991) 161.
47. E.R.Booser, F.D.Ryan, C.L.Linkholler, Lubrication Engineering, 26 (1970) 226.
48. A.M. Mikula, Lubr. Eng., 44 (1988) 969.
49. G.K. Rightmire, V. Castelli, D.D. Fuller, J. Lubr. Techn., 1 (1976) 95.
50. T.R.Choudhary et al., Proc. Instn. Mech. Engrs., 214 Part J (2000) 375.
51. Plain bearing designers handbook, The Glacier Metal Co.
52. S.B.Glavatskih, ASME Paper 2002-Trib-137.
53. A.M.Soifer, D.S.Kodnir, Yu.I.Baiborodov, Izvestia VUZov, Mashinostroenie, 7 (1965) 67.
54. A.E.Aleksandrov, Gidrotekhnicheskoe stroitelstvo, 9 (1981) 12.
55. Private communication with A.E. Aleksandrov.
56. A.E.Aleksandrov, N.G.Platonov, Elektricheskie Stantsii, 9 (1988) 58.
57. J.E.L.Simmons, R.T.Knox, W.OMoss, Proc. Instn. Mech. Engrs., 212 Part J (1998) 345.
58. S.Uno, et. al., Japanese journal of tribology, 42 (1997) 205.
59. R.T.Knox, Hydropower & Dams, 6 (1999) 62.
60. Baiborodov, Yu. I. et al., Gidrotekhnicheskoe stroitelstvo, 6 (1982) 42.
61. Baiborodov, Yu. I. et al., Gidrotekhnicheskoe stroitelstvo, 10 (1977) 28.
62. A.E.Aleksandrov, Gidrotekhnicheskoe stroitelstvo, 9 (1981) 12.
63. A.E.Aleksandrov, N.G.Platonov, Gidrotekhnicheskoe stroitelstvo, 11 (1990) 20.
64. C.M. Laffoon, R.A. Baudry, P.R. Heller, Trans. ASME, 69 (1947) 371.
65. R.A.Baudry, D. N.Reilly, Hydro review, (1991) 126.
66. F.P.Snegovskii, N.G.Bulyuk, Trenie i Iznos, 4 (1983) 322.
67. G.K.Rightmire, R.G.Bill, H.G.Anderson, Proc. of the 14th Leed-Lyon Symp. on Tribology, (1987) 139.
68. Y.Kumada, K.Hashizume, Y.Kimura, Tribology Trans., 39 (1996) 792.
69. J.H.Ferguson, J.H.Juan, J.B.Medley, Proc. 24th Leeds-Lyon Symp. on Tribology, (1997) 187.
70. N.H.New, ASME J. Lubr. Techn., 96 (1974) 22.
71. A.M.Mikula and R.S.Gregory, ASME J. Lubr. Techn., 105 (1983) 39.
72. A.M.Mikula, ASLE Trans., 30 (1987) 220.
73. R.S.Gregory, J. Lubr.Techn., 101 (1979) 154.
74. F.A.Martin, Tribology Convention, Proc. Instn. Mech. Engrs., 184 (1970) 120.
75. A.J.Leopard, Tribology, November (1970) 206.
76. A.M.Mikula, ASME J. Tribology, 107 (1985) 423.
77. C.M.Ettles, Discussion to A.M.Mikula and R.S.Gregory, ASME J. Lubr. Techn., 105 (1983) 45.
78. N.H.New, R.H.Schmaus, Proc. 12th Turbomachinery Symp., (1983) 139.
79. S.B.Glavatskih, M. Fillon, R. Larsson, J. Tribology, 124 (2002) 377.
80. S.B.Glavatskih, R. Larsson, Proc. 2nd World Tribology Congress, (2001).
81. K.W.Kim, M.Tanaka, Y.Hori, Japanese Journal of Tribology, 40, (1995) 1.
82. S.B.Glavatskih, J. Tribology, 123 (2001) 608.
83. R.C.Elwell, R.E.Gustafson, J.C.Reid, ASME J. Basic Engineering, 86 (1964) 483.
84. R.S.Gregory, in Super Laminar Flow in Bearings, Mech. Eng. Publications, Suffolk, 1977, England.
85. M.Tanaka, Y.Hori, R.Ebinume, Proc. JSLE international conf., Tokyo, Japan, (1985) 553.
86. D.R.Garner, A.J.Leopard, Proc. 13th Turbomachinery Symp., (1984) 133.
87. R.C.Elwell, Machine design, June 24 (1971) 79.
88. R.C.Elwell, Machine design, July 8 (1971) 91.
89. R.S.Gregory, ASME J. Lubr. Tech., 96 (1974) 7.
90. J.W.Capitao, R.S.Gregory, R.P.Whitford, J. Lubr. Techn., (1976) 73.
91. J.W.Capitao, J. Lubr. Techn., (1976) 81.
92. P.B.Neal, Proc. 6th Leeds-Lyon Symp. on Tribology, (1980) 137.
93. A.M.Mikula, ASLE Trans., 29 (1986) 173.
94. S.B.Glavatskih, Proc. Int. Conf. AUSTRIB2002.
95. M. Höbel and K.Haffner, Meas. Sci. Technol., 10 (1999) 393.
96. S.B.Glavatskih, Ö.Uusitalo, D.Spohn, Tribology International, 34 (2001) 853.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Study on the development of a thin-film sensor for measuring oil-film pressure in engine bearing and piston pin-boss

Yuji Mihara[a], Kazuyuki Sekine[a] and Tsuneo Someya[b]

[a] Musashi Institute of Technology
1-28-1 Tamazutsumi, Setagaya-ku, Tokyo 158-8557, Japan

[b] University of Tokyo, Professor emeritus
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan

In order to measure accurately the oil-film pressure in machine elements such as plain bearings and rolling contact parts without reducing the rigidity and changing the configuration of sliding surface, we have developed a piezo resistive-type thin-film pressure sensor. Using physical vapour deposition technique (PVD-sputtering), a sensor was deposited with a thickness of approximately 3~5μm on the sliding surface. The electrical resistance of such pressure sensors varies under pressure (piezoresistivity). But when used in machines such as internal combustion engines, sensors are subjected to extraneous influences causing measurement errors such as temperature change and deformation of the sliding surface. In order to reduce these errors, the material and the shape of the sensor were devised: To reduce first the temperature dependence of pressure sensitivity, alloy composition of the sensor was optimized. To reduce the temperature dependence at higher temperature caused by oxidation of sensor material, aluminium-nitride film, instead of oxide films, was applied successfully as an insulating and a protection film. Then, to reduce the error due to sensor deformation, a new sensor shape was devised, that possesses a uniform and low strain sensitivity. This has enabled to correct for error properly or even made the correction unnecessary for practical use. Using this newly-developed sensor, the distribution of oil-film pressure was measured in engine main bearing and in piston pin-boss.

1. INTRODUCTION

Pressure measurement using a piezo-resistive method was first introduced by Bridgeman [1] who used metals, especially an alloy of copper, manganese and nickel for this purpose. Using this alloy, Kannel et al. [2] developed a thin-film pressure sensor to measure the oil-film pressure between two rollers. This sensor had the advantages of very small dimensions and of enabling the measurement of high pressure exceeding 1000[MPa]. It has, therefore, been used especially in research of the elasto-hydrodynamic lubrication (EHL) of machine elements such as rollers and cams [3]-[6]. Using the same technique, the present authors have measured oil-film pressure using thin-film sensors deposited on to the sliding surface of engine bearings [7]-[9]. However, temperature change and deformation of the sliding surface affected the sensor output in engine operation, and made the correction necessary, as the oil-film pressure itself in the plain bearing and piston pin-boss is of the order of only 10MPa to 100MPa. Therefore, we have developed sensor alloys to decrease the error caused by temperature and have then devised the sensor shape to decrease the error caused by strain [10].

In this paper, we present the structure and shape of the new thin-film pressure sensor and results of oil-film pressure distribution measured in an engine main bearing and in a piston pin-boss. In addition, a new sensor structure is reported, using aluminum-nitride to reduce the temperature dependence of pressure sensitivity in high temperature range.

2. THE THIN-FILM SENSOR

2.1. Measurement principle

The electrical resistance of a thin-film sensor made of alloy changes under pressure. This change in resistance is converted into the voltage change through a Wheatstone bridge circuit and is amplified and recorded. The change in resistance is mostly attributable to the piezo-resistive effect, which is the change in electrical resistivity ρ of the alloy caused by pressure. The electrical resistance R of this sensor depends on pressure P, temperature T and strain ε : $R=(P,T,\varepsilon)$. In the linear range, the change in these variables will result in the change ΔR of the resistance R in the following form:

$$\Delta R/R=\alpha_P\cdot\Delta P+\alpha_T\cdot\Delta T+\alpha_\varepsilon\cdot\Delta\varepsilon \qquad [1]$$

where $\alpha_P=(dR/R)/dP$, $\alpha_T=(dR/R)/dT$ and $\alpha_\varepsilon=(dR/R)/d\varepsilon$ are the pressure sensitivity, the temperature sensitivity and the strain sensitivity of the thin-film pressure sensor, respectively. In applying equation [1] for pressure measurement, the second and third terms on the right hand side, which represent errors, should be negligibly small compared to the first term.

2.2 Structure and shape of the sensitive part

Figure 1 shows examples of the shape and basic structure of the developed thin-film sensor. Figure 1(c) is a sectional view of the thin-film sensor, where (1) is a substrate (for example, engine frictional parts such as bearing, piston pin and rocker arm); (2) is a 2~2.5 μm-thick SiO_2 or Al_2O_3 isolation film deposited on the substrate (1) for isolating the sensitive part (3) from the substrate (1); (3) is a 0.1~0.3 μm-thick alloy film that senses the pressure; and (4) is a protective film that protects the sensor alloy (3) during contact with the crank shaft and piston pin boss that are positioned opposite the thin-film sensor.

In the literature, the sensor shape shown in Figure 1 (a) has been used to measure the oil-film pressure in line contact such as in cam and rocker arm and two contacting rollers. The pressure-sensitive area of the sensor is 2~3 mm long and 15~30 μm wide. There are square lead parts and gold wires on both ends of the sensitive part. Using this type of sensor, we found that it is not possible to compensate for the error in measured pressure due to sensor deformation. To solve this problem, we have developed a new sensor shape of "twin-arc type" as shown in Figure 1(b). Applying this sensor, oil-film pressure in the engine main bearing, connecting rod big-end bearing and piston pin boss could be measured. This sensor consists of two arcs with a diameter of 0.4~0.8 mm and its external dimensions of 0.55~0.95 mm. The two semi-circle in the center serves as the lead film connecting these two arcs.

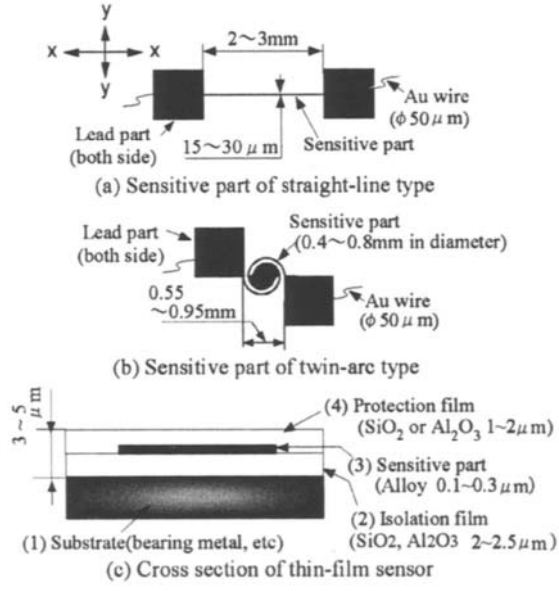


Figure 1 The structure of thin-film pressure sensor

2.3 Fabrication method of sensors

Figure 2 shows a radio frequency magnetron sputtering apparatus for the fabrication of the thin-film sensor shown in Figure 1. The substrate (2) which is engine frictional part is fixed onto the substrate holder (1) in the top of the sputter chamber. The target (3) consists of sputter materials such as Al_2O_3, SiO_2 and sensor alloys. The film was deposited under the conditions listed in Table 1.

Figure 3 shows the sensor fabrication steps. Step (1):The substrate is polished to provide surface roughness Ry in the range of 0.1~0.5 μm and sufficiently cleansed by ultrasonic equipment with organic solvents. Step (2):The substrate is then fixed in the sputter chamber for deposition of the isolation film (2) in Figure 1(c). Step (2)':In order to make a sensor of the shape shown in Figure1 (a) or (b), a resin photo-mask is prepared by photolithography. Step (3):After sputtering the sensor alloy, substrate

was taken out from the sputter chamber, and Step (3)' the resin mask is removed using organic solvents. Step (4): Subsequently, the protection film (4) in Figure 1(C) is deposited again by sputtering. Step (5): After all films are deposited, gold wires of φ50μm are welded to the lead part, as shown in Figure 1(a) and (b) using a micro welder.

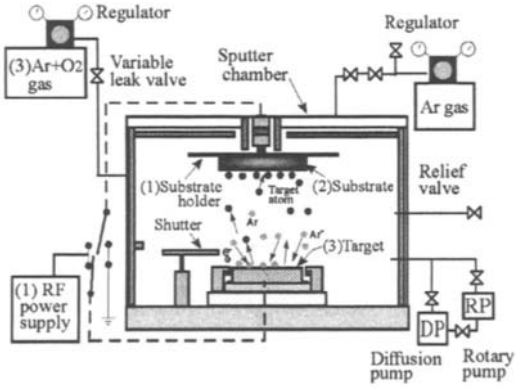


Figure 2 RF magnetron sputtering apparatus for sensor film deposition

Table 1 Sputtering condition

Vacuum attained	8×10^{-5} Torr (1.1×10^{-2}Pa)
Sputtering gas pressure	5×10^{-3} Torr (6.7×10^{-1}Pa)
Sputtering gas	100% Ar for Metal layer (Ni-Cr-Al, Cu-Mn-Ni, Cr) Ar+10%O2 for Isolation and protection layer(SiO2,Al2O3)
Sputtering power	400W for SiO2 , 700W for Metal
Pressure chamber volume	4700cm^3

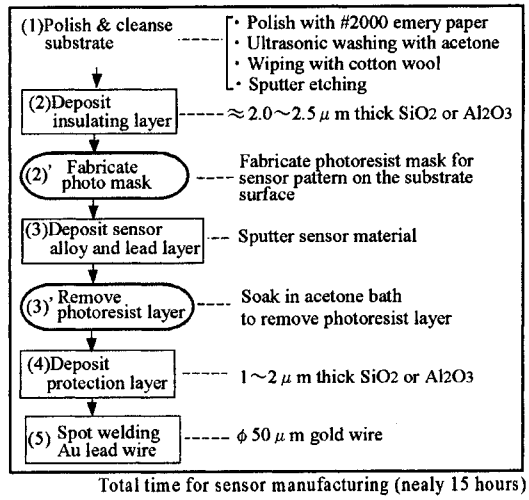


Fig. 3 Sensor fabrication steps (for aluminum alloy bearing metal)

3. CHARACTERISTICS OF SENSORS

3.1 Pressure and Temperature sensitivity of sensor alloy

Figure 4 shows the temperature sensitivity and the pressure sensitivity of some thin-film pressure sensors. The area (1) in Figure 4 is an alloy composed of 86Cu-12Mn-2Ni developed by a German research institute [4]-[6], while the area (2) is an alloy composed of 87.4Cu-9.6Mn-3.0Ni newly developed to reduce the temperature sensitivity of the sensor. As seen from the figure, the temperature sensitivity of alloy (2) is lower than that of alloy (1). Indeed, the temperature sensitivity of alloy(2) can be lowered even to zero. On the other hand, the pressure sensitivities of alloys (1) and (2) in Figure 4(b) fell in the range (18~22)× 10^{-6}[(Ω/Ω)/MPa]. We decided, therefore, to use the 87.4Cu-9.6Mn-3.0Ni alloy (2) with lower temperature sensitivity as our sensing material.

3.2 Characteristics of Ni-Cr-Al alloy for strain measurement

Error caused by the strain of the sensor is included in the pressure measured by a thin-film sensor, demonstrated by the third term of equation [1]. In order to evaluate this error, a Ni-Cr-Al alloy (3) shown in Figure 4 was used. The temperature sensitivity of this alloy (3) is (-4~0) × 10^{-6} [(Ω/Ω)/°C], while its pressure sensitivity is ±1×10^{-6} [(Ω/Ω)/MPa]. These values are significantly smaller

than those of the alloy (2) in Figure 4. Therefore, this alloy (3) is suitable to measure the strain of the sensor without being affected by temperature and pressure.

3.3 Strain sensitivity

Figure 5 gives the strain sensitivity of the thin-film sensors. The strain sensitivity depends on the direction of the strain ε. This fact should be expressed by the formula: Namely, writing the strain sensitivity α_ε as K_{sl} for the straight-line sensor in the x- or longitudinal direction and K_{sw} in the y- or lateral direction, respectively. For the twin-arc type sensor, they should be expressed as K_l or K_w. The last suffix C or N in K_{slC} and K_{slN}, etc. denotes the sensor alloy Cu-Mn-Ni or Ni-Cr-Al.

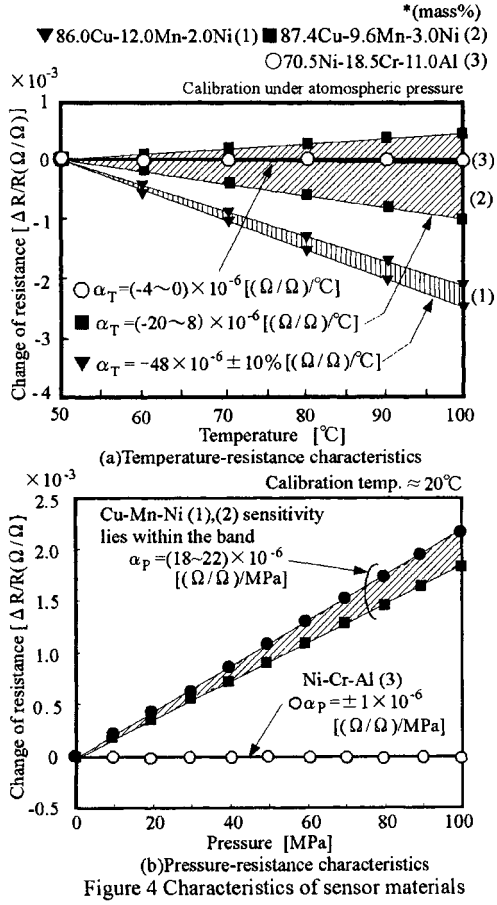


Figure 4 Characteristics of sensor materials

The straight-line type sensor has yielded very different strain sensitivity in two directions: K_{slC} =0.59 and K_{swC} = -1.19. On the other hand, twin-arc type sensor has yielded the same strain sensitivity in two directions: $K_{slC} = K_{swC}$ = -0.3. And these values are even much smaller than those of the straight-line type. Therefore, using the twin-arc type sensor of Cu-Mn-Ni, the hydrodynamic oil-film pressure can be obtained practically without correction in plain bearings, such as engine main bearing, connecting rod big-end bearing and piston pin-boss, because the strain itself is usually also small

3.4 Correction method for pressure measurement error

When the pressure sensor is deformed during the measurement, the pressure measured includes the error and has to be corrected as follows:

	Straight-line		Twin-arc	
(y)Lateral (x) Longi -tudinal	L=2~3mm, w=15~30μm		0.55~0.95mm, w=20~30μm	
Direction	(X)	(Y)	(X)	(Y)
Cu-Mn-Ni	K_{slC} =0.59	K_{swC} = -1.19	K_{lC} =- 0.30	K_{wC} =- 0.30
Ni-Cr-Al	K_{slN} =1.49	K_{swN} =- 0.40	K_{lN} =0.55	K_{wN} =0.54

Fig. 5 Strain sensitivity K of Cu-Mn-Al and Ni-Cr-Al sensors of 2 different types

Assuming that the temperature change is negligibly small during the measurement of pressure by the sensor of Cu-Mn-Ni alloy (2), equation[1] is reduced to equation [2]:

$$(\Delta R/R)_C = \alpha_{PC} \cdot \Delta P + \alpha_{\varepsilon C} \cdot \Delta\varepsilon \qquad [2]$$

As discussed in 3.2, the alloy Ni-Cr-Al (3) shows very small value of α_T and α_P. Therefore, for the sensor of this alloy (3), equation [1] is reduced to equation [3]:

$$(\Delta R/R)_N = \alpha_{\varepsilon N} \cdot \Delta\varepsilon \qquad [3]$$

Accordingly, introducing the ratio of these strain sensitivities by $a = K_C / K_N$, the corrected pressure ΔP is obtained from equation [2] as follows:

$$\Delta P = \{(\Delta R/R)_C - K_C \cdot \Delta\varepsilon\}/\alpha_{PC}$$
$$= \{(\Delta R/R)_C - a \cdot K_N \cdot \Delta\varepsilon\}/\alpha_{PC}$$

[4]

Equation [4] and $(\Delta R/R)_N = K_N \cdot \Delta\varepsilon$ yield:

$$\Delta P = \{(\Delta R/R)_C - a \cdot (\Delta R/R)_N\}/\alpha_{PC} \qquad [5]$$

The first term on the right hand side of equation [5] represents the pressure measured directly by Cu-Mn-Ni pressure sensor, while the second term the correction given by the Ni-Cr-Al strain sensor.

Putting the measured $(\Delta R/R)_C$ and $(\Delta R/R)_N$ into equations [5], the corrected pressure ΔP is obtained. However, using the presssure sensors of the straight-line type, equation [5] can not be applied for strain correction, because the ratio, a, of strain sensitivities depends on the direction. As shown in Table 2, a is 0.4 or 3 in x- or y- direction, respectively. On the other hand, the twin-arc type sensor has the uniqueness that a is the nearly independent of direction. As shown in Table 2, a is -0.55 and -0.56 in x- and y- direction, respectively. Therefore, the twin-arc type sensor enables the correction by equation [5].

Table 2 Strain sensitivity ratio a of sensors of 2 different types

Sensor type		Straight-line	Twin-arc
Strain sensitivity ratio : a	(X)	a (Longitudinal) $= \frac{K_{slC}}{K_{slN}} \approx 0.4$	a (Longitudinal) $= \frac{K_{lC}}{K_{lN}} \approx -0.55$
	(Y)	a (Lateral) $= \frac{K_{swC}}{K_{swN}} \approx 3$	a (Lateral) $= \frac{K_{wC}}{K_{wN}} \approx -0.56$
Remarks		a (Longitudinal) $\neq a$ (Lateral)	a (Longitudinal) $\approx a$ (Lateral)

4. DEPENDENCE OF PRESSURE SENSITIVITY UPON TEMPERATURE AT HIGH TEMPERATURE

Because of oxidation of the sensor, the pressure sensitivity α_P of the pressure sensor made of Cu-Mn-Ni depends more upon temperature in a higher temperature range, when it is coated with oxide films such as Al_2O_3 or SiO_2 as shown in Figure 1(c). In order to prevent the oxidation and so to reduce this temperature dependence, the Cu-Mn-Ni alloy was coated with AlN films, instead of oxide film as shown in Figure 6. Figure 7 depicts the dependence in both cases. The ordinate δ is given by equation [6], and is the percentage change of α_P to the value at 20°C. Where the sensor (1) coated with oxide film Al_2O_3 shows a large change reaching 23% in the range above 90°C, the sensor (2) coated with AlN only 2.5% and is, therefore, suitable for pressure measurement at higher temperature.

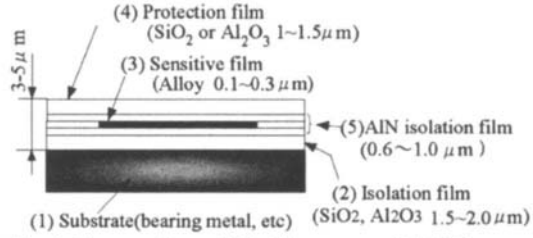
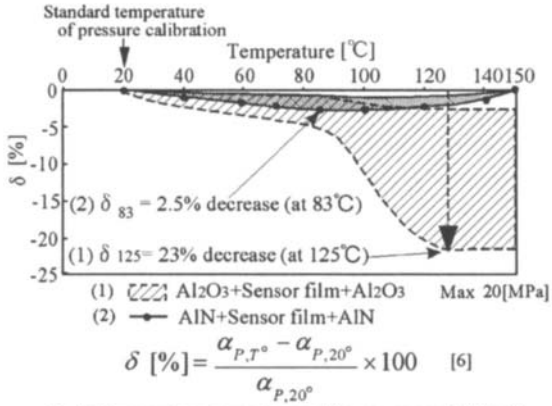


Fig. 6 Cross section of thin-film sensor with AlN



$$\delta\ [\%] = \frac{\alpha_{P,T^\circ} - \alpha_{P,20^\circ}}{\alpha_{P,20^\circ}} \times 100 \qquad [6]$$

Fig.7 Temperature dependence of pressure sensitivity α_P

5. MEASUREMENT IN MAIN BEARING

5.1 Test engine and measurement environment

A water-cooled four-cylinder turbo-charged diesel

engine of 5307 cm^3 piston displacement volume was used for oil-film pressure measurement in the main bearing. As shown in Figure 8(a), the thin-film sensors were deposited on the bottom axial line of main bearing surface at 5 different points with equal distance 4mm. 7 films of J-shape extending from each sensing parts to the bearing ends represent the 0.2μm thick lead film made of aluminum. Figure 8(b) shows a sectional view of the thin-film sensor deposited on the bearing. The basic structure is the same as that of Fig.1(c). In addition an 0.3 μm-thick intermediate film (2) of chromium was deposited between the aluminum alloy (1), the material of the bearing metal, and the SiO_2 isolation film (3) to improve the sensor adhesivity.

In order to investigate the effects of the bearing metal deformation on the measurement of oil-film pressure, we deposited the Ni-Cr-Al strain sensor in the same positions as the pressure sensors in Figure 8(a). Figure 9 shows the locations of the installed thin-film sensors. The measurement was made on the fourth (from the front side of the engine) main bearing.

5.2 Measurement result

Figure 10 shows an example of oil-film pressure measured in the engine main bearing. The measurement location was in the bottom center of the bearing. Experiments were carried out under no load and half load at 1500[rpm], with supply oil temperature maintained at 75[°C]. Oil feed pressure was about 0.5[MPa]. The curve (1) represents the measured strain using a thin-film strain sensor of Ni-Cr-Al alloy described in chapter 3.2, 3.3 and 3.4. Using this measured strain and equation [5], the correction curve (2) was obtained. The curve (3) represents the pressure measured directly by the Cu-Mn-Ni pressure sensor, and the curve (4) the corrected pressure given as the sum of (2) and (3) according to equation [5]. The curve (5) represents the bearing load calculated by a simple method, and (6) the gas pressures in the second, third and fourth cylinders. For comparison, the corrected oil-film pressure (4)' under no load is added to Figure 10.

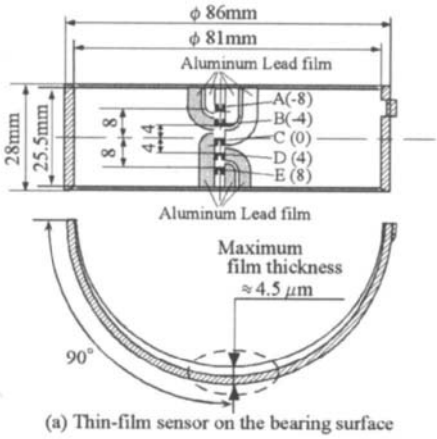
For the directly measured pressure curve (3), the output pressure level at crank angles ±360 [deg], at which no bearing load should theoretically act on the fourth main bearing, was defined as the zero level. The oil-film pressure (4) increases with the increase of the cylinder pressure (6) and the main bearing load (5). As corrected oil-film pressure, maximum value of 45MPa was obtained under half load, and 36MPa under no load.

Figure 11(a) shows the variation of the axial distribution of oil-film pressure at crank angles from -33 to 14 degree when the #4 cylinder combusts. As parameters for each curve, specific bearing load are indicated at crank angle in brackets. Figure 11(b) shows the variation of oil-film pressure when the oil hole (7mm in diameter) of the crank journal passes over the sensor at crank angles from -107 to -100.

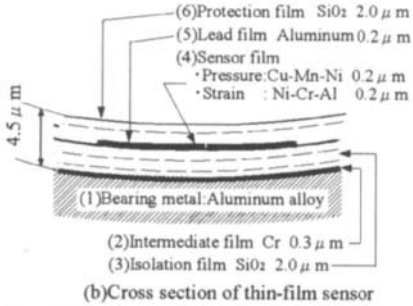
6. MEASUREMENT IN PISTON PIN BOSS

6.1 Thin-film sensor for piston-pin

Figure 12(a) shows the thin-film sensor, the lead films and the terminals fabricated on the piston pin. The axial location of sensing part, measured from the R-edge as shown in Figure 13(a), was varied from -0.6mm to 14.6mm by changing the length L of the lead part. The sensitive part used in this experiment was twin-arc type shown in Figure 1(b). Figure 12(b) shows the sectional view of the thin-film sensor on the piston pin. The basic film structure is the same as that of Figure 1(c). In addition an Al2O3 stepfiller film (5) was adopted to balance the thickness of the sensor alloy film (3). To investigate the effects of the piston pin deformation

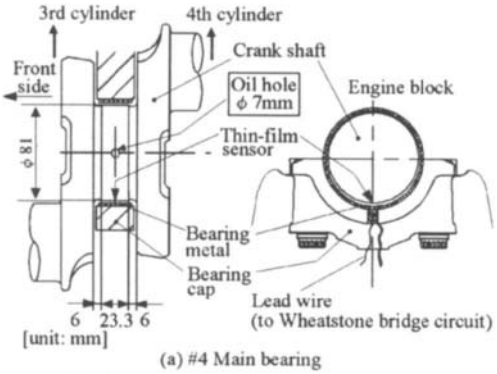


(a) Thin-film sensor on the bearing surface

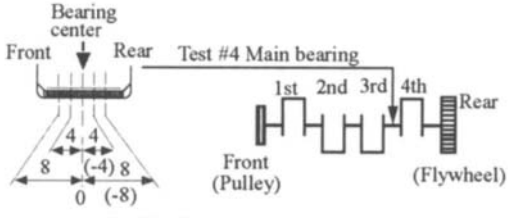


(b)Cross section of thin-film sensor

Fig. 8 Structure of thin-film pressure sensor on engine main bearing (Total 5 sensors)



(a) #4 Main bearing



(b) Test bearing & sensor position

Fig.9 Test main bearing in diesel engine

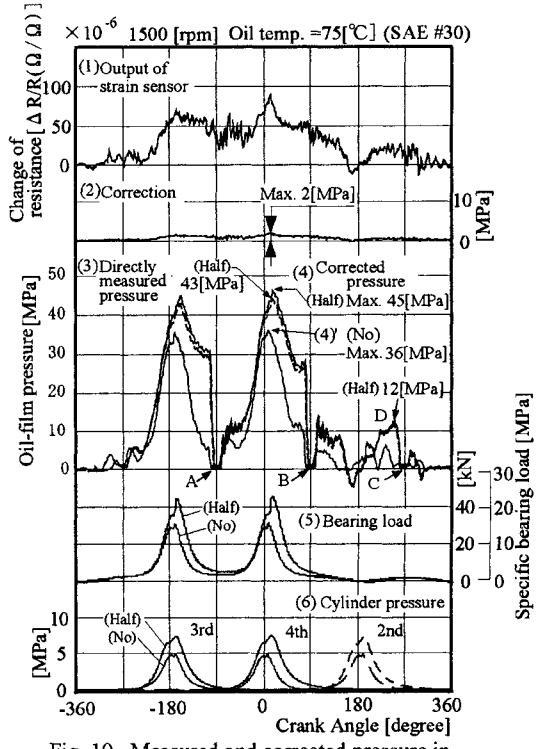
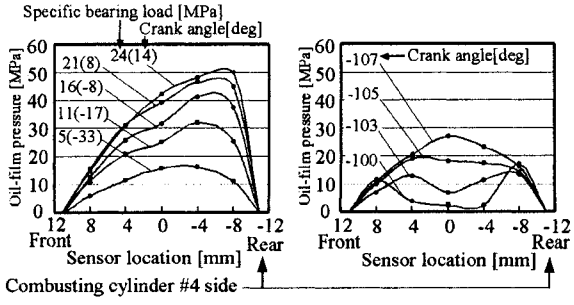


Fig. 10 Measured and corrected pressure in #4 main bearing



(a) Variation of oil-film pressure distribution at #4 cylinder combustion

(b) Effect of passing oil hole

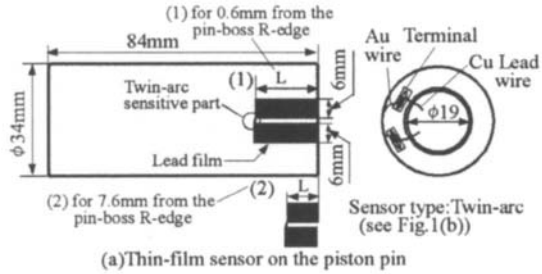
Fig.11 Axial oil-film pressure distribution in #4 main bearing

on the measurement of oil-film pressure, we deposited the Ni-Cr-Al strain sensors in the same positions as the pressure sensor.

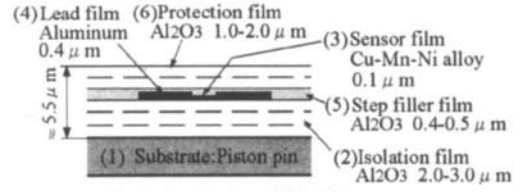
6.2 Pin-boss shapes of test pistons

Figure 13(b) shows three shapes of the top line of the piston pin-boss inner surface used in the test. The shape (1) is a straight-type piston pin-boss. The axial length of the pin-boss at the top, measured between the end of the snap ring groove and the R-

edge inside the piston is 23.6mm. The basic shapes of (2) and (3) are the same as that of the straight-type pin-boss (1). However, they are tapered by 0.1° beginning at the points distant 20 or 15mm from the snap ring groove. As shown in Table 3, the clearance Y at the inner edge increases in the order of (1), (2) and (3). The diametral clearance between the pin and the inner cylindrical pin-boss surface was 15-20μm in all types. In addition, the reference point 0 of the measurement location z for the pin-boss (2) and (3) was adjusted to the same position as the R-edge in a straight-type as shown in Figure 13(b).

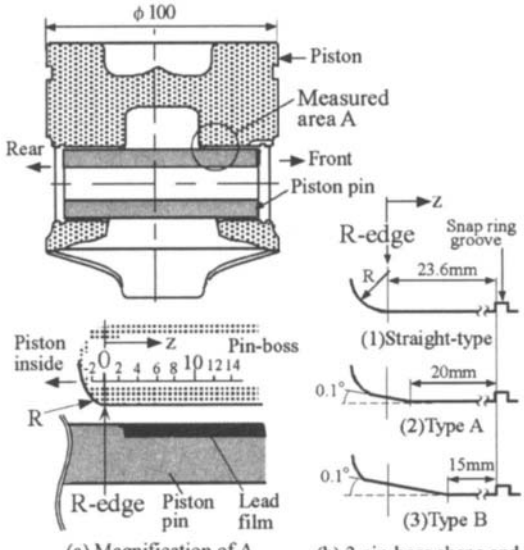


(a)Thin-film sensor on the piston pin



(b) Sectional view of thin-film sensor

Fig. 12 Thin-film sensor for measuring oil film pressure in piston pin-boss

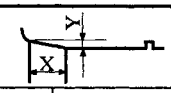


(a) Magnification of A

(b) 3 pin-boss shape and R-edge location

Figure 13 Measurement point and pin-boss shape of test piston

Table 3 Amount of clearance in 3 type pin-boss

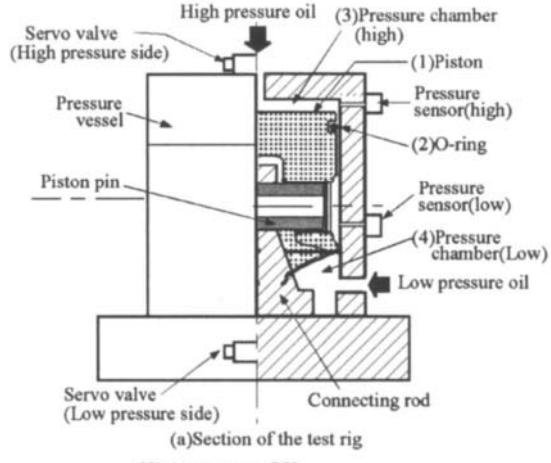


X: mesured data
Y:calculated data

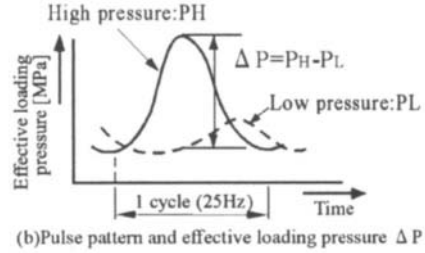
(1)Straight	(2)Type A	(3)Type B
X: 0 mm Y: 0 mm	X : 3.8mm Y : 6.6 μm	X : 8.9mm Y : 15.5 μm

6.3□ Test rig

A piston pin boss fatigue tester shown in Fig.14 was used to measure the oil-film pressure in the piston pin boss. As illustrated in Fig.14(a), the rig is divided by an O-ring (2) installed in the piston (1) which devides the higher-pressure side of an oil chamber (3) and the lower-pressure side (4). The higher-pressure side oil chamber was subjected to pressure variation simulating that in engine cylinder, while the lower-pressure side oil chamber was subjected to pressure variation simulating the inertia force. Therefore, the pressure which acts on the piston pin-boss part is considered to be $\Delta P=P_H-P_L$, as shown in Figure 14(b). These pressures were applied at a frequency of 25 Hz, corresponding to an engine rotation of 1500 rpm.



(a)Section of the test rig



(b)Pulse pattern and effective loading pressure Δ P

Fig.14 Test rig for measuring pin-boss oil film pressure

6.4 Measured pressure and its correction

Figure 15 shows an example of the result of measurement in the straight-type pin-boss shown in Figure 13(b)(1) and Table 3(1), using the sensors with the sensitive part located on the piston pin 1.6mm apart from the R-edge shown in Figure 13(a) and (b). ISO VG68 oil was used as a lubricant at temperature 50°C, and the maximum of the effective loading pressure ΔP on the piston amounted to 12MPa.

In Figure 15, the curve (1) represents the strain measured by the thin-film strain sensor of Ni-Cr-Al. This strain causes the maximum error of 7MPa in the measured oil-film pressure according to the second term of equation [5]. The curve (2) with its maximum 155MPa represents the oil-film pressure measured directly by the pressure sensor of Cu-Mn-Ni according to the first term of equation [5]. The curve (3) with its maximum 162MPa is the corrected pressure and is the sum of the curves (1) and (2).

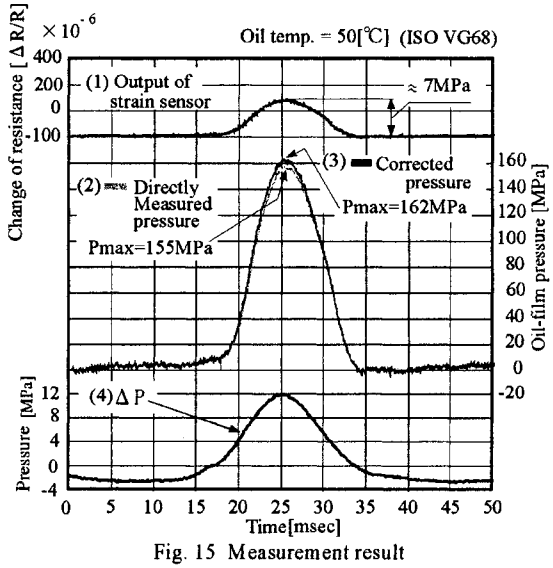


Fig. 15 Measurement result

6.5 Comparison of oil-film pressure distribution between 3 different pin-boss shapes.

Figure 16 shows the axial distribution of the corrected oil-film pressure along the top line of the inner surface of the piston pin-boss for 2 cases of loading pressure ΔPmax=6MPa and 12MPa. In case of ΔPmax=6MPa, the maximum oil-film pressure of each pin-boss type shown in Figure 13(b) and Table 3 became 80MPa at 1.6mm in straight type, 81MPa at 3.6-5.6mm in type A and 79MPa at 5.6mm in type B, respectively. In case of ΔPmax =12MPa, the oil-film pressure in the straight-line type piston showed the highest value, approximately 195MPa, at the location 0.6mm away from the R-edge. These values were 155MPa at 1.6-3.6mm in type A and 130MPa at 3.6-5.6mm in type B.

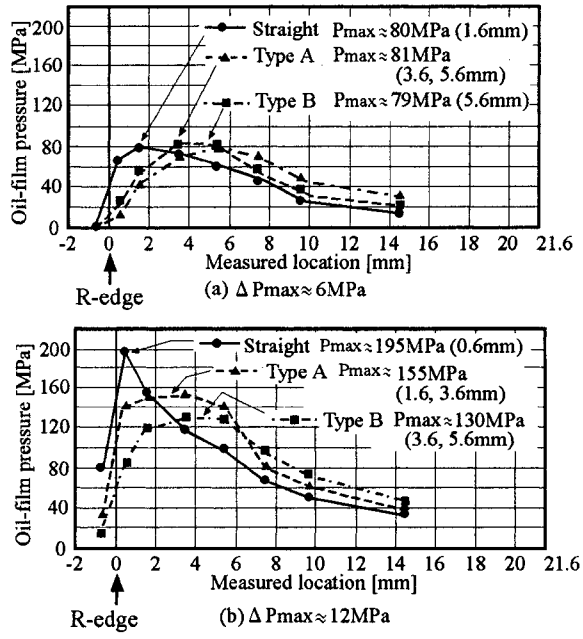


Fig.16 Measured oil-film pressure distribution in the piston pin-boss

7. DISCUSSION

7.1 Main bearing

1. As shown by curve(2) in Figure 10, the maximum of the measurement error due to bearing deformation was approximately 2MPa at most or less than 5%. This means, the twin-arc type pressure sensor, as shown in Figure 1(b) and Figure 5, can be used as a pressure sensor without correction, though it may depend on the measurement accuracy required.
2. The oil-film pressure falls instantaneously to approximately 1 MPa or lower at the crank angles A, B, and C in Figure 10. This is because the oil hole (7 mm in diameter) of the crank journal shown in Figure 9(a) passes over the sensing parts of the sensor. The result shows speed of response of the sensor.

This is also seen in Figure 11(b), where the oil-film pressure decreases as the oil hole in the crank

journal approaches the pressure sensor situated at the bearing bottom and becomes approximately 0~2MPa at a crank angle of 100 deg where the oil hole meets with the sensor.

3. The oil-film pressure distribution shown in Figure11(a) is almost symmetrical when the cylinder gas pressure is 5MPa or lower. However, when the gas pressure exceeds 5MPa, the oil-film pressure distribution becomes unsymmetrical with the peak shifted to the combustion cylinder side.

4. In Figure 10, the oil-film pressure (4) builds up and reaches approximately 12MPa at crank angle D, while the calculated bearing load (5) is almost 0MPa. This peak oil-film pressure is due to the influence of second cylinder combustion pressure, which was not considered in the calculation of bearing load.

7.2 Piston pin-boss

1. It is seen from the curve (1) in Figure 15, that the maximum error of pressure due to piston pin deformation was 7MPa, measured under the loading pressure ΔPmax=12MPa by sensor position 1.6mm away from the R-edge. This error 7MPa was the highest among the errors obtained at all other measurement positions. It corresponds to 4.3% error relative to corrected pressure 162MPa.

2. As shown in Figure 16, the straight-type boss showed the highest oil-film pressure, followed by the type A and type B on the inner side of the boss. This order of magnitude reverses in the center and on the outside of the boss.

3. With increasing load ΔPmax, the peak value of the oil-film pressure increased and the axial position of the peak pressure approached the inside of the boss.

8. CONCLUSION

In order to measure the oil-film pressure on the sliding surfaces in internal combustion engines, we have developed a thin-film sensor and measured the distribution of oil-film pressure in the main bearing and piston pin-boss. The results obtained are summarized as follows:

(1)The best composition of Cu-Mn-Ni alloy for measuring pressure was 87.4Cu-9.6Mn-3.0Ni [mass%]. It showed the smallest temperature sensitivity α_T =(-20~8)×10^{-6} [(Ω/Ω)/°C] among the tested alloy compositions. On the other hand, pressure sensitivity of Cu-Mn-Ni alloys fell in the range of (18~22)×10^{-6}[(Ω/Ω)/MPa], regardless of alloy compositions. Thus the 87.4Cu-9.6Mn-3.0Ni alloy showed the smallest error in measured pressure due to temperature change: α_T / α_P =(-1.1~0.4) [MPa/°C].

(2) The longitudinal strain sensitivity K_{sl} of all tested Cu-Mn-Ni alloys was between 0.50 and 0.70, while its lateral strain sensitivity K_{sw} was between -1.40 and -1.10.

(3) We have devised a twin-arc type sensor, of which the longitudinal and lateral strain sensitivity, K_l and K_w are equal. This type gave the results K_l=-0.3 and K_w=-0.3 for Cu-Mn-Ni alloy and K_l=0.55 and K_w=0.54 for Ni-Cr-Al alloy. Thus the strain sensitivity ratio a of both materials could be made equal in longitudinal and lateral directions, so that the error in measured pressure due to strain could be corrected properly.

(4)With the introduction of the twin-arc type sensor, the absolute value of strain sensitivity could also be reduced nearly to 50% or 75% of its original value.

(5) The Cu-Mn-Ni alloy used for pressure measurement has yielded a large temperatrure dependence of pressure sensitivity when it was coated with an oxide film. However, this dependence could be reduced drastically by coating the alloy with a nitride film AlN.

(6) Using the above-mentioned sensor material and shape, a single-point and a multi-point pressure sensor were developed and applied successfully to measure the oil-film pressure in a main bearing of a firing diesel engine. As expected, the measured oil-film pressure fell instantaneously below 1Ma when the oil hole in the crank shaft passed over the pressure sensor. The position of the peak oil-film pressure shifted axially towards the neighboring combusting cylinder.

(7) In a dynamic test rig, the change in axial distribution of the oil-film pressure in the piston pin-boss could be studied for varied clearance geometries in the pin-boss.

REFERENCES

[1] Bridgeman, P, W., *The Physics of high Pressure*, G.Bell and Sons, Ltd., London(1931).
[2]Kannel, J,W., Bell, J.C. and Allen, C.M.,"Methods for Determining Pressure Distributions in Lubricated Rolling Contact," ASLE Trans.,8, pp 250-270(1965).
[3]Köhler, A."Die Entwicklung von Aufgedampften Meßwertaufnehmern und deren Anwendung zur Druck- und Temperaturmessung in geschmierten Wälz- und Gleitkontakten," Dissertation, T. H. Aachen(1981).
[4]Bauer, P., "Theoretische und experimentelle Untersuchungen zu tribologisch relevanten Betriebsgößen an verkanteten Zylinderrollenlagern," *Dissertation*, Technische Hochschule Aachen(1987).
[5]Kirschner, J., "Tribologische und kinematische Untersuchungen an einem Nocken-Rollenstößel-Trieb," *Dissertation*, Universität Karlsruhe(1988).
[6]Staiger, J., "Untersuchungen über den Einfluß keramischer Dünnschichten auf die Reibung in elastohydrodynamischen Kontakten," *Dissertation*, Universität Karlsruhe(1991).
[7]Mihara, Y., Kajiwara, M., Fukamatsu, T., Someya, T., "Study on the measurement of oil-film pressure of engine main bearing by thin-film sensor - The influence of bearing deformation on pressure sensor output under engine operation -," *JSAE Review*, Vol.17, No.3, pp281-286 (1996).
[8]Mihara, Y., Hayashi, T., Nakamura, M., Someya, T., "Development of measuring method for oil-film pressure of engine main bearing by thin-film sensor," *JSAE Review*, Vol.16, No.2, pp125-130 (1995).
[9] Someya, T., Mihara, Y., Gaul, L., Mahrenholtz, O., "Entwicklung eines Dünnschichtsessors zur Messung des Schmirfilmdrucks im Motorgleitlager," *Motortechnische Zeitschrift*, 58, Jahrgang, Nr. 2, pp110-116(1997).
[10]Mihara, Y., Someya, T., "Measurement of oil-film pressure in Engine Bearings Using a Thin-Film Sensor", STLE Trans.,Vol.45, 1, pp11-20(2002).

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Optimisation of static behaviour and stability level of a multilobe journal gas bearing

G. Grau, I. Iordanoff, B. Bou Saïd, Y. Berthier
Laboratoire de Mécanique des Contacts, Bâtiment Jean d'Alembert, 20, av. Albert Einstein
69621 Villeurbanne Cedex, France

This work carries out a study of the static and dynamic behaviour of a three lobe offset journal gas bearing. The influence of the internal geometry on both load carrying capacity and stability behaviour is investigated. The load carrying capacity is characterised from the minimum film thickness within the bearing when submitted to an imposed load (magnitude and direction are fixed). A numerical non linear dynamic model is presented. The specific approach developed allows to point out and quantify the inner geometry influence on both stability limit and minimum film thickness in the bearing. The typical result is that a bearing able to support a high load will not necessary fit a given application because of its low level of stability. Using this original approach, it is possible to design a profile with optimised static and dynamic behaviour in connection with a specific application.

1 Introduction

Gas lubricated hydrodynamic bearings have been of increasing interest for high speed turbomachinery. As the rotational speed tends to increase, they offer a good alternative to conventional rolling elements. Moreover, using ambient gas (air for example) as a lubricant leads to put offside all the lubrication system. Load capacities of air lubricated bearings are lower than conventional bearings. Therefore their field of application are limited to high speed, low load machines. Due to the high rotational speed, the level of stability is a characteristic of real importance.

Experiments with journal bearing have put forward a higher load capacity for cylindrical bearings, and, on the other hand, a better whirl resistance for lobed bearings [4,5] (especially the full offset bearing).

Therefore, this study is focused on a three lobe journal gas bearing. This analysis is carried out using a non linear dynamic model. This analysis enables to understand the static behaviour of the bearing when submitted to a given load (fixed magnitude and direction) but also to determine if the operating condition is stable or not. The influence of the bearing profile parameters on both static performances (film thickness) and stability will be discussed.

2 Description of the model

2.1 Assumptions and equations

The geometry of the bearing is reported on Fig. 1.

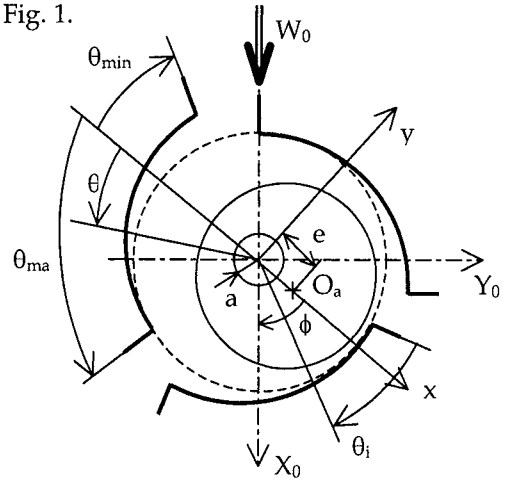


Fig. 1 Geometry of a three lobe journal bearing

The film thickness distribution is:

$h(\theta,z)=C_0(1+\varepsilon cos(\theta-\theta_{min}))+a(1+cos(\pi-\theta_i))$

The normalised film thickness can be expressed as:

$H(\theta,Z)=(1+\varepsilon cos(\theta-\theta_{min}))+m(1+cos(\pi-\theta_i))$

With dimensionless variables,

$Z=(z/R)$, $H=(h/C_0)$, $m=(a/C_0)$

When the external load acting on the shaft is not constant in direction or magnitude, the shaft center describes a trajectory within the housing.

The analysis of this trajectory requires the resolution of the dynamics equations that needs the evaluation of the hydrodynamic effect in the film. The main difficulty comes from the dependence of the hydrodynamic force on both the position of the shaft within the housing and the velocity of the shaft centre. In most industrial applications, the external force is the weight of the shaft, with or without superimposed dynamic forces from mass unbalance for example.

The rotor assembly is modelled as a rigid shaft, symmetrically supported by two identical bearings. Each bearing supports a mass M, a static load W_0 and a synchronous excitation due to an unbalance, characterised by its eccentricity e_b, Fig. 2.

The application of the basic principle of dynamics gives:

$$\begin{cases} M\ddot{x} = W_0 + Me_b\omega^2\cos(\omega.t) + F_x \\ M\ddot{y} = Me_b\omega^2\sin(\omega.t) + F_y \end{cases} \quad (1)$$

where x and y denote the shaft centre co-ordinates

$\overrightarrow{OO_a} = x\vec{X} + y\vec{Y}$

and F_x, F_y are the components of the hydrodynamic force, calculated from the pressure distribution. The pressure is obtained from the solution of the time-dependent Reynolds equation [2]. In dimensionless variables it writes:

$$\frac{\partial}{\partial\theta}\left(PH^3\frac{\partial P}{\partial\theta}\right)+\frac{\partial}{\partial Z}\left(PH^3\frac{\partial P}{\partial Z}\right)=\Lambda\left(\frac{\partial(PH)}{\partial\theta}\right)+2\Lambda\left(\frac{\partial(PH)}{\partial T}\right) \quad (2)$$

where $P=(p/p_a)$, $T=t.\omega$, and Λ is the bearing number

$$\Lambda=\frac{6\mu\omega}{p_a}\left(\frac{R}{C_0}\right)^2$$

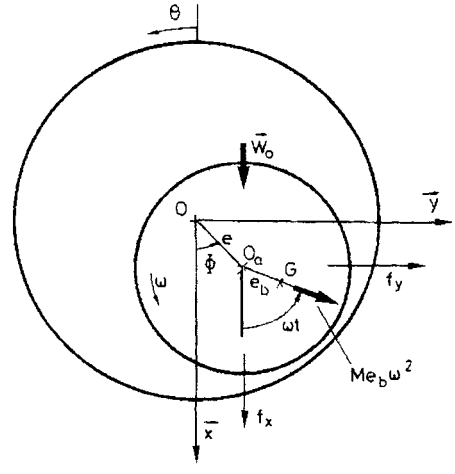


Fig. 2 Nomenclature for dynamic analysis

The basic principles of dynamics, [3], is written in dimensionless form:

$$\begin{cases} \overline{M}\,\ddot{\overline{X}} = \overline{W}_0 + \overline{M}\varepsilon_b\cos(T) + \overline{F}_x \\ \overline{M}\,\ddot{\overline{Y}} = \overline{M}\varepsilon_b\sin(T) + \overline{F}_y \end{cases} \quad (3)$$

where:

$$\overline{F},\overline{W}_0 = \frac{F,W_0}{p_a.R^2},\ \overline{X},\overline{Y} = \frac{x,y}{C_0},$$

$$\varepsilon_b = \frac{e_b}{C_0},\ \overline{M} = \frac{MC_0\omega^2}{p_a R^2}$$

The computation of the shaft centre trajectory requires the solution of the Reynolds equation (2), and the system of equations (3).

2.2 Method of solution

The resolution scheme used to compute the trajectory of the shaft centre is as follows:

- at time t, the co-ordinates $X(t)$, $Y(t)$, the speed components $\dot{X}(t)$, $\dot{Y}(t)$ and the components of the acceleration $\ddot{X}(t)$, $\ddot{Y}(t)$ of the shaft centre, the components of the hydrodynamic force $F_x(t)$ and $F_y(t)$ are known,
- the shaft position at time $t+\Delta t$ is calculated using the relation:

$$X(t+\Delta t) = X(t) + \Delta t\,\dot{X}(t) + \frac{\Delta t^2}{2}\ddot{X}(t)$$

- this new position allows the calculation, with (2), of the new components of the hydrodynamic force $F_x(t+\Delta t)$ and $F_y(t+\Delta t)$. Using (3) one obtains $\ddot{X}(t+\Delta t)$ and $\ddot{Y}(t+\Delta t)$. Therefore, the calculation of the speed components at time $t+\Delta t$ is obtained from:

$$\dot{X}(t+\Delta t) = \dot{X}(t) + \Delta t\left(\frac{\ddot{X}(t) + \ddot{X}(t+\Delta t)}{2}\right)$$

This algorithm is known as "Verlet scheme".

For the numerical solution of the Reynolds equation (2) a similar method is used as the one presented in [1]. The finite differences method is applied for the discretization of space (inner bearing surface) and time (shaft position).

Indeed, the time – dependant Reynolds equation features the term $\frac{\partial(PH)}{\partial T}$ expressing the variation of the pressure with the time. This implies a certain difficulty for the solution of equation (2).

The boundary conditions for the pressure distribution are:

at $\theta=\theta_{min}$ and $\theta=\theta_{max}$, $P=(p/p_a)=1$

at $Z=\pm(L/D)$, $P=(p/p_a)=1$

An implicit method, (4), is chosen for the calculation of P at time $t+\Delta t$ to reduce computation time (about 10 times faster). This is principally due to the large number of steps (to ensure numerical stability) necessary when using an explicit method.

$$\frac{\partial}{\partial\theta}\left(PH^3\frac{\partial P}{\partial\theta}\right)^{t+\Delta t} + \frac{\partial}{\partial Z}\left(PH^3\frac{\partial P}{\partial Z}\right)^{t+\Delta t} = \Lambda\left(\frac{\partial(PH)}{\partial\theta}\right)^{t+\Delta t} + 2\Lambda\left(H^{t+dt}\frac{\left(P^{t+\Delta t}-P^t\right)}{\Delta T} + P^{t+dt}\frac{\left(H^{t+\Delta t}-H^t\right)}{\Delta T}\right) \quad (4)$$

The convergence is achieved once the trajectory of the shaft becomes stable (transient motion completed).

3 Analysis

3.1 Stability

The model developed can be used to determine the trajectory of the shaft centre within the housing. Such computation requires the operating condition to be stable. It is possible to evaluate the stability of the rotor assembly when it is only submitted to its proper weight, i. e, with no external excitation.

Three situations can be observed; the geometry of the bearing is described in Table 1 and the operating condition is M=1 kg, W_0 equals to the shaft weight.

- *stable* : the shaft centre moves toward its equilibrium position, Fig. 3,
- *limit* : the trajectory of the shaft centre is an orbit of a given amplitude, Fig. 4,
- *unstable* : a divergent orbit is obtained, Fig. 5.

Table 1 Bearing geometry

Number of pads	3
Pad extent	119°
L=D	50 mm
C_0	20 µm
a	30 µm

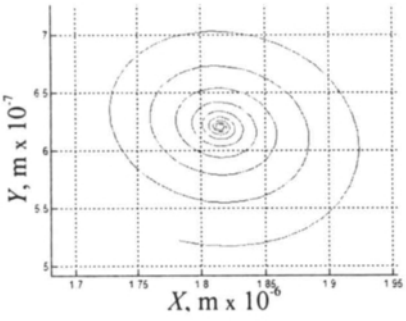


Fig. 3 Converging orbit – 30,000 rpm

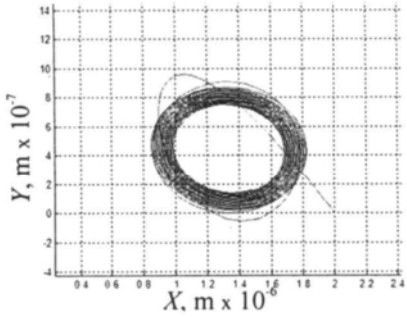


Fig. 4 Limit orbit – 45,000 rpm

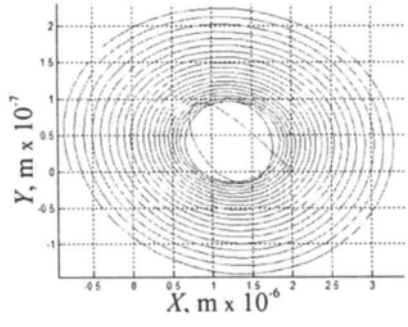


Fig. 5 Diverging orbit – 49,000 rpm

3.2 Influence of the geometry

The inner geometry of a bearing is well known to be an important parameter for the definition of the wedge and therefore the pressure distribution.

The clearances of the bearing (radial clearance and geometrical preload) affect both load capacity and stability as they depend on the pressure distribution.

In order to understand how much the inner geometry of the bearing changes its behaviour, dynamic simulations have been conducted.

The bearing tested is described in Table 1. The operating conditions, i. e, rotational speed and rotor mass are described in Table 2.

Table 2 Operating conditions

Rotational speed	30,000 rpm
Rotor mass	5 kg

The influence of the radial clearance is presented on Fig. 6. The minimum film thickness in the bearing is plotted against the geometrical preload for different values of the radial clearance. The curves presented are limited for decreasing values of the geometrical preload because the operating condition becomes unstable.

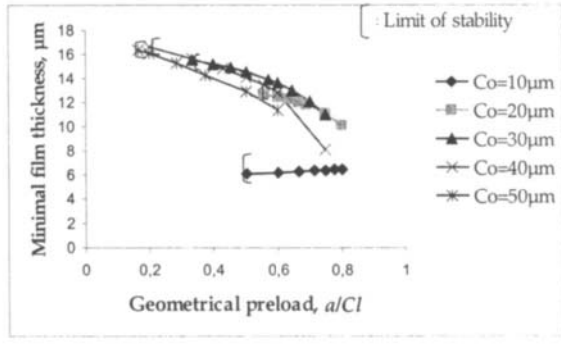


Fig. 6 Hmin Vs geometrical preload - Influence of C_0

One can observe that when the radial clearance is small (10 and 20µm) the stability domain (with respect to the geometrical preload) is quite narrow and the minimal film thickness in the bearing is low. Increasing the radial clearance results in enlarging the range of *possible* geometrical preload, i.e, better stability level. But that doesn't necessary lead to higher film thickness.

Hence, the clearance can be chosen to ensure a good compromise between the stability level and a high film thickness.

For a given radial clearance, the influence of the geometrical preload on the stability of the bearing operating at 30,000 rpm is investigated for different values of the rotor mass.

The minimum film thickness in the bearing is plotted against the geometrical preload, Fig. 7. As in Fig. 6, the curves presented are limited for the small values of the geometrical preload.

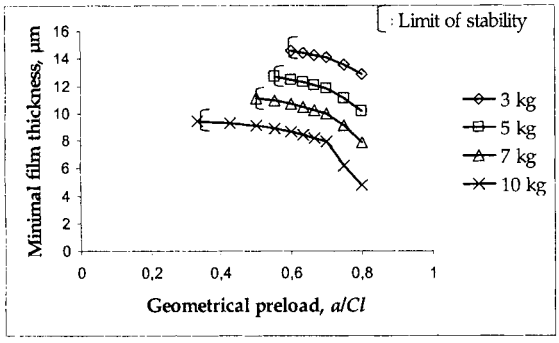


Fig. 7 Hmin Vs geometrical preload – Influence of M

These curves are very interesting for different reasons. They clearly put forward that for a given rotor mass and rotational speed, all the geometrical preload will not fit. When increasing the mass of the shaft (and therefore its weight) the rotor assembly is much more stable regarding to the preload. In other words, the heavier the rotor the smaller the geometrical preload can be chosen.

But this choice has also to be conducted to ensure high film thickness. For example, if the rotor has a mass of 10kg, the operating condition will be stable if the geometrical preload is higher than about 0.35. The range is reasonably limited to 0.7 because for higher values the minimum film thickness decreases substantially.

The originality of the present approach compared to the well known stability map is the possibility to choose the inner geometry parameters with regard to both static (minimum film thickness) and stability behaviour with only one curve. This is interesting and useful for the design of specific bearing.

4 Conclusion

The static behaviour and the stability of a three lobe offset journal gas bearing have been investigated with a non linear dynamic model. The analysis was focused on the effect of the internal geometry on the bearing performances.

The typical result is that a bearing able to support a load with high film thickness will not be chosen for a given application because its level of stability is not appropriated. The originality of the approach presented is based on the elaboration of curves allowing to choose the inner geometry with regard to both static and stability criteria.

The model developed is intended to be adapted to a compliant profile and the same kind of study will be carried out.

Nomenclature

a	geometrical preload, $a=C_l-C_0$
C_0	radial clearance
C_l	radial clearance on the lobe
D	diameter of journal bearing
e	eccentricity
e_b	unbalance eccentricity
L	bearing lenght
m	a/C_0
M	rotor mass
p_a	ambient pressure
R	radius of journal bearing
t	time
T	$T=\omega.t$
U	linear velocity
W_0	static load on the bearing
x, y	shaft centre co-ordinates
$\dot{x}, \dot{y}$	shaft centre velocities
$\ddot{x}, \ddot{y}$	shaft centre accelerations
z	axial co-ordinate
Λ	bearing number
ε	e/C_0
ε_b	e_b/C_0
μ	gas viscosity
θ	angular general co-ordinate
θ_i	angular co-ordinate on a pad
θ_{min}	start of the pad
θ_{max}	end of the pad

References

1. Heshmat, H., Walowit, J.A., Pinkus, O., "Analysis of Gas-Lubricated Foil Journal Bearings", Journal of Lubrication Technology, October 1983, vol 105, pages 647-655.
2. Arakere, N., Nelson, H. D., "An Interior Collocation Method for Static and Dynamic Analysis of Finite Length Gas Journal Bearings", Journal of Tribology, July 1988, vol 110, pages 456-461.
3. Abdul Wahed, M. N., Nicolas, D., Pascal, M. T., "Stability and Unbalance Response of Large Turbine Bearings", Journal of Lubrication Technology, January 1982, vol 104, pages 66-75.
4. Li, D. F., Choy, K. C., Allaire, P. E., "Stability and Transient Characteristics of Four Multilobe Journal Bearing Configurations", Journal of Lubrication Technology, July 1980, vol 102, pages 291-299.
5. Allaire, P. E., Li, D. F., Choy, K. C., "Transient Unbalance Response of Four Multilobe Journal ", Journal of Lubrication Technology, July 1980, vol 102, pages 300-306.

SESSION XI
BIO-TRIBOLOGY (2) - BIOLOGICAL SYSTEMS

Chairman : Dr Z M Jin

Paper XI (i) **Biotribological Approaches to the Lubrication of Engineering Systems**

Paper XI (ii) **Eyelid/Lens Contact During a 100 metres Sprint. A Tribological Approach**

Paper XI (iii) **Biological Lubrication of Hydrated Surface Layer in Small Intestine**

Paper XI (iv) **Tribology of a Snail (Terrestrial Gastropod)**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Biotribological Approaches to the Lubrication of Engineering Systems

Nicholas D. Spencer[a], Scott S. Perry[b], Seunghwan Lee[a], Markus Müller[a], Stéphanie Pasche[a], Susan M. De Paul[a], Marcus Textor[a], Xiaoping Yan[b] and Min Soo Lim[b]

[a]Laboratory for Surface Science and Technology, Department of Materials,
Swiss Federal Institute of Technology, ETH-Zürich, Switzerland

[b]Department of Chemistry, University of Houston, Texas 77204-5003, USA

Nature relies on aqueous lubrication for critical tribosystems, such as hip, knee, shoulder, ankle, and finger joints. Although water is clearly an attractive lubricant for economical, ecological, thermal management and other reasons, it suffers from a considerable disadvantage compared to oil, namely its low pressure-coefficient of viscosity, which has serious consequences for elastohydrodynamic lubrication. In nature, alternative lubrication mechanisms are involved, such as the pressure response of cartilage in articulated joints. In this paper we describe the use of the polyelectrolyte graft copolymer, poly(L-lysine)-*g*-poly(ethylene glycol) as a surface coating that could potentially mimic some of these natural lubrication mechanisms in technological applications. Both nanotribological (friction force microscopy) and macrotribological (pin-on-disk tribometry) measurements show the interesting tribological properties of this system, which resembles a grafted brush on the sliding surfaces. In particular, a substantial reduction of friction between a number of different sliding partners has been observed at low sliding velocities, suggesting that the polymer is functioning as an effective boundary lubricant.

1. INTRODUCTION

Water might seem at first sight to be a good candidate as a lubricant. It is non-toxic, environmentally friendly, and has excellent heat-transfer properties due to its relatively high specific heat capacity. On the other hand, it has several distinct disadvantages, such as its very low pressure-coefficient of viscosity [1]. The latter has serious consequences for elastohydro-dynamic lubrication by water and thus for the load-carrying capacity of water-lubricated tribosystems. Nature deals with this issue by using "smart", pressure-responsive cartilage surfaces as sliding partners. The mechanism of natural joint lubrication—still being the subject of intense debate—is remarkably effective; friction coefficients of as low as 0.003 have been measured in natural synovial joints [2].

Clearly, an analogous pressure-responsive aqueous lubrication mechanism might be achievable for engineering systems, through the immobilization of appropriately designed macromolecules onto the sliding surfaces. In this study, we have used the polyelectrolyte graft copolymer, poly(L-lysine)-*g*-poly(ethylene glycol) ("PLL-*g*-PEG") as a surface coating, since it is both strongly adsorbed to many oxide surfaces and coordinates large amounts of water. The molecule has already found application in several biomedical devices, in cases where protein adhesion needs to be inhibited [3].

The structure of PLL-*g*-PEG is depicted in Figure 1. It consists of a polylysine backbone with multiple side-chains consisting of polyethylene glycol, which have been attached to the backbone via the amine groups on some of the lysine units. Over a wide pH range, the remaining, unreacted amine groups develop a positive charge, which leads to the PLL-*g*-PEG

molecule's facile attachment to oxide surfaces bearing a negative change. At neutral pH, this includes many oxides (and, naturally, oxide-coated metals) of engineering interest, with the notable exception of Al_2O_3. The adsorption of PLL-*g*-PEG onto an oxide surface is shown schematically in Figure 2.

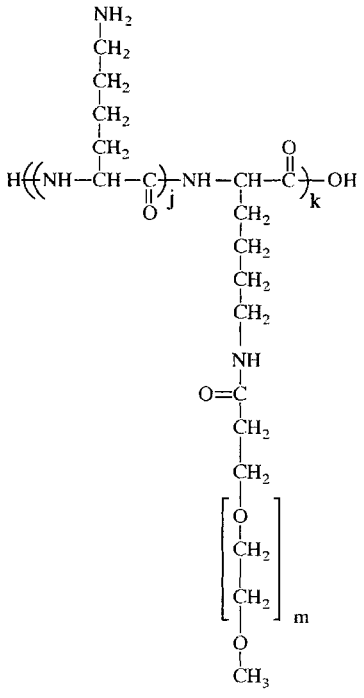


Figure 1. Structure of PLL-*g*-PEG

Previous studies involving the surface forces apparatus have shown that sliding surfaces covered with grafted polymer chains under "good" solvents, such as polystyrene under toluene, can display extremely low coefficients of friction [4,5]. To our knowledge, such brush-solvent systems have not been previously investigated in an aqueous context, nor have they been tested under macrotribological conditions.

2. EXPERIMENTAL

2.1. Materials

PLL(x)-*g*[y]-PEG(z)—a grafted co-polymer with a PLL backbone of molecular weight x kDa, a grafting ratio of (lysine-mer)/(PEG side chains) of y, and a PEG molecular weight of z kDa—was synthesized according to a previously described method [3]. In this study, the composition was PLL(20)-*g*[2.9]-PEG(2) for the tribometry investigation and PLL(20)-*g*[3.5]-PEG(2) for the friction force studies. Briefly, poly(L-lysine) hydro-bromide (MW 20 kDa, Sigma, USA) was dissolved in 50 mM sodium borate buffer solution and the solution filter sterilized (0.22 μm pore-size filter). In order to graft PEG onto PLL, the N-hydroxysuccinimidyl ester of methoxypoly-(ethylene glycol) propionic acid (MW 2 kDa, SPAPEG, Shearwater Polymer, Inc., USA) was added to the dissolved PLL. The reaction was allowed to proceed for 6 hr at room temperature, after which the reaction mixture was dialyzed (Specrea-Por, MW cutoff size 6-8 kDa, Spectrum, USA) for 24 hr against deionized water. The product was freeze-dried and stored at -20°C. Detailed analytical information for the product obtained via this method is available in previous publications [3].

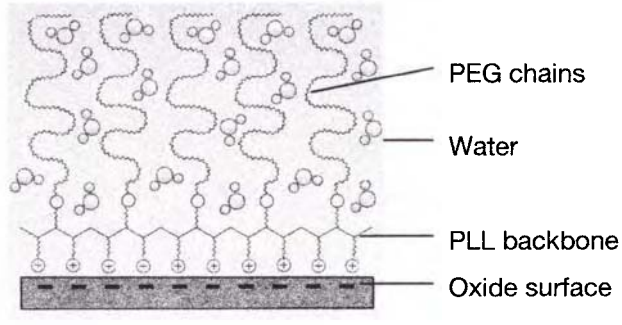


Figure 2. Schematic diagram of PLL-*g*-PEG adsorption onto a negatively charged surface

The PLL-*g*-PEG was dissolved in 10 mM HEPES (4-(2-hydroxyethyl)piperazine-1-ethane-sulfonic acid, adjusted to pH 7.4 with 1.0 M NaOH solution) at a concentration of 0.25 mg/ml and this solution was used for all pin-on-disk experiments. For the cases where the comparison of the PLL-*g*-PEG with the PLL(20) alone was examined, PLL(20) solution was prepared by dissolving the poly(L-lysine) hydrobromide in 10 mM HEPES at a

concentration of 0.045 mg/ml, in order to provide the same PLL concentration as in the PLL-*g*-PEG solution.

2.2. Methods

Pin-on-Disk Tribometry. A pin-on-disk tribometer (CSM, Neuchâtel, Switzerland) was used to characterize the friction and lubrication properties of aqueous PLL-*g*-PEG solutions. A fixed, spherical steel ball (AISI 52100, dia. 6 mm) was placed in contact with a flat glass disk (SuperFrost soda-*g*lass, Menzel-Gläser, Braunschweig, Germany), which could be rotated at constant angular velocity. The load was controlled by placing dead weights on top of the ball holder (0.5 to 5.0 N) and the friction forces measured by means of a strain gauge. To investigate the lubricating properties of the polymer, both load and velocity dependence of the frictional properties were explored. For each friction measurement, different fresh tracks were selected. For this reason, the comparison of the frictional properties with or without polymers could be performed with exactly the same set up of the tribo-pair. All the measurements were performed at room temperature (~25°C).

All the instrumental parts to be coated with the polymer, were oxygen-plasma cleaned for ca. 2 min in a Harrick Plasma Cleaner/Sterilizer PDC-32G instrument (Ossining, NY, USA) prior to the measurements. Before the plasma cleaning, the ball and disks were sonicated in ethanol for 10 min and dried under nitrogen.

Friction Force Microscopy (FFM). All AFM measurements were performed with a beam-deflection instrument in which the sample is scanned with respect to a fixed tip position. The microscope makes use of a single-tube scanner, employs a Digital Instruments (Santa Barbara, CA, USA) liquid cell/tip holder, and is controlled by STM 1000 and AFM 100 electronics and software (RHK Technology, Inc., Troy, MI, USA). Greater details of this instrumental design have been reported previously [6]. The frictional data have been recorded for the interaction of sodium borosilicate microspheres with clean and polymer-coated silicon oxide surfaces. The microspheres (5.1 micron diameter) are attached to AFM cantilevers (Novascan Technologies, Ames, IA, USA) thus making them entirely compatible with common AFM instruments. Normal forces have been recorded using the manufacturer's reported spring constant (0.58 N/m) while frictional forces, ascribed to the lateral torsion of the cantilever during scanning, have been recorded as the raw voltage output of the photodetector. For all measurements of interfacial friction, data reported in the same plot have been recorded with the identical cantilever tip assembly in order to avoid errors introduced through differences in cantilever force constants.

The frictional properties of interfaces containing PLL-*g*-PEG were investigated by measuring the frictional force (ascribed to the lateral torsion of the cantilever while scanning) as a function of normal load. This was accomplished by rastering the sample in a line-scan mode while first increasing and then decreasing the applied load. During this procedure, frictional and normal forces were measured simultaneously with a scan speed of 1400nm/s over a distance of 100 nm. The reported frictional data represent the average of a number of results obtained at different locations across the surface. Additional details of this procedure can be found elsewhere [6]. All measurements were performed under 10 mM HEPES solutions.

Silicon and microsphere surfaces were prepared for measurement and/or polymer adsorption through a series of steps. These included ultra-sonication in toluene for 1 minute, ultra-sonication in iso-propanol for 10 minutes, rinsing with HPLC grade water, and exposure to an oxygen plasma (see above) for 2 minutes. Following this procedure, samples and/or tips were immediately transferred to a 10mM HEPES solution for storage or polymer adsorption. The polymer was adsorbed onto surfaces prepared in this fashion by immersion for 30 minutes in 10 mM HEPES solutions containing ~1mg/ml of PLL(20)-*g*[3.5]-PEG(2).

3. RESULTS

3.1. Pin-on-Disk Tribometry

The lubricating properties of PLL-*g*-PEG in HEPES buffer, PLL in buffer, and buffer solution alone were measured as a function of both load and velocity under macrotribological conditions. In Figure 3, the friction measured over a narrow velocity range is displayed and is seen not to vary (0.2-0.8 mm/s) with loads varying from 0.5 to 5N (0.32 – 0.70 GPa Hertzian pressure). The data have been averaged over five rotations. It can be readily observed that the HEPES buffer (similar in its behavior to water—results not shown) shows the highest values of friction (μ = 0.35), while PLL displays lower friction (μ = 0.15) and PLL-*g*-PEG even lower values (μ = 0.07).

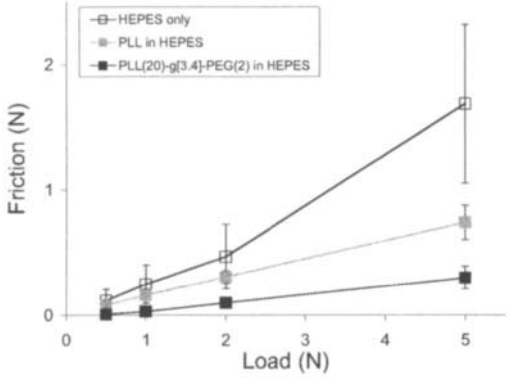


Figure 3. Load-dependent pin-on-disk friction measurements of HEPES, PLL in HEPES, and PLL-*g*-PEG in HEPES. Averaged over five rotations, velocity range 0.2-0.8 mm/s.

The velocity-dependent measurements (Figure 4) were carried out at a normal load of 2 N and averaged over 100 rotations. Again, in general HEPES displays the highest values of friction, with PLL showing reduced values and PLL-*g*-PEG the lowest values. However, while this holds true over most of the velocity range, all three lubricants displayed similar friction coefficients of around 0.05 – 0.07 at the highest measurable velocities. Of particular interest is that while both pure HEPES and PLL show increasing friction with decreasing velocity, the PLL-*g*-PEG solution shows a constant low value of around μ = 0.06(±0.02) over the entire velocity range investigated.

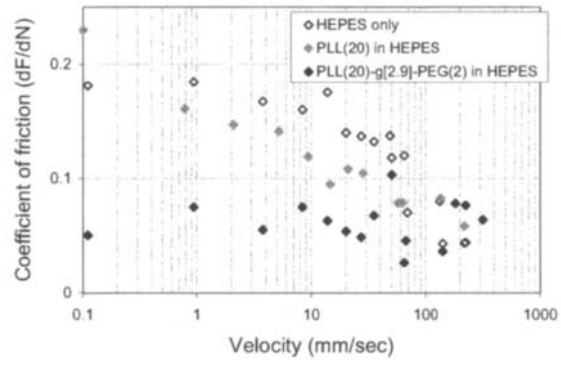


Figure 4. Velocity-dependent pin-on-disk friction measurements of HEPES, PLL in HEPES, and PLL-*g*-PEG in HEPES. Averaged over 100 rotations. Normal load 2N.

3.2. Friction-Force Microscopy

The influence of PLL-*g*-PEG adsorption on the frictional properties of a silicon oxide/sodium borosilicate interface immersed in a 10 mM HEPES buffer solution is illustrated in Figure 5. The data portray distinct differences in the frictional response of the interface depending on the specific geometry of the polymer coating. Three different adsorption geometries were investigated. These included a PLL-*g*-PEG coated microsphere sliding on a bare substrate, a bare microsphere sliding on a PLL-*g*-PEG coated substrate, and a PLL-*g*-PEG coated microsphere sliding on a PLL-*g*-PEG coated substrate. Although the data is not shown, each of these interfaces exhibited a significantly lower frictional response than that of the bare microsphere/bare substrate contact under HEPES solution. For the interfaces represented in Figure 5, both the interfacial adhesion (judged through the negative load at which tip-sample separation occurs) and frictional response (slope of the friction-load plot) are lowest for the contact in which polymer has been adsorbed on both surfaces.

For the surfaces coated on only a single side of the interface, these forces are highest for the polymer-coated microsphere. This result is thought to arise from the greater interaction of the polymer localized at the microsphere apex, relative to the polymer coated on the substrate that experiences relatively less contact during data collection. The shape of the friction-load plots (curvature at low loads) further indicates a difference in the degree of deformation of the interfacial contact as a function of where the polymer has been adsorbed.

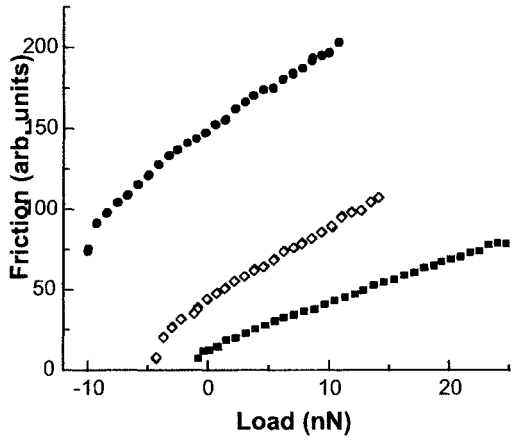


Figure 5. The frictional response of PLL-*g*-PEG-coated interfaces immersed in an aqueous buffer solution (HEPES), measured using an AFM-cantilever assembly modified with a sodium borosilicate microsphere (5.1μm dia.). Friction was measured as a function of applied normal load using the same microsphere/cantilever assembly for: coated tip/bare substrate (filled circles), bare tip/coated substrate (open diamonds), and coated tip/coated substrate (filled squares).

4. DISCUSSION

The tribometry results clearly indicate that PLL-*g*-PEG and PLL can both act to reduce friction to a value below that of buffer (or water). In the case of PLL, this can probably be ascribed to the adsorption of the positively charged macromolecules onto the negatively charged steel and glass surfaces. This leads to the formation of two highly positively charged surfaces, due to excess charge present on those parts of the adsorbed PLL chain that have not been neutralized by surface charge from the oxide. This enhancement (and simultaneous inversion) of the charge leads to an electrostatic repulsion between the surfaces, which lowers the local effective load and therefore the friction. This type of effect has been previously reported for both macro [7] and nano [8] systems.

While PLL-*g*-PEG also contains a charged PLL backbone, it should be remembered that approximately one third of the PLL units have been reacted with the PEG chain and thus the total charge per chain is concomitantly reduced. Nevertheless, it is clear from the pin-on-disk tribometry results that the friction coefficients of surfaces exposed to PLL-*g*-PEG are consistently lower than those exposed to PLL alone. This suggests that presence of the PEG chains is also playing a significant role in reducing the friction, analogously to the polystyrene chains in toluene described by Klein et al [4].

The velocity-dependent data strongly suggest that the PLL-*g*-PEG is functioning as a very effective boundary lubricant in this aqueous environment, since the expected increase in friction at low velocities—frequently observed in the Stribeck curve—was not observed in the presence of PLL-*g*-PEG.

The FFM results demonstrate that two PLL-*g*-PEG-covered surfaces under aqueous solution show substantially reduced adhesion, when compared to bare surfaces, as well as displaying reduced $\Delta F/\Delta L$ values. Thus, the macrotribological behavior of PLL-*g*-PEG is mirrored on the nanoscale, in terms of $\Delta F/\Delta L$, with the addition of an adhesion-reduction component, which could be significant for low-load systems such as micro- or nanodevices. The FFM results also imply that it is important for both surfaces to be covered with PLL-*g*-PEG in order for friction reduction to be effected, and thus it is likely to be the resistance to interpenetration of two brush-

covered countersurfaces [4] that leads to the greatest friction-reduction effect.

5. CONCLUSIONS

Adsorbed PLL-*g*-PEG molecules are clearly capable of functioning as boundary lubricants in an aqueous environment on oxide surfaces with appropriate surface charge. They seem to function both by reducing adhesion (important under nanotribological conditions), and by reducing $\Delta F/\Delta L$, presumably by means of the resistance to mutual chain interpenetration. The net result of PLL-*g*-PEG adsorption is the reduction of friction to very low levels under conditions that normally produce high friction under water lubrication.

These observations suggest that a technological application of such bioinspired, aqueous boundary lubricant systems might be possible, in macro-, micro-, or nano-systems.

6. ACKNOWLEDGMENTS

The authors are grateful for financial support received from the TopNano21 Program of the Council of the Swiss Federal Institutes of Technology (ETH-Rat), as well as the US Air Force Office of Scientific Research under contract #F49620-02-1-0346.

REFERENCES

1. M. Ratoi-Salagean and H.A. Spikes, Tribol. Trans. 42 (1999) 479.

2. M. Scherge and S.N. Gorb "Biological Micro- and Nano-tribology" (Springer, Berlin 2001).

3. D.L. Elbert and J.A. Hubbell, Chem. Biol. 5 (1998) 177; D.L. Elbert and J.A. Hubbell, J. Biomed. Mater. Res. 42 (1998) 55; G.L. Kenausis, J. Vörös, D.L. Elbert, N.-P. Huang, R. Hofer, L. Ruiz-Taylor, M. Textor, J.A. Hubbell, N.D. Spencer J. Phys. Chem. B 104 (2000) 3298; N.-P. Huang, R. Michel, J. Vörös, M. Textor, R.Hofer, A. Rossi, D.L. Elbert, J.A. Hubbell, and N.D. Spencer, Langmuir 17 (2001) 489; S. Lee, M. Müller, M. Ratoi-Salagean, J. Vörös, S. Pasche, S. De Paul, H.A. Spikes, M. Textor, and N.D. Spencer, Trib. Lett. submitted.

4. J. Klein, Annu. Rev. Mater. Sci. 26 (1996) 581; J. Klein, E. Kumacheva, D. Perahia, and L.J. Fetters, Acta Polym. 49 (1998) 617; U. Raviv, R. Tadmor, and J. Klein J. Phys. Chem. B. 105 (2001) 8125; S. Granick, A.L. Demirel, L.L. Cai, and J. Peanasky, Isr. J. Chem. 35 (1995) 75; J. Klein, D. Peraha and S. Warburg, Nature 352 (1991) 143; G.S. Grest, Phys. Rev. Lett. 76 (1996) 4979; T. Kreer, M.H. Müser, K. Binder, and J. Klein, Langmuir 17 (2001) 7804.

5. K. Feldman, PhD Dissertation, TU Eindhoven, 2000.

6. S. Lee, Y.-S. Shon, R. Colorado, R.L. Guenard, T.R. Lee, and S.S. Perry, Langmuir 16 (2000) 2220; H.I. Kim, T. Koini, T.R. Lee, and S.S. Perry, Langmuir 13 (1997) 7192.

7. G.H. Kelsall, Y.Y. Zhu and H.A. Spikes, J. Chem. Soc. Faraday Trans. 89 (1993) 267

8. G. Hähner, A. Marti and N.D. Spencer, Trib. Lett. 3 (1997) 359.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Eyelid/ lens contact during a 100 metres sprint. A tribological approach

B. Bou-Saïd

Laboratoire de Mécanique des Contacts Bât Jean d'Alembert
INSA 20 Avenue A. Einstein 69621 Villeurbanne Cedex France
Tel : + 33 4 72 43 84 47
Fax : + 33 4 78 89 09 80
Benyebka.Bou-Saïd@insa-lyon.fr

ABSTRACT

Contact lens practitioners know the importance of adequate soft lens movement to maintain comfort, good vision and corneal integrity during lens wear. This contact is a fascinating subject with interesting features that relate to the science of tribology. The interaction of the eyelid with a contact lens during the event of blinking can be considered in a first approach as the tribological problem of a conformal contact where the lens represents the moving surface , the eye plays the role of the stationary pad, and the tear film is the lubricant. It is clear that the motion of the contact lens on the cornea is greatly affected by the blinking action. Modeling this contact requires a consideration of the lens, the cornea and the eyelid as porous visco-elastic media and the lachrymal fluid composed of three phases : a mucous phase (0.02 to 0.04 μm), an intermediate phase (6 to 10 μm) and finally a lipid film of 500 angstroms. During a sprint of 100 meters which takes 9.79 seconds (world record) the racer touches the ground 44 times that means 44 shocks on the eyelid/lens contact. Modeling this contact during a shock needs the lubricant to be considered as non-Newtonian with specific viscoelastic properties in order to obtain correctly the response of the fluid. This shock creates an overloading and modifies the contact equilibrium. It can lead to trouble with vision and upset the performance of the sprinter. Thus it is possible to optimize the lens mechanical properties to avoid this kind of disturbance. Starting from a tribological approach of an eyelid/lens contact a steady regime is anlysed initially and then the response of this contact to a shock is considered with the lubricant as non-Newtonian.

1. Introduction

To optimize the performances of a lens, it is necessary to extract its intrinsic characteristics allowing one to study the mechanics of the eye. All the elements present have to be considered and their properties, namely, the lachrymal fluid, which is a nearby particular fluid with water behavior, the eyelid, visco-elastic and the porous lens or cornea. The lachrymal film is constituted of three layers :

- a lipid layer of 0.1 μm of thickness
- an intermediate water like layer of 7 μm of thickness
- a mucous layer deep from 0.4 to 1 μm of thickness, connected to the cornea by the epithelial micro-villosities.

In 1966, Mishima [1] determined the lachrymal volume and the lachrymal flow by fluorometric methods. He measured a lachrymal volume of 6.2 ± 2 μl and a lachrymal flow of 1.2 μl / min with variations from 0.5 to 2.2 μl / min. The role of the eyelid is to insure :

- the mechanical protection of the globe against attacks
- the hydration corneo-conjonctival by distribution of the lachrymal film
- the limitation of its evaporation during the sleep
- the evacuation towards the lachrymal ways
- the regulation of the penetration of light in the eye in the waking state according to the luminous intensity and according to the visual sequences.

The skin of the eyelid is extremely fine, delicate and flexible [7]. It presents a high elasticity and deformability. The subcutaneous tissue is also extremely fine, devoid of fat, contributing to the exceptional flexibility of this region and to the requirements of strong mobility. The eyelid opens, closes, is animated by rhythmical, spontaneous blinking or reflexes. These movements are anatomically and physiologically connected to the eye movements.

The eyeball is a hollow organ about 25 mm in diameter. The cornea is the front of this eye shell. Its fundamental properties are its transparency and its porosity. It allows the light to penetrate into the eye and reflects or refracts them. The iris is situated just behind the cornea [9].

The palpebral movement is characterized by the amplitude of movement of the free edge in the center of the eyelid, between a downward glance then an upward glance. On average, the values according to Fox [2] are 14.4 mm for the upper eyelid and 5.3 mm for the lower eyelid. The upper eyelid is thus more mobile than the lower eyelid. During a spontaneous blinking, the global duration of the movement is about 260 ms. The superior eyelid moves in 80 ms by accelerating, reaching the maximal speed about 19 cm/s at the time of the crossing of the optical axis. It slows down in the approach of the free edge which mostly is not touched, stands out a pause from 3 to 5 ms and goes back up a little less fast by slowing down just before the total opening [4,5]. The amplitude of the moving upper palpebral during blinking is 9 mm, the lower eyelid staying about at the same level. Every individual has, in a sense , a personal basic rhythm, but the normal frequency is from 12 to 20 blinkings per minute as an adult [6].

Different works were published to deal with this very complex problem. We can cite mainly the papers of Hayashi [11], Conway et al [12] and Hayashi et al [3]. They present nevertheless restrictive approaches in term of modeling (geometry, mechanical properties,...). In this article a global approach is presented using a Finite Element idealization which allows a relaxation from many restrictive hypotheses and enables the problem of a runner of 100 m wearing contact lenses to be analyzed.

2. Basic equations

This contact between the eye and the eyelid through the liquid lachrymal as the third body leads to a tribological approach to this phenomenon. This contact between three bodies is coupled . It takes into account the contact of the eyelid on the lachrymal fluid as well as that of the lachrymal fluid on the cornea and the lens .The intrinsic performances of every element are determined using basic equations. Thus, the equation governing the behavior of the lachrymal fluid is that of a modified Reynolds equation [13], assuming the following hypotheses :

- thin film (h/L = 10μm / 10mm=1 E-03)
- non-Newtonian fluid
- continuous medium
- laminar flow
- body and inertia forces are neglected
- no slip at the walls

The general form of this equation is :

$$\frac{\partial}{\partial x}\left(G\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(G\frac{\partial p}{\partial y}\right)=$$

$$\frac{\partial}{\partial x}\left(U_2(R_2-F)+U_1F\right)-\rho_2U_2\frac{\partial h}{\partial x}+\frac{\partial}{\partial y}$$

$$\left(V_2(R_2-F)+V_1F\right)-\rho_2V_2\frac{\partial h}{\partial y}+\rho_2W_2$$

where R, F, G et R_2 are :

$$R=\int_0^z \rho(x,\xi,y,t)d\xi$$

$$F=\frac{1}{J_2}\int_0^h\frac{R}{\mu}dz \text{ with } J_2=\int_0^h\frac{dz}{\mu}dz$$

$$G=\int_0^h\frac{Rz}{\mu}dz-I_2F \text{ with } I_2=\int_0^h\frac{z}{\mu}dz$$

$$R_2=\int_0^h\rho\,dz$$

with :
h = film thickness
p = pressure
U, V, W = surface velocity components in x, y, z directions

The weak form is :

$$\int_D \left(G\frac{\partial p}{\partial x}\delta\left(\frac{\partial p}{\partial x}\right) + 2U(h-F)\delta\left(\frac{\partial p}{\partial x}\right) - G\frac{\partial p}{\partial y}\delta\left(\frac{\partial p}{\partial y}\right)\right) dV$$
$$- \int_{\Gamma_q} q\,\delta p\,dl = 0$$

with :

Γ_q : boundary where the flow is imposed
q : exiting flow at the boundary

Coupled with the modified Reynolds equation, the equation of the motion of the thin fluid film with regard to the porous cornea or lens requires [7,8] the resolution of the Darcy's equation which reads :

$$\frac{\partial}{\partial x}(k_x\frac{\partial p}{\partial x}) + \frac{\partial}{\partial y}(k_y\frac{\partial p}{\partial y}) = 0$$

with :
p = pressure
k_x, k_y = diffusion coefficients in the x and y directions respectively

The weak form of Darcy's equation is :

$$V_2 = \int_D \left[k_x\frac{\partial p}{\partial x}\frac{\partial}{\partial x}(\delta p) - k_y\frac{\partial p}{\partial y}\frac{\partial}{\partial y}(\delta p)\right] dxdy$$
$$+ \int_{\Gamma_1} k_x\frac{\partial p}{\partial x}\delta p\,dy + \int_{\Gamma_1} k_y\frac{\partial p}{\partial y}\delta p\,dx$$

with $\Gamma_1 \cup \Gamma_2 = D$

We can also determine the deflection of the eyelid subject to a field of pressure stemming from the fluid and from the movement of the eyelid. The equation of the membrane type is for an oblong membrane of the shape :

$$w = \frac{pa^2}{S_0 h}\frac{8}{\pi^4}\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}$$

with : w = deflection
a, b, h = membrane dimensions
p = pressure
S_0 = stress of the membrane not influenced by the load

3. Geometry and kinematics

The geometry and the schematic of the contact are presented in figures 1, 2 and 3. Let us consider that the palpebral motion is uniform, of average speed 12 cm/s. Furthermore, we do not differentiate the phase of the closing and opening of the eyelid. We consider the lachrymal fluid as being composed of a single homogeneous layer. The viscosity of the human lachrymal fluid varies between 0.0013 to 0.0059 Pa.s. To consider the three layers we assume a single homogeneous non-Newtonian third body. Different rheological laws have been tested. They are the following :

1 – Newtonian model

$$= \frac{1}{\mu}\tau_{ij}$$

2 – Linear visco-elastic model

$$= \frac{1}{G}\frac{d\tau_{ij}}{dt} + \frac{\tau_{ij}}{\mu}$$

3 – First nonlinear visco-elastic model

$$= \frac{1}{G}.\frac{d\tau_{ij}}{dt} + \frac{\tau_0}{\tau_e}.\frac{\tau_{ij}}{\mu}sh\left(\frac{\tau_e}{\tau_0}\right)$$

4 – Second nonlinear visco-elastic model

$$= \frac{1}{G}.\frac{d\tau_{ij}}{dt} + \frac{\tau_0}{\tau_e}.\frac{\tau_{ij}}{\mu}th\left(\frac{\tau_e}{\tau_0}\right)$$

5 – Visco-elastic-plastic model

$$= \frac{1}{G}.\frac{d\tau_{ij}}{dt} + \frac{\tau_{ij}}{\tau_e}.\frac{\tau_L}{\mu} In\left|1 - \frac{\tau_e}{\tau_L}\right|$$



Figure 1 : Schematic of the geometry

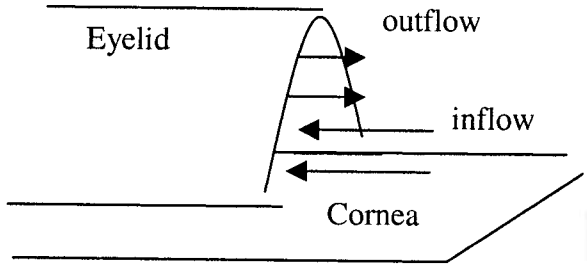


Figure 2 : Schematic of the contact

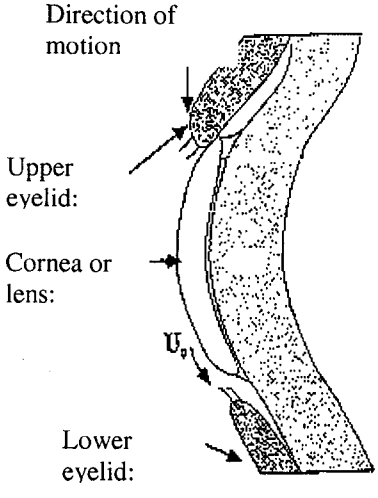


Figure 3 : Contact idealization

4. Some results

We first simulate the movement of blinking. Figure 4 represents the contact idealization. The data are the following :

- displacement amplitude :
- 14.4 mm (upper eyelid)
  5.3 mm (lower eyelid)
- cornea diameter : 25 mm
- blinking time : 26 ms
- blinking speed : 19 cm/s
- frequency : 12 to 20/min
- imposed load : 1.2 N
- lachrymal fluid flow : 1.2 μl/min
- eyelid flexibility : 0.2 10-9 m3/N

Figures 4 and 5 present, respectively , the lachrymal field pressure during motion and the field pressure in the cornea. In figures 6,7 and 8 we can see the evolution of the pressure gradients in x and y directions and the shear stress representation.

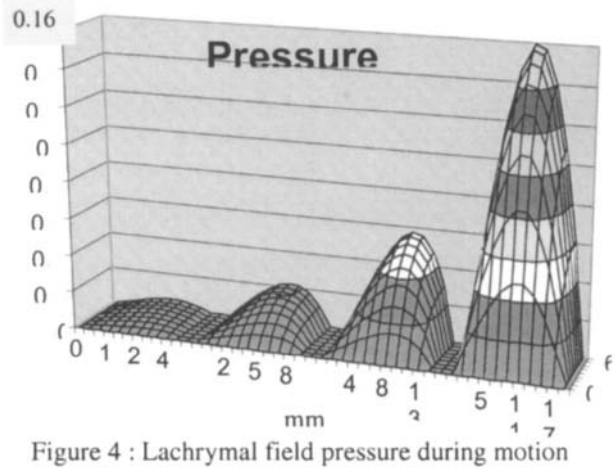


Figure 4 : Lachrymal field pressure during motion

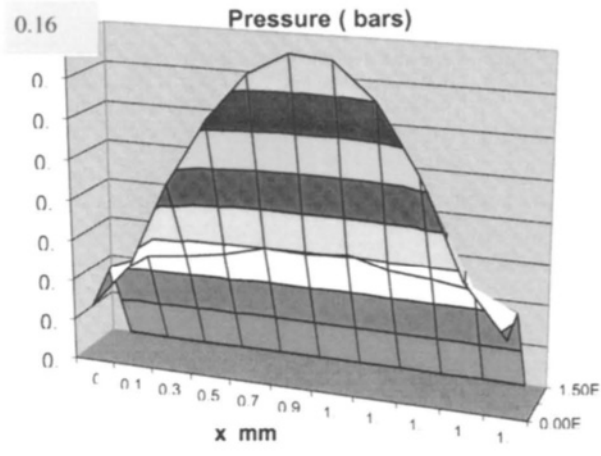


Figure 5 : Field pressure in the cornea

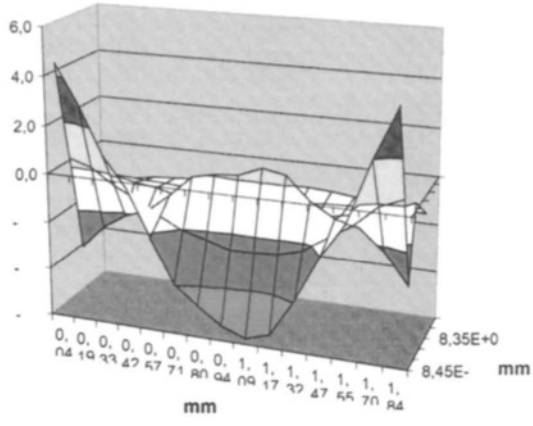


Figure 7 : Pressure gradient in y direction (bars/m)

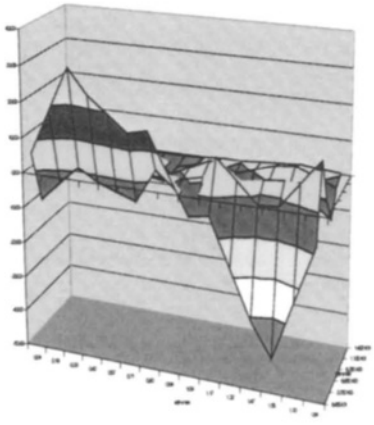


Figure 6 : Pressure gradient in x direction (bars/m)

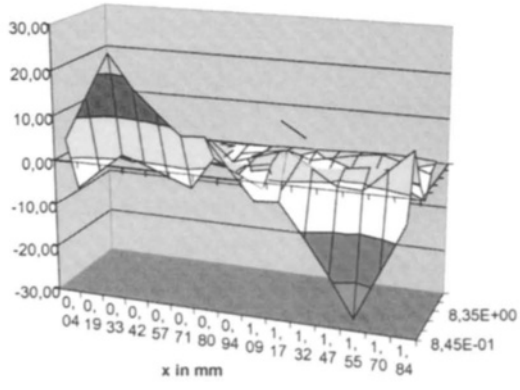


Figure 8 : Shear stress (bars/m)

It should be noted that the model developed in order to characterize the contact between the eyelid and the lens or the cornea gives realistic and useful results. A minimum film thickness of 6 μm is predicted and a lacrymal flow of 1.3 μl/min It is interesting to point out that numerical simulations can be performed to guarantee a maximum comfort to people wearing lenses by changing ,for instance,

the geometrical parameters while respecting the physiological constraints.

5. Eyelid/ lens contact during a 100 meters sprint

We consider now the example of a runner wearing contact lenses. The following data were used:

- time for 100 m : 9.79 s
- weight: 76 kg
- number of strides : 44
- inpulse tangential force : 3.2 N [14]

The problem is to characterize the contact during the shock of a stride. The rheological model of Johnson and Segalmann [10] were considered to take into account possible stretching and overshooting stress effects. The expression of the shear stress is :

small shearing

$$S_{xy}(t)=G(t)\overset{0}{\gamma} \quad \text{with} \quad G(t)=\frac{\eta}{\lambda}e^{-t/\lambda}$$

high shearing

$$S_{xy}(t)=\frac{\eta}{\lambda}e^{-t/\lambda}\frac{\sin\left(\sqrt{\xi(2-\xi)\overset{0}{\gamma}}\right)}{\overset{0}{\gamma}\sqrt{\xi(2-\xi)}}$$

with η the viscosity, λ the relaxation time, t the time, ζ a sliding coefficient and $\overset{0}{\gamma}$ the shear strain. Figures 11 and 12 present the dimensionless first normal difference stress against dimensionless time with $Wc = \lambda\overset{o}{\gamma}$

A stress overshoot during the shock is observed which resutlts in an increase of the gap between the lens and the cornea. Figures 9 and 10 give the dimensionless shear stress. Consequence: approximately 30 % of the supplementary forces generated a 25 % increase of the distance of separation between lens and cornea.

For a normal eye : $\cong$ 15 % of variation introduces a visual disturbance.

In conclusion : this phenomenon does not affect the race of Maurice Green and he would have the world record for the 100 m.

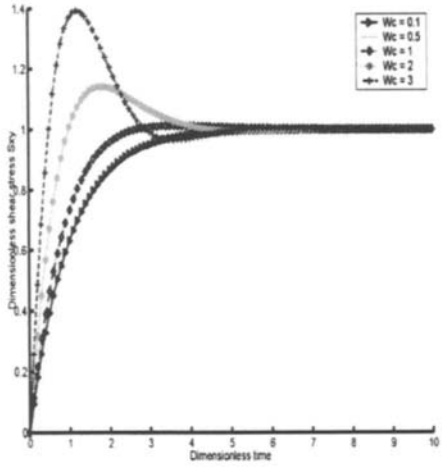


Figure 9 : $\zeta = 0.1$

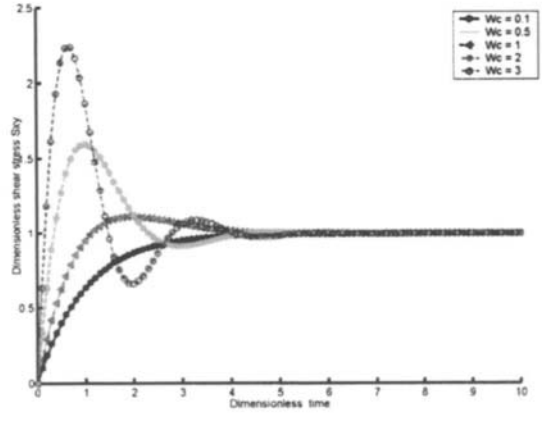


Figure 10 : $\zeta = 0.4$

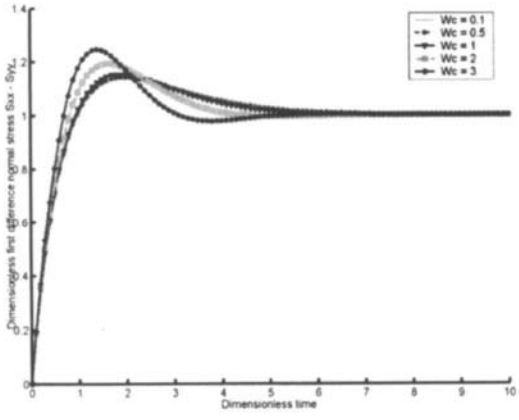
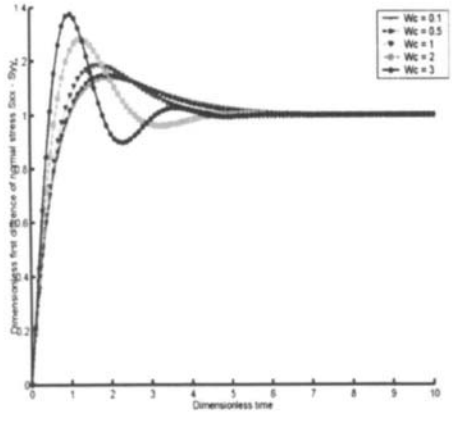


Figure 11 : $\zeta = 0.1$



6. General conclusion

In this approach, a hydrodynamic analysis of a soft contact lens during the sliding motion as well as during a shock is presented. The model developed allows a parametric study of the contact in order to give an adequate soft lens movement to maintain comfort, good vision and corneal integrity during lens wear. During a sprint a visual disturbance occurs due to the existence of normal stresses and the increase of shear stress.

Bibliography

[1] JP Adenis, S Maes , J-M. Ebran, 1992- « Physiologie de l'excrétion lacrymale », encyclopédie Médicale Chirurgicale, Ophtalmologie, 21-020-B-10, p1-8.

[2] Jean-Luc George, 1998- « Physiologie des mouvements palpébraux »,Encyclopédie Médicale Chirurgicale, Ophtalmologie, 21-020-A-10, p1-6.

[3] S.Kamiyama , M.M.Khonsari, 1999- « Hydrodynamics of a Soft Contact Lens during Sliding Motion ».

[4] Donald K. Martin, 1987-« Empirical Analysis of the Motion of Soft Contact Lenses across the human eye », Australian Physical & Engineering Sciences in Medicine, Vol.10,No.4, p214-221.

[5] Peter E. Raad , Adrian S. Sabau, 1994- « Permeable Contact Lens Motion during Blinking », Advances in Bioengineering, BED-Vol.28, p75-76.

[6] John A. Nairn , Tong-bi Jiang, 1995- « Mesurement of the friction and lubricity properties of contact lenses »,ANTEC, p3384-3388.

[7] John C.Corzine , Stanley A. Klein, 1997- « Factors determining Rigid Contact Lens Flexure », Optometry and Vision Science, Vol.74, No.8.

[8] Alexandra J.Taylore , Simon D.R. Wilson, 1994-« Post-Lens Tear Film Thinning in Rigid Gas Permeable Lenses », Optometry and Vision Science,Vol.7 2, No.12, p849-856

[9] JP Adenis, PY Robert , MP Boncoeur-Martel, 1996-« Anatomie des glandes et des voies lacrymales », Encyclopédie Médicale Chirurgicale, Ophtalmologie, 21-006-A-25, p1-9.

[10] N.W Johnson, J. Segalmann "A model for viscoelastic fluid behaviour which allows Non- Affine deformation " Journal of Non-Newtonian Fluid Mechanics ,1977, Vol. 2 ,p. 255-270.

[11] Hayashi T., T. ' Mechanics of contact lens motion ' Ph. D. thesis . University of California. Berkeley. 1977

[12] Conway H. D. , RichmanM. ' Effects of contact lens deformation of tear film pressure iinduced during blinking ' Am. J. Optom. Physio. Opt. 59, N0 1, pp1-20

[13] B. Najji, B. Bou-Saïd, D. Berthe New formulation for lubrication with non-newtonian fluids". Journal of Tribology Vol. 111 n° 1 p. 29 Jan. 1989

[14] Y.C. Fung ' Biomechanics : mechanical properties of living tissues ' Second edition . Springer . 1993

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Biological lubrication of hydrated surface layer in small intestine

H. Yoshida, Y. Morita and K. Ikeuchi

Institute for Frontier Medical Sciences, Kyoto University,
53 Kawaharacho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

Frictional properties of the internal surface of small intestine (lamina epithelialis) were investigated. Friction between the internal surface of small intestine of porcine and silicon nitride ceramic was measured using an end-face friction test apparatus. The internal surface of small intestine was found to have a highly hydrated layer to protect tissue and transport food under digestion. Friction depends on water content of the surface layer because the water absorbed by the layer plays an important role in keeping friction low. The layer can recover low friction condition after water exudation by absorbing water again. At low sliding speed, friction is kept low and the tissue surface is protected from injury because the surface layer absorbs water. The mechanism of lubrication results from unique properties of water and is entirely different from industrial lubrication.

1. INTRODUCTION

The tissue surfaces in a human body, such as articular cartilage, tendon, fascia and digestive tract, slide with very low friction, consequently preventing damage. Many investigations have been reported about the tribological properties of living tissues, especially articular cartilage [1-4]. Human joints indicate low friction coefficients in the range 0.005 to 0.03. Articular cartilage with water content 70-80% is tender and slides slowly under enormous contact stress, nevertheless its lifetime is longer than that of the average human. The main lubrication mechanism of joints is elastohydrodynamic lubrication (EHL) [5], while supplementary lubrication is necessary because a human being walks and stops repeatedly and the load on joints changes irregularly.

Sasada et al. proposed a model named "surface gel hydration lubrication" in order to explain the lubrication mechanism of cartilages [6,7]. In this study, it was assumed that a part of proteoglycan molecules sticks out of the surface of cartilage and forms the layer of gel hydrated with lubricant.

Some tissues have similar surface layers composed of hydrophilic macromolecules which are combinations of polysaccharides and proteins. Though several studies of polysaccharides have been carried out by biochemists, few investigations have been made on the mechanical properties of the surface layer.

The purpose of this study is to investigate the tribological properties of the internal surface of the small intestine (lamina epithelialis) and to examine the lubrication mechanism of the hydrated layer on the tissue.

2. MATERIALS AND METHODS

Friction between the internal surface of the small intestine of porcine and silicon nitride ceramic was measured using an end-face friction test apparatus. Figure 1 illustrates the shapes and the dimensions of specimens.

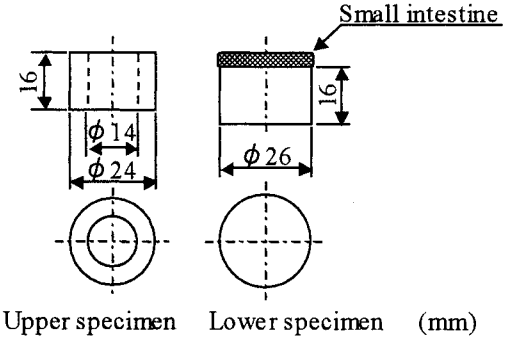


Figure 1. Shapes and dimensions of specimens

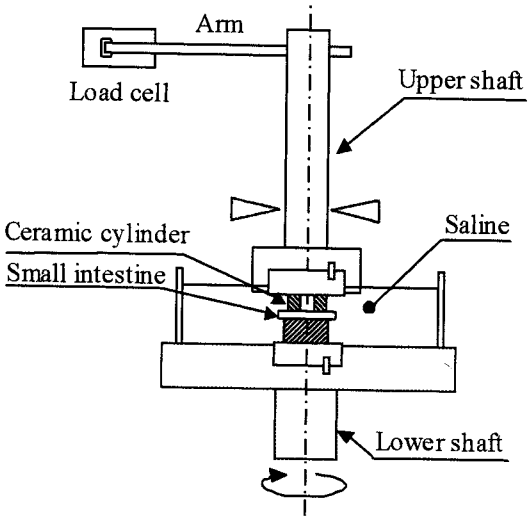


Figure 2. End-face friction test apparatus

Small intestine preserved at -30℃ was thawed out at room temperature before tests. Small intestine was attached to the lower disk with tension. The upper ring-shaped specimen is made of silicon nitride ceramic. The surface roughness of the upper specimen was 0.02 μm Ra. Figure 2 shows schematic drawing of the apparatus. The lower part rotates with a water vessel filled with saline solution. Mean contact pressure was 12 kPa and the sliding speed was 20 mm/s. Sliding speed was calculated using the equivalent radius r_m (mm) defined by Ikeuchi et al. [8] as follows.

$$r_m = \int_{r_i}^{r_o} \frac{r^2 \pi r\, dr}{\pi(r_o^2 - r_i^2)} = \frac{2(r_o^2 + r_o r_i + r_i^2)}{3(r_o + r_i)} \qquad (1)$$

where r_0 (mm) is the outer diameter and r_1 (mm) is the inner diameter of upper specimen. Torque was measured using a load cell placed at the tip of the arm and converted into frictional force at r_m, and then friction coefficient was calculated. The tests were carried out at room temperature (25℃).

(a) Effect of the surface conditions on friction

Three kinds of surface condition were prepared to investigate how the conditions of surface affect friction. (i) Intact condition: The state of the surface was not treated and it was covered with food under digestion; (ii) Soft wiped condition: Mucus on the surface was wiped off softly with gauze; (iii) Hard wiped condition: The surface was rubbed so strongly with gauze as to remove the surface layer.

(b) Role of absorbed water in lubrication

It is considered that much water is attracted by hydrophilic macromolecules in the surface layer. Sliding was started after a certain rest under a load in order to vary water content of the layer. Resting times were 15, 30 and 45min.

(c) Water absorbing ability of surface layer

Intermittent sliding test was carried out in order to investigate whether the surface layer absorbed water. 4 min sliding - 1 min resting was given repeatedly. Specimens were allowed to make contact during the rest. On the other hand, they were separated in order that lubricant might come into the surfaces.

(d) Sliding speed dependency of friction

Sliding speed was changed to study the effect of sliding speed on friction. Sliding speed were 5, 10, 20, 40, and 80 mm/s.

3. RESULTS

Figure 3 shows the effect of the surface conditions on friction.

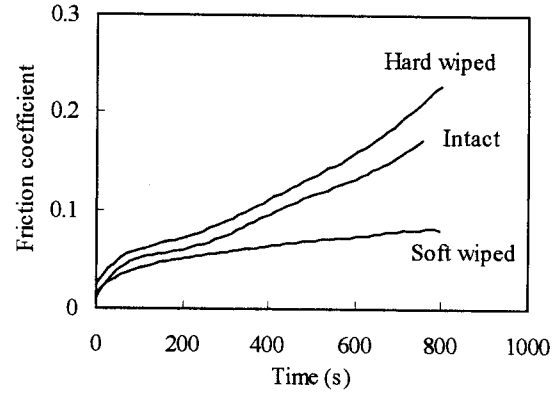


Figure 3. Effect of surface conditions on friction

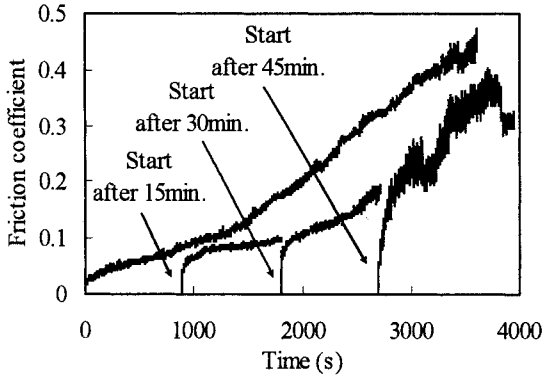


Figure 4. Effect of resting time under a load on friction

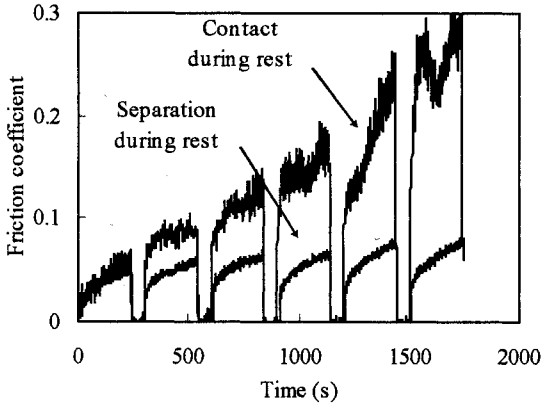


Figure 5. Water absorbing ability of surface layer

The friction coefficient of soft wiped surface is lower than that of intact tissue. The friction coefficient of hard wiped condition is highest among all.

Figure 4 shows the effect of resting time under a load on friction. The start-up friction and the rate of increase of friction coefficient are higher with increasing resting time.

Figure 5 shows the result of intermittent sliding test. If specimens are contacted each other during the rest, the friction coefficient is not changed before and after the rest. On the other hand, if specimens are separated each other during the rest, the friction coefficient after the rest is as low as that at start of the test.

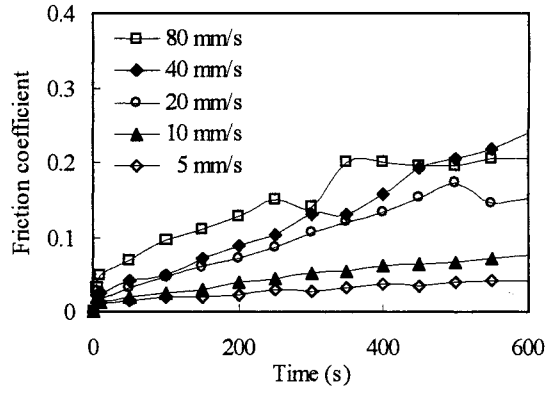


Figure 6. Effect of sliding speed on friction

Figure 6 shows the effect of sliding speed on friction. Coefficients of start-up friction and dynamic friction increase with an increase of sliding velocity.

4. DISCUSSION

Friction coefficient decreases when only the viscous fluid coating the surface of small intestine is removed. However, the friction coefficient increases again after the surface is rubbed more strongly (Fig. 3). This result suggests that small intestine has surface layer by which friction is kept low.

Figure 7 shows a schematic model of hydration lubrication.

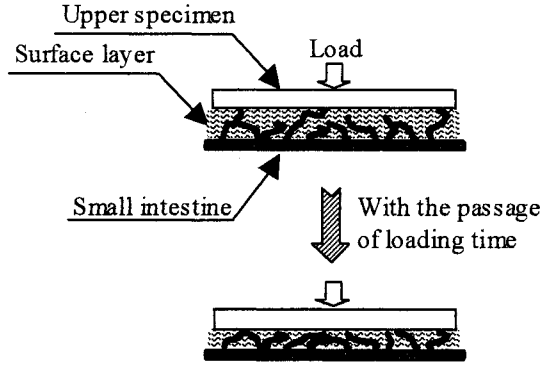


Figure 7. Schematic model of hydration lubrication

Water attracted by hydrophilic macromolecules in the surface layer plays an important role in lubrication. As the macromolecules are flexible, they cannot support a load by themselves. The water in the surface layer would support most of the load because the water are attracted by the macromolecules. Frictional force arises due to adhesion of macromolecules to the mating surface.

The time dependent property of frictional forces can be interpreted as follows. Under a load, water exudes slowly from the surface layer with or without sliding. As the result of water loss, thickness of the surface layer reduces and water content of the surface layer decreases. Consequently, the degree of adhesion to the upper specimen increases and frictional force increases. Therefore, it may be concluded that friction depends essentially on water content of the surface layer.

As Fig. 6 shows, a lower sliding speed results in a lower friction coefficient. The frictional property which is entirely different from that of the engineering tribology also seems to be owing to the water absorbed by the surface layer. Since the fluidity of the water absorbed by the layer is low, it can lubricate the surface as well as a highly viscous fluid can, even if the sliding speed is low.

According to the intermittent test (Fig. 5), the surface layer is restored to low friction condition when it is exposed to lubricant during the rest. This would be due to water absorption by the surface layer and the decrease of adhesion. The surface layer seems to have the ability to recover the condition of low friction.

5. CONCLUSION

The frictional properties of small intestine were investigated in this study. The conclusions are as follows.

1. The internal surface of small intestine (lamina epithelialis) has a highly hydrated layer to protect tissue and transport food under digestion.
2. Friction depends on water content of the surface layer because the water absorbed by the layer plays an important role in keeping friction low.
3. The layer can recover low friction condition after water exudation by absorbing water again.
4. At low sliding speed, friction is kept low and the tissue surface is protected from injury because the surface layer absorbs water.

REFERENCES

1. D. E. Bereznick, J. K. Ross and S. M. McGill, Clin. Biomech., 17 (2002) 297
2. F. C. Linn, J. Biomech. 1 (1968) 193
3. M. D. Longfield, D. Dowson, P. S. Walker and V. Wright, Biomed. Engng., 4 (1969) 517
4. H. Forester and J. Fisher, Proc. Instn. Mech. Engrs., 210 (1996) 109
5. K. Ikeuchi, J. Jpn. Soc. Tribologist, 45, 2 (2000) 108
6. T. Sasada, J. Jpn. Soc. Clin. Biomech., 21 (2000) 17
7. Y. Ishikawa, T. Sasada and K. Ikeuchi, J. Jpn. Soc. Biomater., 19, 3 (2001) 89
8. K. Ikeuchi, T. Takii, H. Norikane, N. Tomita, T. Ohsumi, Y. Uyama and Y. Ikada, Wear, 161 (1993) 179

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribology of a Snail (Terrestrial Gastropod)

Akiyoshi Kobayashi[a], Ikuo Yamamoto[b] and Tsuyoshi Aoyama[b]

[a]Departtment of Mechanical Engineering, Meijo University,
Shiogamaguchi1-501, Tenpaku-ku, Nagoya, Japan 468-8502

[b]Graduate School of Meijo University

The locomotion characteristics of a snail (terrestrial gastropod) that can move forward extremely smoothly not only on the horizontal substratum but also even on walls and reverse on ceilings, are examined experimentally in the tribological standpoint. The locomotion can be achieved by the combination of the progressive wave (pedal wave) of the ventral surface of the foot and the secreted viscous mucus. From pulling off force experiments, the holding force of a snail to a substratum does not depend on the sole area. According to NMR analysis, viscous secreted mucus of consists of about93% of water and 7% of glycoprotein. From the experimental examinations of mechanical properties such as viscosity variation versus shear rate, the relationship between shear strength and mucus film thickness and the relationship between shear rate and shear stress, the secreted viscous mucus of a snail is a non-Newtonian fluid that has visco-plastic characteristics. The movement of snails is obtained by a fine peristalitic motion, based on the combination of the concave shaped contraction of the ventral surface muscle during the passage of the pedal wave and the stretched ventral surface muscle without pedal wave passing. The driving force is generated at the anterior part of the ventral surface as a whole and the posterior part is only dragged. The ventral surface of the foot has a sucking in function and snails circulate secreted mucus around the sole and reuse it skillfully.

1. INTRODUCTION

Animals have adapted themselves to the natural environment and evolved over a long time in order to survive in very strict nature. As a result, animals have an excellent motor function. On the other hand, the economization of energy and the preservation of natural environment are the most important thesis and problem at present. Therefore, animals have many learning points on how to create a high performance machine for the future.

Snails (terrestrial gastropod) can move forward extremely smoothly on the surface (even wall and ceiling) through the secreted viscous mucus. It is said that the locomotion is achieved by the combination of the progressive wave (pedal wave) of the ventral surface of the foot and the secreted viscous mucus. Lissman [1] investigated the mechanics of snail's locomotion, especially the pedal wave motion. Trueman examined the relationship between mucus and muscle of snail [2]. Denny also examined the role of mucus [3]. Ikeuchi et al. tried to model movement mechanics on mucus [4]. However, the locomotion mechanism and the reason why snails can move extremely smoothly not only on the horizontal substratum but also on a wall and even on a ceiling in reverse state is still not clear.

In this paper, the locomotion characteristics of snails were examined experimentally from the tribological viewpoint. Namely, the holding force of snails to the substratum was examined initially in relation to the ventral surface area. Next, the various characteristics (the ingredients, mechanical properties) of the secreted

viscous mucus were examined. Moreover, the locomotion mechanism of the ventral surface was analyzed experimentally based on the observation by CCD camera.

2. HOLDING FORCE ON SUBSTRATUM

Snails can move smoothly even on the wall and the ceiling, because primarily snails can hold themselves on the substratum by the attaching force generated by the combination of the molluscan ventral surface of the foot and the secreted mucus film. It is considered that this force is generated based on the adhesional effect of the viscous secreted mucus itself and the effect of atmospheric pressure.

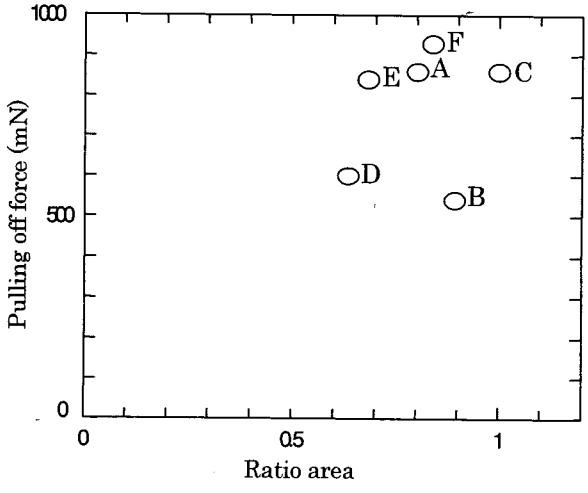


Figure 1. Relationship between contact area and pulling off force

Table 1 Weight and pulling off force [mN]

	Weight	No. 1	No. 2	No. 3
A	64.68	1182	914.3	844.8
B	72.52	703.6	590	551.7
C	84.28	1282	933.9	858.5
D	50.96	1064	661.5	586
E	52.92	1158	902.6	836.9
F	67.62	1487	991.8	933.9

In order to examine the relationship between this force and the area of the ventral surface of the foot, the pulling off force when a snail on a glass plate was detached was measured by a leaf spring and strain gages and also the contact area of a snail on a glass plate was measured by a CCD camera at the same time. Experiments were done with 6 snails. As shown in Table 1, the scatter of the pulling off force of each snail decreased with the number of pulling off tests. Therefore, the third value of the pulling off experiments was chosen. Fig.1 shows the relationship between pulling off force and the area ratio of the ventral surface of 6 snails. The area ratio is decided as the ratio that the area of each ventral surface to the ventral surface of the biggest snail (84.28 mN weight). As shown in the figure, the pulling off force does not depend on the area ratio.

3. SECRETED MUCUS

Snails can move forward extremely smoothly on surfaces (even wall and ceiling) always through the secreted viscous mucus. Therefore, the viscous secreted mucus of a snail influences the locomotion characteristics extremely strongly. As an obvious actual fact, it was observed that snails could not go forward at all when they were set on the highly absorbing water material like a tissue paper.

Firstly, the ingredients of the secreted mucus were measured. The secreted mucus collected from the shell consisted of about 93% of water and the remaining 7% of glycoprotein. Whether the remaining 7% is glycoprotein or not was analyzed by NMR. Denny [3] also reported almost the same result. The density of the mucus was 940.25 (kg/m^3). Secondly, the thickness of the mucus on a glass plate traced after horizontal locomotion was measured by laser microscope. Although the secreted mucus after crawl dries very rapidly from the edge, the average thickness was about 6 μm

3.1. Mechanical properties of secreted mucus

Fig.2 shows the relationship between the shear rate

and the viscosity change of the secreted mucus of 20 μm thickness measured by the cone-plate type of viscometer at the temperature of 298 K. The viscosity of the secreted mucus is very different by the secreted conditions. Because the collection from the trace of secreted mucus after crawl was very difficult, the mucus was collected from in the shell. When the mucus is stored in the test tube, the mucus separated into two layers of different viscosities, the higher viscous mucus floated on the upper side and the lower viscous mucus sank to the bottom side. As seen in the figure, the higher viscous mucus shows values of viscosity as large as 10 times that of the lower viscous mucus. The viscosities of both mucus decreased rapidly as the shear rate increased until 50 /s, and the subsequent decrease was very slow. On the other hand, the measurement of the viscosity by the vibration type viscometer was impossible by Weissenberg effect [5], due to the separation of glycoprotein and water during the test.

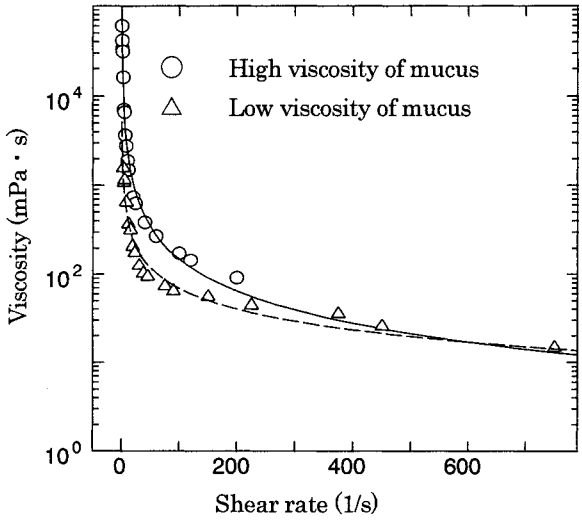
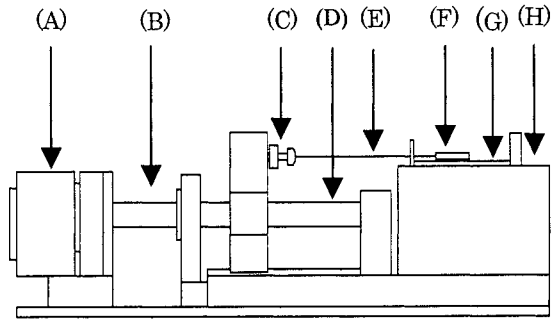


Figure 2. Relationship between viscosity and shear rate

Next, the shear strength of the mucus was measured. The measurement was done with the apparatus as shown in Fig.3. This apparatus can measure the tangential forces of mucus sandwiched between two glass plates (F) and (G) by a load cell (C) at various velocities.



(A) Speed control motor (E) Wire
(B) Ball screw (F) Glass slider
(C) Load cell (G) Glass base
(D) Liner guide (H) Base

Figure 3. Experimental device

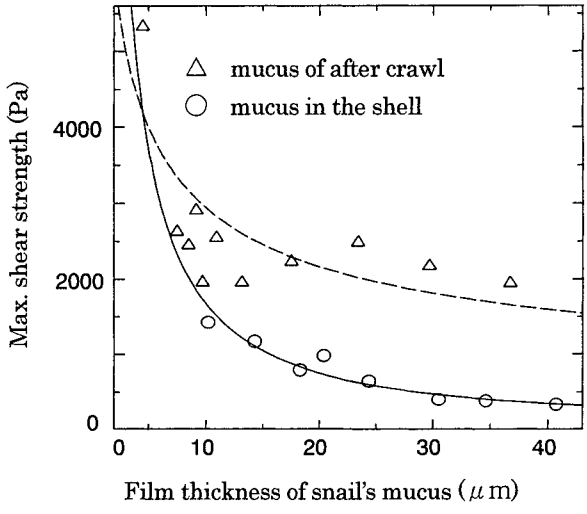


Figure 4. Relationship between film thickness and max. shear strength

Fig.4 shows the maximum shear strength versus the mucus film thickness. The average mucus film thickness was decided as follows. The mucus of 10 μl was set on a base glass of 5cm sq., and pressed and spread by another glass plate. The average film thickness was decided from the spread area and the volume of 10 μl. The pulling velocity was 1.26 mm/s,

that is an average movement velocity of snails. In the figure, the dotted curve is the result of the secreted mucus on the glass plate after crawl, and the solid curve is the result of the mucus collected directly from within the shell of the snail. The viscosity of the mucus in the shell is lower than that of the secreted mucus on a glass plate after crawl. As shown in the figure, the maximum shear strength decreases rapidly for the film thickness to 10 μm and then decreases slightly as the film thickness increases. From this result, the shear strength of the average film thickness (6 μm) is about 2500 Pa and the force becomes about 1.2 N multiplied by the average area of ventral surface of the foot. Therefore, a snail does not slip down and can easily hold itself on the wall in the rest state because a snail's weight is 0.08 N.

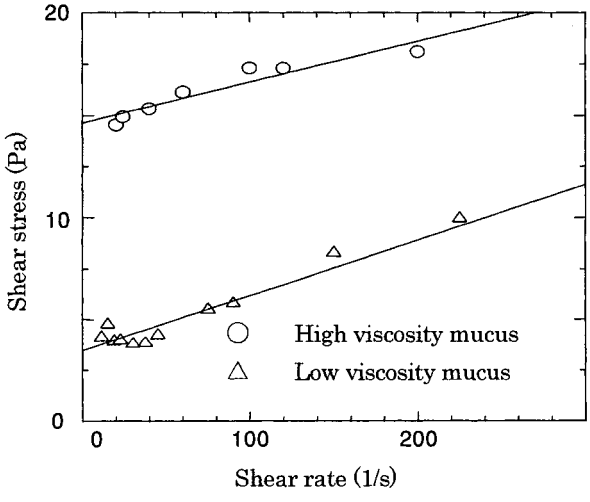


Figure 5. Relationship between shear stress and shear rate

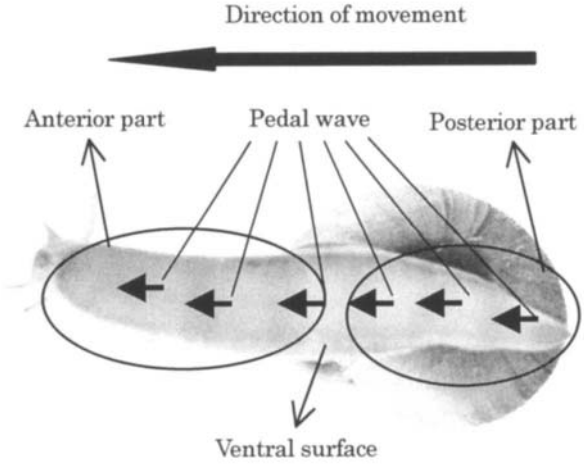
Fig. 5 shows the variation of the shear stress versus shear rate obtained by the cone-plate type viscometer. As shown in the figure, the shear stress has a plastic yield value and increases almost linearly as shear rate increases.

Secreted mucus of a snail is a kind of non-Newtonian fluid that has the characteristics of a visco-plasitic (Bingham) fluid because the viscosity decreases with the increase of shear rate and the shear stress has a plastic yield value and then increases with shear rate.

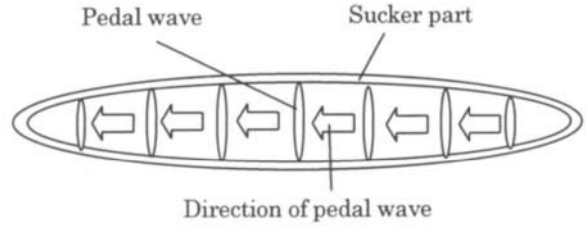
4. LOCOMOTION CHARACTERITICS

4.1. Observation of pedal wave

The snail's movements (horizontal, vertical and reverse on ceiling) on a transparent plate were observed from behind by a CCD camera. Fig. 6 shows a picture and schematic illustration of a ventral surface of the foot. As shown in the figures, snails can move forward by the pedal waves that go forward from the back part of the body to the front. [1,2,4]. The pedal wave on a horizontal surface has a wavelength of about 9 mm, a frequency of 1.3Hz, periodic time 0.8 1/s on average and the propagation speed of 17.4 mm/s is about 6 times of movement velocity of a snail.



(a) Picture of ventral surface



(b) Schematic illustration of ventral surface

Figure 6. Picture and illustration of ventral surface

The observation of marks put on the ventral surface of the foot was also done. One mark was put on the center of the ventral surface and another was put on the side. In about 1 mm width of the periphery of the ventral surface as shown in fig.6 (b), the pedal waves do not generate and that part moves smoothly with the snail's locomotion. It is considered that this part plays the role to hold the whole body on the surface (namely sucker effect). According to the enlarged observation of this part, it has many wrinkles in the radial direction that have a kind of suction action in the peripheral water or mucus. In fact, it is observed that peripheral water of the snail is sucked in when a snail is set on a wetted surface.

Figs.7, 8 and 9 illustrate the relationship between time and moving distance of a snail under each surface condition obtained by the observation of the mark's motion on the ventral surface of the foot. As shown in the figures, the mark moves forward when pedal waves pass through and that part of the ventral surface of the foot at which the pedal wave does not pass through is stopped. Fig.10 shows the locomotion characteristics in each surface condition (horizontal, vertical and reverse on ceiling) obtained from figs.7, 8 and 9. As seen in the figure, when a snail climbs on the wall (vertical condition), the velocities of movement and the pedal wave are most slow and the amplitude of the pedal wave is smallest. The velocities of movement and the pedal wave for horizontal and reverse motion on a ceiling are almost the same, and the amplitude is different.

Next, it was examined how a mark put on the ventral surface changes during the passage of a pedal wave. The mark begins to move slowly at the beginning of the passage of the pedal wave and is accelerated gradually. After reaching a constant speed (pedal wave speed), it decelerates again slowly. On the other hand, the color of the skin is darkened when the snail's muscle contracts. From this observation, it is considered that the pedal wave is a shrunk shape as a whole because the pedal wave is dark in comparison with the color of other parts of the ventral surface and it

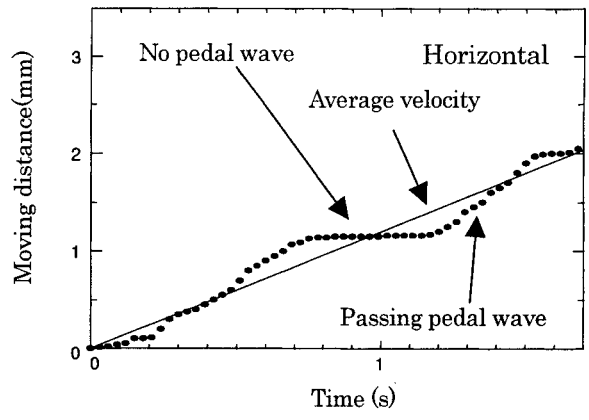


Figure 7. Relationship between time and moving distance

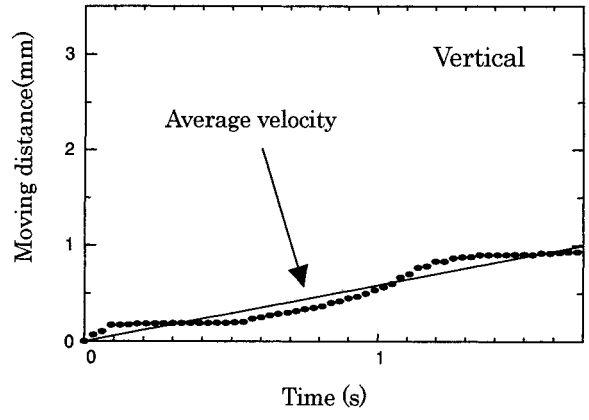


Figure 8. Relationship between time and moving distance

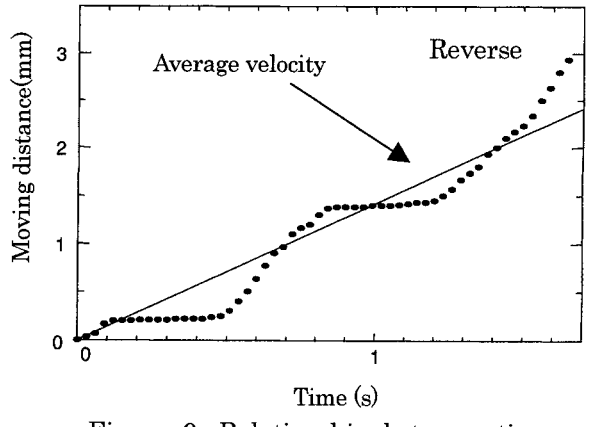


Figure 9. Relationship between time and moving distance

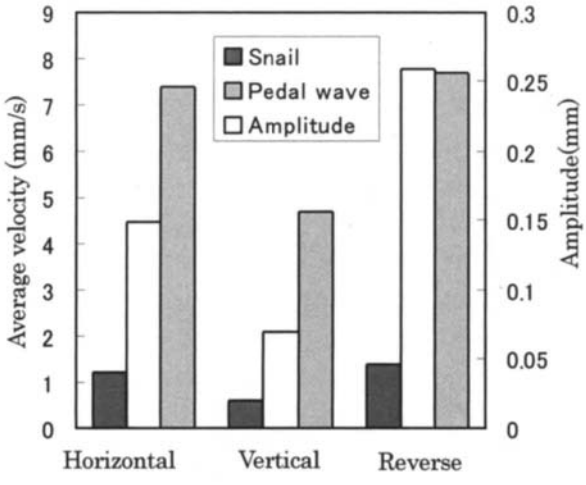


Figure 10. Kinetic characteristics of snail

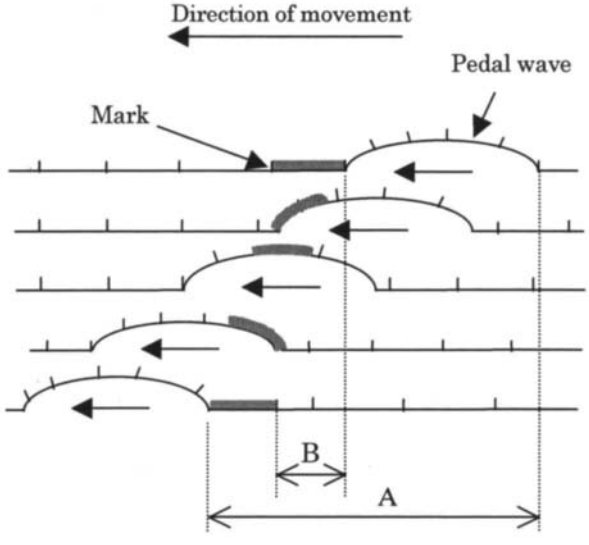


Figure 11. Locomotion of pedal area

looks like dark stripes run forward across the ventral surface. Lissman [1] suggested that a pedal wave is a concaved shape. From these results, the locomotion mechanism of a snail can be illustrated as shown in Fig. 11. When the pedal wave does not pass through, the ventral surface is elongated and attached to the substratum. When the pedal wave passes through, the ventral surface begins to contract and achieves the same speed at that of the pedal wave. Based on this mechanism of one pedal wave's passing, the progress distance of the snail is B in the figure. That is to say, snails can go forward extremely smoothly by very fine peristaltic movement of the ventral surface of the foot through viscous secreted mucus.

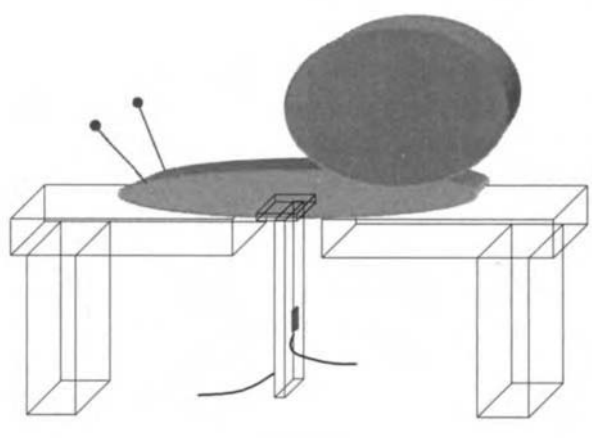


Figure 12. Schematic illustration of driving force measurement

4.2. Driving force of a snail

In order to examine how snails generate the driving force at the ventral surface, the driving force was measured by the method shown in Fig. 12. In the figure, the force is measured by a leaf spring type cantilever in which a small plastic plate (width=5mm, length=10mm, thickness=1mm) is attached on the upper free end. Fig.13 shows an example of the measured driving force variation for one pass in the horizontal locomotion. Snails generate the driving force at the anterior part of sole and generate resistance force in the posterior part. Namely, the posterior part from shell is only dragged. Here, the border between anterior and posterior is near to the shell. The fine vibrations in the variation curve correspond to pedal waves. Fig.14 shows the driving strength and the resistance at each surface locomotion. As shown in the figure, those strengths are largest in

horizontal locomotion and are smallest in the reverse one. The secreted viscous mucus is a non-Newtonian fluid (Bingham fluid) as mentioned before. Therefore, the driving force relates very closely to the film thickness of mucus decided in relation to the weight of a snail and atmospheric pressure. The weight of a snail acts most strongly in horizontal locomotion. Consequently, it is considered that the driving and resistance force in horizontal locomotion is largest.

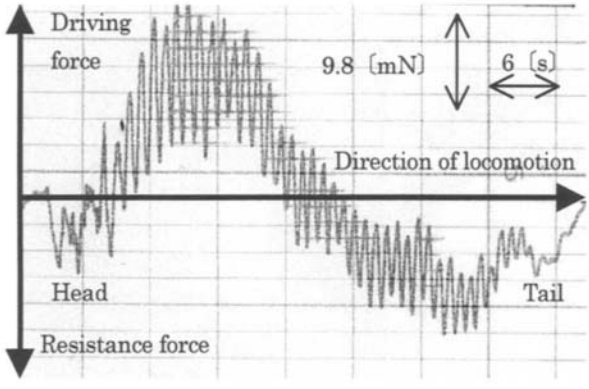


Figure 13. Driving force of snail

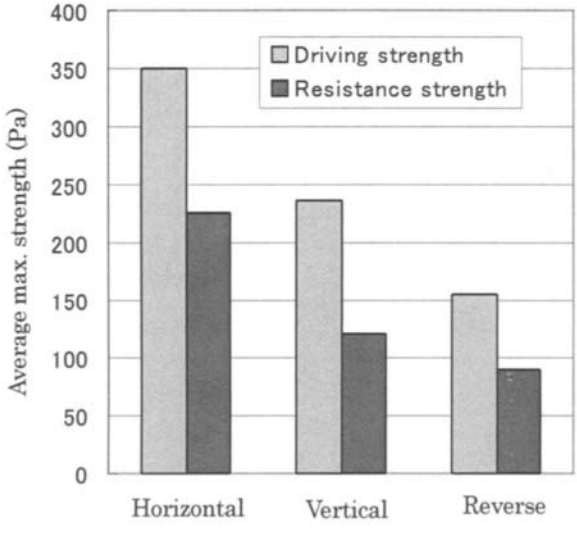


Figure 14. Comparison with three situations

Fig.15 shows the mucus film thickness in the anterior and posterior part of a snail's ventral surface during locomotion and after crawl. As shown in the figure, the mucus film thickness of the anterior part is thinner than that in the posterior part in every surface situation, because the driving force is generated in the anterior part.

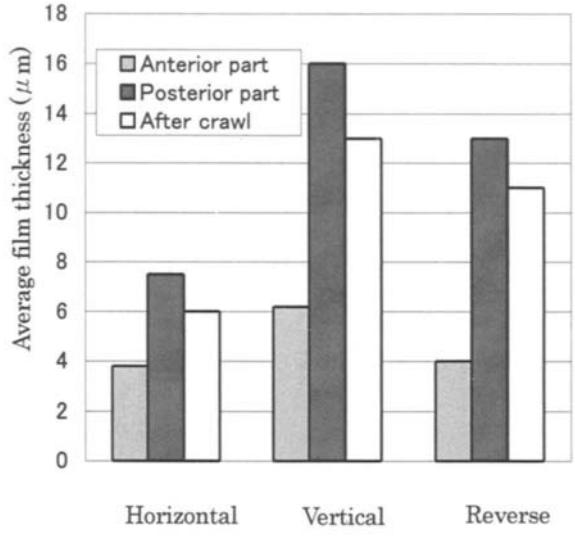


Figure 15. Film thickness of pedal mucus during crawling

5. MUCUS FLOW IN VENTRAL SURFACE AND PERIPHERAL

Fig.16 illustrates the mucus flow in the ventral surface and the periphery of the ventral surface of the foot obtained by the visualized mucus in which blue ink was injected. It was observed that much viscous mucus was mainly discharged from the mouth and some is also discharged from the ventral surface itself. It is said that mucus is also discharged from the mantle. However, this could not be observed clearly. Mucus from the mouth is spread the periphery of the mouth. As illustrated in the figure, it was observed that colored mucus in the ventral surface flows forward using pedal wave at the same speed of pedal wave and mucus around the sole is mainly sucked in from the posterior part. In such a manner, snails circulate mucus and reuse

around the sole is mainly sucked in from the posterior part. In such a manner, snails circulate mucus and reuse it skillfully.

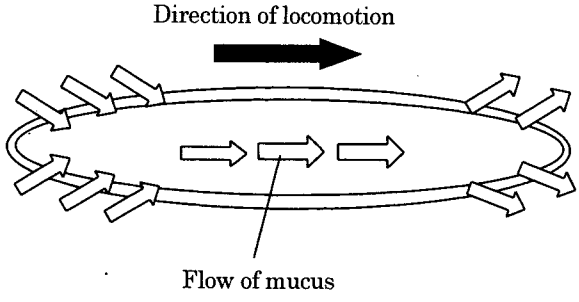


Figure 16. Flow of mucus in ventral surface

6. CONCLUSIONS

Snails (terrestrial gastropod) can move forward extremely smoothly on a surface (even wall and ceiling) through the secreted viscous mucus. Then, the locomotion characteristics of snails were examined experimentally from the tribological standpoint. Consideration of the results shows that;-

(1) The holding force of snails to a substratum does not depend on the ventral surface area.
(2) Ingredients of mucus in the shell consist of about 93% of water and the remaining 7% of glycoprotein.
(3) Secreted mucus of a snail is a kind of non-Newtonian fluid that has the characteristics of a visco-plasitic (Bingham) fluid because the viscosity decreases with an increase of shear rate and the shear stress has a plastic yield value and increases with shear rate.
(4) The basic locomotion mechanism of a snail's body itself is based on the very fine peristalitic motion of the ventral surface of the foot by the pedal wave that generates by the contraction and stretching of muscle
(5) Driving force is generated in the anterior part of the sole and the posterior part is only dragged.
(6) In about 1 mm width of the periphery of the ventral surface, the pedal waves do not generate and that part has many wrinkles in the radial direction that have a kind of suction action on peripheral water or mucus.
(7) Snails circulate secreted mucus around the sole and reuse it skillfully.

Acknowledgement

Authors thank Prof. K. Harada and Dr. K. Fujii of the Faculty of Pharmacy of Meijo University, for analysis and discussions of mucus by NMR.

REFERENCES

1. Lissman, H. W., *J. Exp. Biol.* Vol.21, (1945). pp.58-69.
2. Trueman, E.R., *The MOLLUSCA*, Vol.4, (1983). Academic Press. Inc.
3. Denny, M. R., *Nature London* Vol. 285, (1980). pp.160-161
4. Ikeuchi, K, and Tomita, N., *Trans. of JSME, B series*, Vol.60, 579. (1994). pp.3618-3622.
5. Nakamura K., Non-Newtonian hydrodynamics, (1997). Korona Press. Inc.

SESSION XII
TRIBOLOGICAL STUDIES OF DRY AND LUBRICATED SURFACES (3) – LUBRICATED

Chairman : Professor Y Berthier

Paper XII (i) **The Role of Tribology Research in the Development of Advanced Lubricants**

Paper XII (ii) **Ultrasonic Characterisation of an Interference Fit**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

The Role of Tribology Research in the Development of Advanced Lubricants

A. Jackson and M. N. Webster

ExxonMobil Research and Engineering Company, Corporate Strategic Research, Annandale, NJ, USA

Economic and operational drivers, such as efficiency, reliability, maintenance costs, productivity and longevity, encourage new technology development in engineering equipment. Advanced lubricants are engineering components that are designed to support these drivers and to provide benefits to equipment owners and users well beyond what is expected from commodity lubricants. Tribology contributes to the full spectrum of advanced lubricant development, from evaluating and guiding the design of additive and base stock molecules to the formulation of lubricants that offer superior performance in full-scale machinery. This paper discusses the role of tribology research in this process, and gives examples of its application in solving field problems, balancing conflicting lubrication requirements and creating new fluids and business opportunities.

1. INTRODUCTION

Lubrication research came into its own in the late 19th century as railways competed with each other for business. Track and loading gauges limited locomotive weight and boiler size and safety limited boiler pressure so friction reduction was the best way to increase profits. The automotive and aviation industries were in their infancy and required new types of lubricants to cope with combustion products. Steam ships vying for higher speeds and tonnage and industrial plant looking for higher productivity also challenged lubricant suppliers.

Simultaneously, the developing mineral oil industry was finding new ways to make lubricants along with fuels, offering opportunities for advancement in lubrication technology. Many of the lubricants developed for these applications would have been considered advanced in their day, but are now long obsolete.

During the 20th century, the demand for higher power to weight ratios led the move from coal-fired steam engines to diesel engines, from aviation piston engines to jet turbines, and encouraged the wide scale availability of electric power. Continuous improvements in engine technology have required increases in lubricant capabilities and offered opportunities for lubricant companies to differentiate their products. To stay in business, industrial productivity needs to continuously improve, which in turn produces opportunities for advanced lubricants to demonstrate an economic advantage.

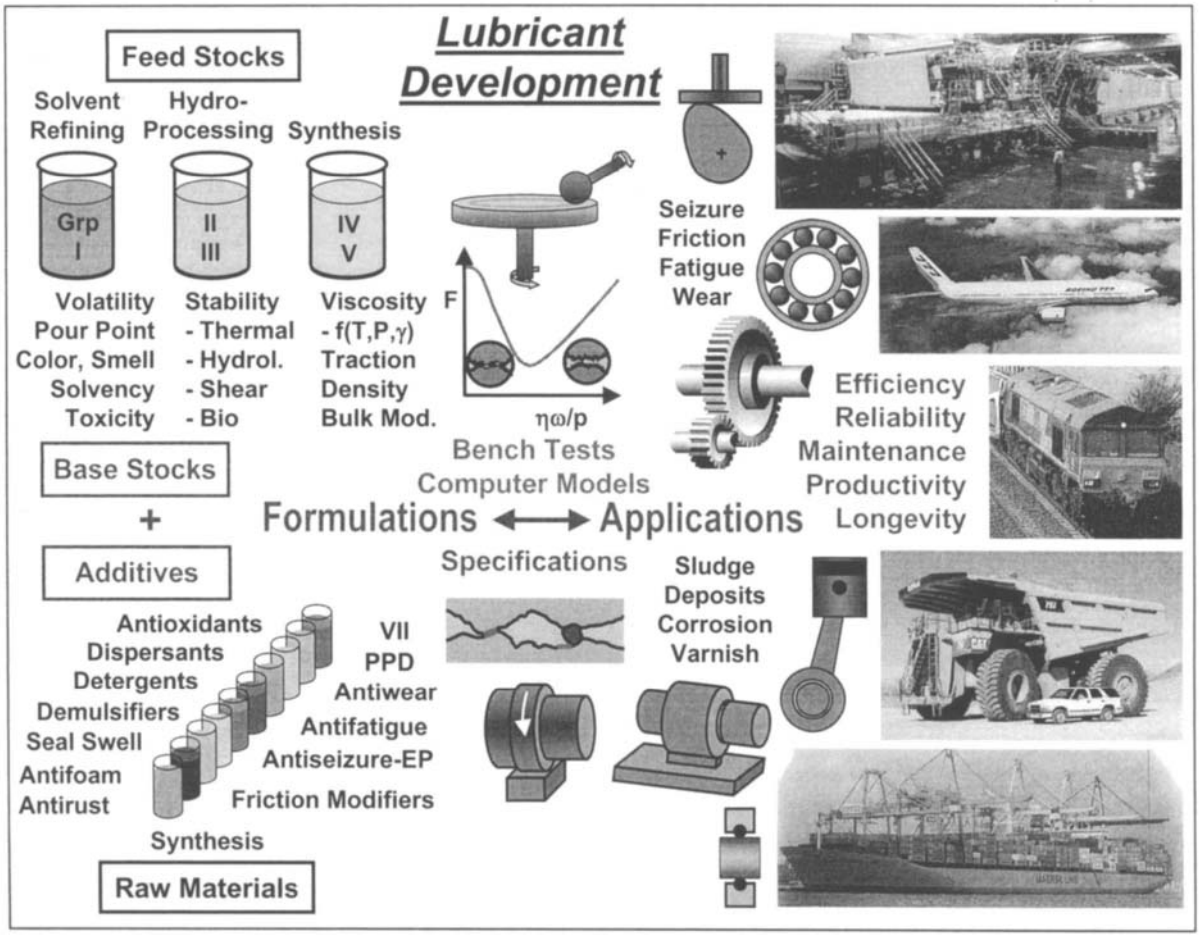
This presentation reviews the role that tribology research plays in the development of advanced lubricants and lubrication technology. The scope will be restricted to advances in large volume lubricants used in the automotive, industrial, marine, railroad and aviation industries, although it is recognised that there are many advances in tribology outside these fields.

2. THE LUBRICANT DEVELOPMENT AND FORMULATION PROCESS

It is instructive to consider the general process of lubricant development before tackling the topic of advanced lubricants. Figure 1 is a schematic of the process. Economic drivers in engineering equipment are efficiency, reliability, maintenance costs, productivity and longevity. The relative importance of these will depend on the user. For personal cars, reliability is likely to be top of the list - a car is not much use if it can't be driven. Fuel and maintenance costs are also important. For industrial equipment, all the above drivers will be important, but weightings will vary by industry and location.

Tribologists consider equipment in terms of the components where tribological interactions take place, such as cams, rolling bearings, gears, piston assemblies, hydrodynamic bearings and seals. Lubrication is intended to keep at bay gremlins like seizure, friction, fatigue, wear, sludge, carbon deposits, corrosion and varnish. Formulation design is targeted at a variety of applications and is often

Figure 1. Schematic of the lubricant development and formulation process.



driven by industry or customer specifications. Bench tests are developed to simulate conditions in components. To facilitate the formulation process, models are used to understand the type of lubrication involved, hydrodynamic, elastohydrodynamic, mixed or boundary and the relative roles of base stocks and additives in a given application.

A wide variety of base stocks and additives is available. Current terminology groups base stocks by API classification, with manufacturing type generalised as solvent refining, hydroprocessing and synthesis. Properties that base stocks bring to lubrication include solvency, low temperature fluidity, volatility, viscosity, traction and density. The stability of base stock properties can also vary with thermal stress, shear and exposure to water and contaminants. Additives include antioxidants, dispersants, detergents, corrosion inhibitors, VI-improvers, antiwear, anti-scuffing (EP) and friction modifiers. Properties of synthetic base stocks and additives can be varied widely by design through feed stock and manufacturing process selection.

3. PRESERVATION VERSUS TRIBOLOGY

Part of the effort of lubricant formulation has little to do with tribology and can be considered as supporting the "preservation" function of lubricants. This includes preservation of the lubricant itself and its properties over its projected lifetime as it is exposed to oxidation, combustion products and contaminants and to ensure it remains capable of functioning as a lubricant within the temperature range it is designed for. It includes preservation of the lubricant's acceptability to customers in areas such as colour, clarity, foaming, and toxicity. It also includes preservation of the components exposed to

the lubricant by protection against corrosion, sludge build-up, carbon deposits and varnish.

Base stock functions that fall into the preservation area include volatility, pour point, solvency, colour, smell, toxicity, and stability against thermal stress, hydrolysis and shear. Bio-stability or bio-degradability may be required for different applications. Additive functions that fall under preservation include oxidation inhibition, dispersancy, detergency, demulsification, seal swelling, anti-foam and corrosion inhibition.

The tribology role of lubricants includes the avoidance of seizure, the control of friction - usually but not always trying to reduce it, the minimisation of wear and slowing fatigue crack initiation and growth.

Base stock functions important to tribology include viscosity and its variation with temperature, pressure and shear rate, traction, viscoelasticity, and to lesser extents, density, bulk modulus and thermal properties.

Additive tribology functions are clearly antiwear, anti-fatigue, anti-seizure (also known as anti-scoring anti-scuffing and extreme pressure, EP) and friction modifiers. Rheology modifiers such as polymeric viscosity index improvers and pour point depressants can also impact the tribology function when temperature and shear rate are varied.

4. THE ROLE OF TRIBOLOGY RESEARCH IN LUBRICANT DEVELOPMENT

Having differentiated between the preservation and tribology functions of a lubricant, it is now possible to explore the extent to which tribology research can impact lubricant development.

Both base stock and additive tribological performance features are functions of their molecular structure and a key role for tribology research at the input to the component supply pipeline is to develop structure-performance relationships to guide processing and synthesis. This requires close co-operation between the organic chemist and the tribologist and entails a considerable amount of experimental tribological evaluation of model compounds as well as commercially viable components.

Since the costs of creating large samples of new lubricant components, and testing in full scale equipment are both very high, it is clearly impractical to test all but a few formulations in equipment. A key role of tribology research is to develop small-scale tests which mimic equipment operation and to develop a correlation between these tests and equipment performance. Bench tests must also link to the economic drivers of efficiency, reliability, maintenance cost and component life and ideally should be supplemented by macro-, micro- and molecular-scale models of performance.

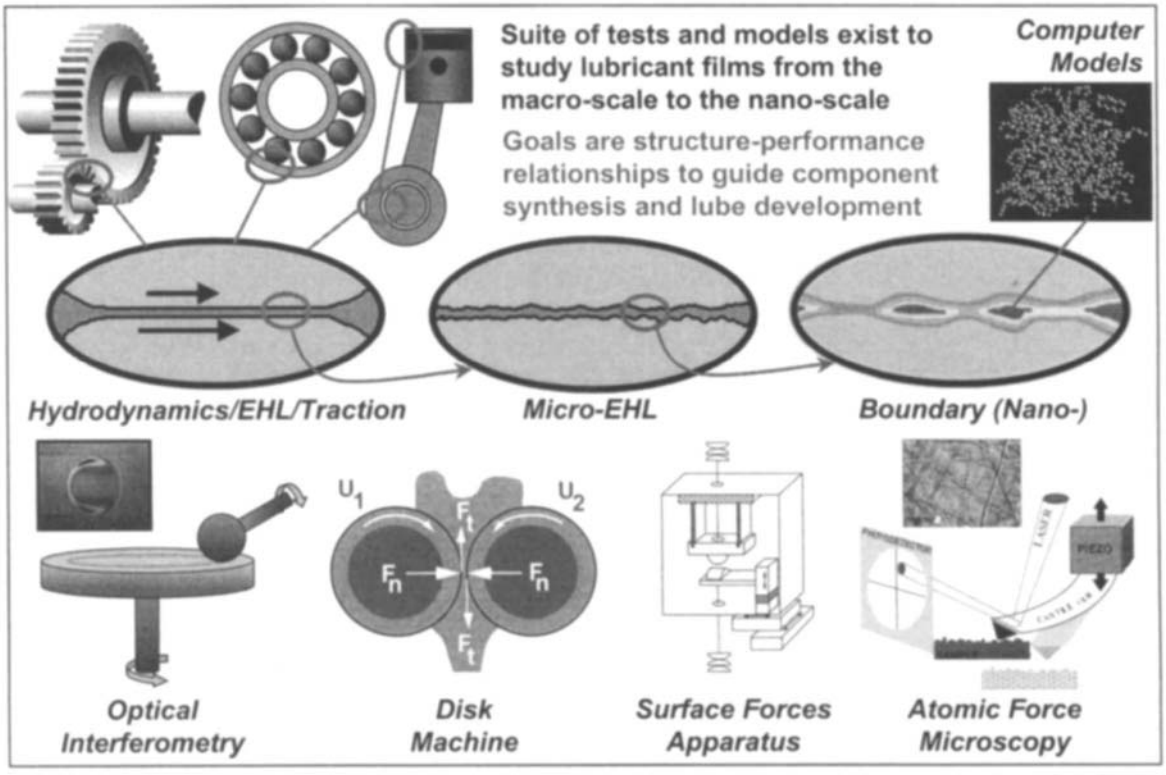
Linking equipment performance to lubricant molecular structure clearly requires a broad range of disciplines. Tribology research can best be conducted by an interdisciplinary team of mechanical engineers, materials and surface scientists, mathematicians and modellers, physicists, chemists and chemical engineers. To maximise the chances of developing truly step-out lubricants, the team needs to be closely connected to the lubricant component manufacturers, the lubricant formulators and to the equipment designers and users. Lubricants are critical engineering components which, as a minimum, should be considered in the design process. Ideally they should be allowed to influence design in order to maximise the economic advantages of new equipment. In other words equipment and lubricant design should advance together.

5. MACHINES TO MOLECULES

Figure 2 looks more closely at the tribological processes involved in surface protection under rolling and sliding. The first and preferred line of defence against surface contact which can lead to high wear, high friction and eventual failure is a hydrodynamic or elastohydrodynamic (EHD) film. Reasonable theoretical prediction of film thickness has been possible for over 50 years and has been validated experimentally using capacitance for the relatively thick hydrodynamic films and optical interference for EHD films. These techniques have allowed lubricant manufacturers to quantify the ability of their lubricants to generate protective hydrodynamic and EHD films in machinery. This largely depends on viscosity and its variation with temperature, pressure and shear rate.

Early analytical models have been replaced by finite element models, which allow more variables to be considered and a greater number of cases to be investigated. Transient models can now study the passage of rough surfaces through an EHD contact whilst simultaneously including thermal effects

Figure 2. From machines to molecules.



and a range of rheological models of lubricant behaviour. Predictions of contact pressure variation and traction forces in micro-EHL are proving valuable in understanding friction and wear in rolling bearings and gears and such models are now being included in full scale models of operating equipment. Experimental methods such as disc machines and ball on plate devices have been developed to measure traction forces.

At the molecular scale, the surface force apparatus is being used to study lubricant behaviour in very close approach of surfaces. Atomic force microscopy can study local frictional variation due to additive deposition on surfaces in addition to giving detailed topographical maps. Nano-indenters can probe the elastic and plastic properties of surface films. These techniques are enhancing the understanding of boundary lubrication and supplement results from the multitude of more widely known boundary friction and wear testers, such as 4-ball, ball-on-plate and ring-on-block.

Finally computer models of molecular behaviour in lubricated contacts are starting to shed light on the long sought after relationship between structure and performance. Whilst there are many approaches that can be used to model the behaviour of liquids, one which has great appeal is the use of molecular dynamics. Molecular dynamics simulations of hydrocarbons treat carbon atoms and associated hydrogen atoms as spheres with rules on the behaviour of bonds within a molecule and on how the spheres from different molecules can approach each other. Newtonian mechanics is used to predict the motion of molecules in time steps of the order of femto-seconds and statistics are used to compute averages of physical properties from the simulations.

Computing power is now sufficient to allow the modelling of lubricant range molecules in large enough quantities for simulations to be meaningful. Molecular dynamics models have been used recently to model viscosity and its variation with temperature, pressure and shear rate. Absolute prediction of viscosity is difficult, but relative changes in viscosity with temperature and pressure have been found to be quite accurate, and have given insight into structural features controlling this behaviour. Figures 3 and 4 show examples of

molecular dynamics prediction of viscosity and its variation with shear rate and pressure.

Figure 3. Equilibrium molecular dynamics (EMD) prediction of relative viscosity of octane isomers.

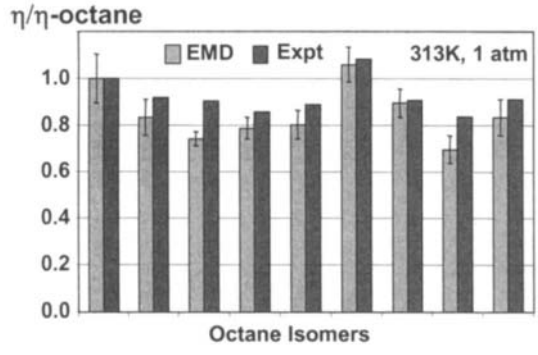
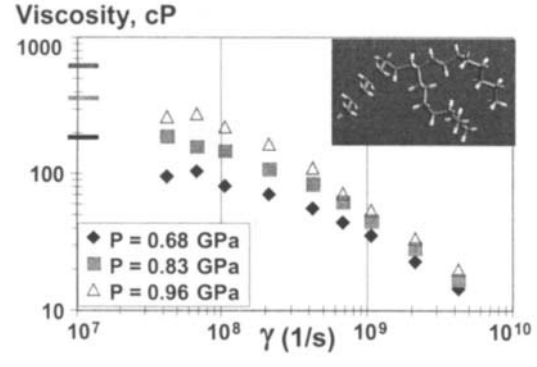


Figure 4. Non-equilibrium molecular dynamics prediction of viscosity variation with shear rate at various pressures (1).



A major challenge is how to handle the large number of isomers of hydrocarbons that are possible in lubricant range materials. These range from a billion at C30 to a billion billion at C50. Few of these have relevance as lubricants and judicious selection of representative isomers and groups of isomers is needed. Experimental validation of computed results is also problematic since there are few pure isomers available in this range. Nevertheless, the simulations done so far have been very helpful in guiding synthesis.

6. LUBRICANT COMPONENTS

The opportunity to develop advanced lubricants depends on the availability of a wide variety of components with different tribological characteristics and on the skill of the formulator in assembling them into working lubricants. The variety of synthetic base stocks is limited only by the availability of feed stocks and intermediates, and by the imagination of the organic chemist. There are three major families of synthetic lubricants, hydrocarbons, polyglycols and esters, and a variety of other families of lesser importance. Within each family there are different classes and within each class there are fluids with a wide variety of tribological and other properties. There is no shortage of tools to work with. The major issue is the high cost of many of these fluids and the ability to demonstrate their value in engineering applications.

The range of molecular structures available from hydrocarbon synthesis parallels those available in mineral oils. Group I mineral oils are manufactured by separation processes, vacuum distillation for viscosity, solvent separation of aromatics and deasphalting for higher viscosity bright stocks. Wax removal is handled either by low temperature solvent separation or catalytic dewaxing. The resulting mixture is a wide range of acceptable molecules but with physical properties and performance which can only be an average for the mixture. There are few opportunities to optimise properties of relevance to tribology, other than viscosity. Conventional hydroprocessing for Group II products saturates aromatics and reduces sulfur levels, increasing oxidative stability but doing little to modify tribological properties. Severe hydroprocessing can open up ring structures and increase viscosity index, VI. Hydroisomerisation of slack wax from mineral oil solvent dewaxing creates a high VI product but can not match the low temperature fluidity of synthetic hydrocarbons such as polydecenes.

Esters, polyglycols, fluorocarbons and other fluid families with valuable tribological properties are not available in mineral oils. Functionalization of synthetic fluids expands the opportunities for performance enhancement and can bridge the gap between base stocks and additives. There is an equally wide variety of chemistry used in additive synthesis.

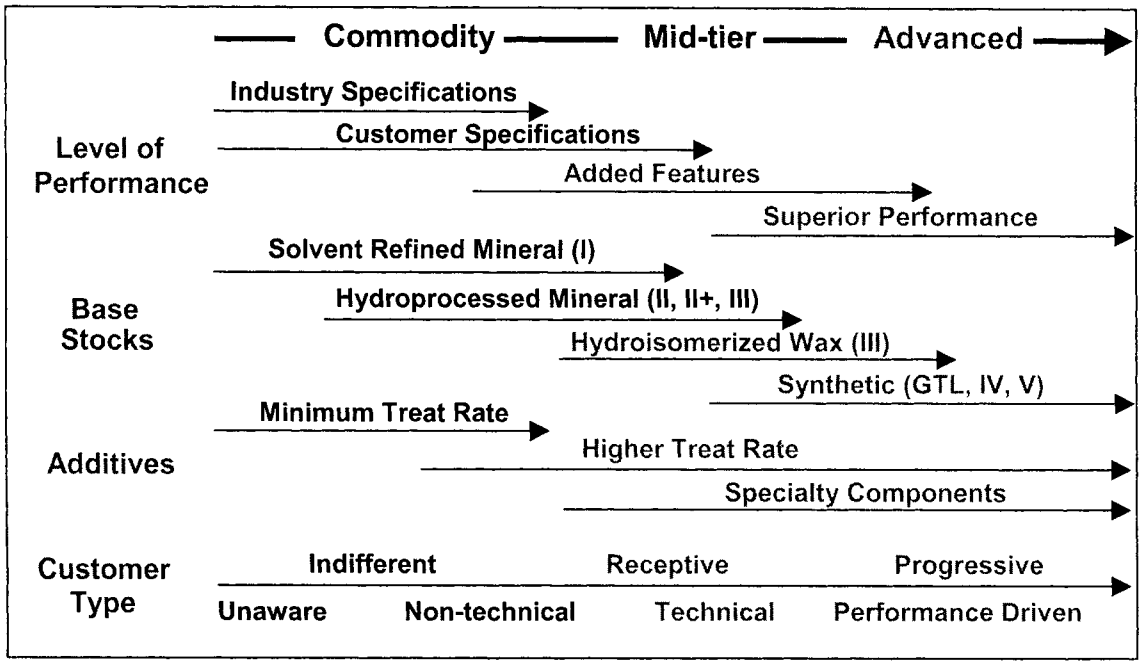


Figure 5. Characteristics of commodity, mid-tier and advanced lubricants

7. ADVANCED LUBRICANTS

Having explored the wide range of components available to lubricant formulators and the tribological investigations that can characterise and differentiate performance, it is time to look at what might be called advanced lubricants. Advanced is a term that only applies to a formulation for a limited time as today's advanced products become tomorrow's commodities. Continuous development of new technology is needed to satisfy the increasing demands of equipment.

Marketing frequently has a good, better, best approach to products and many names are used to describe each level. For lubricants, the words commodity, mid-tier and advanced are often used. Characteristics of each are shown in Figure 5. Commodity products generally meet industry or customer specifications with little quality give-away. They are formulated with Group I base stocks and the minimum acceptable level of additives. In certain markets, hydroprocessed stocks may be needed to meet specifications. Mid-tier products will have some added features, gained from higher additive treat levels or the use of specialty additives, perhaps with hydroprocessed base stocks. To be called advanced, products must have demonstrably superior performance in many areas, which comes from higher additive treat levels, specialty additives and the use of high end Group III or synthetic base stocks.

The term premium is often used to describe both upper level product types. Premium is a meaningful word since it recognises that the superior performance of such products can only be made commercially available if the additional costs of research and higher quality components can be supported by a premium price. This in turn requires that the customer can recognise and measure the economic or competitive advantage to be gained from the use of such a product. Customers must at least be receptive to technical advances and those that can understand the value are often technical and driven by performance.

Nature did the human race a great service by providing mineral oils as a by-product of fuels refining but their low cost may be limiting progress by discouraging the commercial development of non-mineral fluids that may bring greater overall engineering benefits.

Opportunities to develop lubricants with superior performance include:

1. Improve energy efficiency and reduce CO_2 emissions through viscosity optimisation in hydrodynamic lubrication, traction reduction in elastohydrodynamic lubrication, and boundary friction reduction using surface-active friction modifying additives.
2. Reduce maintenance costs through longer component lives, longer oil life, and tolerance of soot and contaminants.
3. Reduce downtime by eliminating component failure and seizure.
4. Enable new designs using novel lubricant features, novel lubrication approaches, and lubricants designed to work with new materials and coatings.
5. Increase design ratings through higher loads, higher speeds, and increased productivity.
6. Enhance environmental performance through exhaust after-treatment system compatibility, the use of recyclable fluids, and by using bio-friendly fluids.
7. Extend the temperature range of operation by improved low temperature rheology, reduced dependence of viscosity on temperature (higher VI), higher oxidative stability, and lower volatility.
8. Increase safety margins through reduced toxicity and increased fire resistance.

Satisfying many needs can often be a balancing act since lubricant properties which bring an advantage in one machine component may bring a disadvantage to another component and both might be found in the same machine. Also molecular structural features that bring a performance benefit in one area may bring a reduced performance in another area.

Four examples follow which demonstrate how tribologists and formulators can work together in the development of advanced products.

8. TRIBOLOGY RESEARCH IN THE DEVELOPMENT OF ADVANCED LUBRICANTS – FOUR EXAMPLES

8.1. Balancing Performance – Fuel Efficiency and Engine Wear Protection

Lubrication engineers have used the Stribeck curve for over a hundred years to describe the friction behaviour of lubricated components and machines as a function of operating conditions. The characteristic V-shaped curve can clearly be seen in a plot of fuel consumption against viscosity obtained in an ASTM Sequence VI fuel economy engine test used in a screening mode, shown in Figure 6.

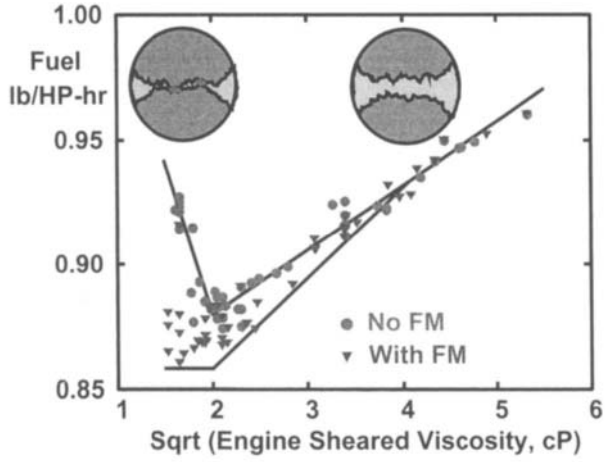


Figure 6. Engine Stribeck Curve. Upper lines are least-squares fits through the data on oils with no friction modifier (FM) at viscosities above and below 4cP. Lower lines are lower bounds for oils with friction modifier.

The engine is run at constant speed and load but at three different temperatures and with various lubricants so the speed times viscosity divided by load term for the x axis reduces to viscosity alone. The viscosity has been corrected for non-Newtonian behaviour deduced from a different series of experiments run on the same engine. A clear minimum is seen at a viscosity of 4 cP.

The benefit of various friction modifiers (FM) can be seen, especially at low viscosity, but the effect disappears at about 15 cP as hydrodynamic lubrication takes over throughout the engine. It is clear that the amount of friction reduction varies and that some friction modifiers are not effective in the engine even though they demonstrated friction reduction in bench tests. Part of the problem is that friction modifiers must compete for the surface to be effective and often, the competition is against antiwear additives, which do not necessarily have low friction. The inverse is that friction modifiers can displace antiwear additives and that can result

in wear when operating near the minimum on the Stribeck curve.

Using bench testing and engine testing an advanced synthetic engine oil formula was developed which gives very low friction as well as excellent wear protection and has additional credits for oxidation inhibition. This was combined with a novel approach to VI and hydrodynamic film thickness boosting which had been extensively tested in drag, road and circuit racing. This advanced synthetic engine oil was released on the market this year.

8.2. New Fluid Development – High Traction and Low Traction

A very important lubricant base stock property that depends on molecular structure is the traction force generated in a rolling-sliding EHL contact. The variation of maximum shear stress developed by various fluids as a function of contact pressure is shown in Figure 7. Paraffinic mineral oils, which contain a mix of hydrocarbon structures, not surprisingly fall into the middle of the group. Synthesis can generate fluids with very high traction and others with very low traction.

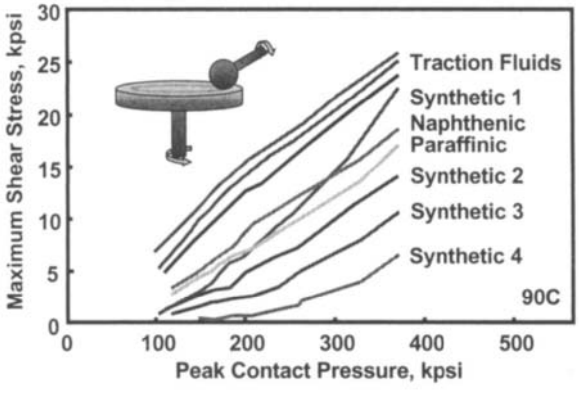


Figure 7. Maximum shear stress versus contact pressure for a variety of fluids.

Low traction fluids are particularly beneficial for all types of gears, where the mesh contact has varying levels of sliding superimposed on rolling. Mesh losses can be reduced by 30-70% depending on the fluid used and the operating conditions and this can amount to energy savings of 1-3% in multiple ratio steel gears to 10% or more in high ratio worm gears compared with mineral oils. The reduced traction also results in less heat generation, higher scuffing loads and longer pitting lives due to reduced tangential forces on the surface.

High traction fluids are enablers for rolling contact traction drives. The power that can be transmitted depends on the traction force carried by the fluid. When traction drives are incorporated into automotive transmissions, the continuously variable ratio allows the road speed to be disconnected from the engine speed. For a given vehicle power requirement, this allows drive-by-wire optimisation of fuel efficiency or of emissions by forcing the engine through its maximum efficiency or minimum emissions operating regions. Resulting efficiency improvements have been reported in the 15-20% range.

For both applications, good low temperature fluidity is required, and this can be a challenge. Some low traction fluids are linear structures that can solidify at low temperature. Most fluids with high traction have a low viscosity index and their viscosity becomes too high to flow at low temperatures. The challenge for the synthesis chemist and the tribologist is to develop fluids with the desired high or low traction properties and with good low temperature performance.

8.3. Problem Solving – Micropitting in Hardened and Ground Gears

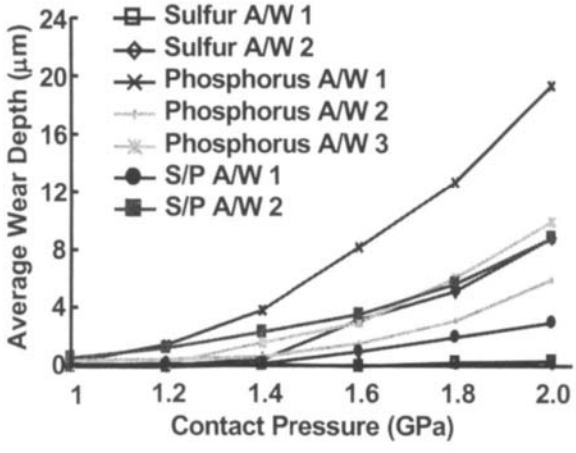
New applications often place demands on mechanical systems that can create unexpected problems. In response to the prospect of a free and clean energy source, along with tax incentives, the number of wind turbine generator systems has rapidly increased. For highest efficiency, the slow speed of the blade rotor needs to be increased for input to the generator, using a gearbox.

The rotor shaft, gearbox and generator are housed in the nacelle of the wind turbine. The gearbox needs to be small to save weight for structural reasons, and must need little maintenance due to difficult access. Case hardened gears are almost always used. Unfortunately many of the early designs suffered gear failures. These were partly due to a lack of understanding of the gearbox loads, but many gears failed by a mechanism identified as micropitting, in far greater proportion than was observed in other applications. While case hardening increases spalling life, on which gear life is often calculated, it can also result in micropitting. It was discovered that some lubricants appeared to perform better than others

despite being developed around the same specifications.

A roller-disc machine was used to reproduce the conditions in typical wind turbine gearboxes (2). It was initially considered that micropitting was governed only by the degree of asperity contact but it was confirmed that lubricant components could also have an impact, as shown in Figure 8. Lubricant components that help reduce micropitting have been identified. Advanced gear oils have been

Figure 8. Additive effects on micropitting. (A/W = antiwear, S/P= Sulfur/Phosphorus)



formulated which reduce the risk of micropitting while offering protection against scuffing. These oils have been developed in collaboration with wind turbine and gearbox manufacturers. The effort is ongoing as knowledge of the demands that wind turbines place on gear systems improves.

8.4. New Opportunities – Environmentally Acceptable Open Gear Lubes

Open gears, such as those used in cement mills, have been lubricated with asphalt gear oils diluted with chlorinated solvent to help flow. Such gear lubricants adhere to the surface, which allows low feed rates while producing a relatively thick EHL film. Environmental concerns now limit the use of chlorinated solvents. Replacement organic solvents could be used but often their volatility created a fire hazard. As a result several replacement options have been investigated.

EHL analysis of typical open gears confirmed that the combination of slow speed and relatively rough surface finish required the use of high viscosity lubricants. A simple replacement for asphalt is a semi-fluid grease since much of the existing lubricant delivery system can be used. However, the viscosity of the base stock that can be used is limited to the conventional paraffinic mineral oil range, well below the requirement for open gears. Combining this with the use of a solid lubricant such as graphite allows the build up of a film over time, and can protect the gear from wear.

A significant advance is the use of extremely high viscosity (e.g. ISO 46,000) synthetic base stocks. These remain sufficiently fluid at low temperature to be pumped into the gear mesh and stay long enough on the gear to allow low feed rates. This approach has the added advantage that the gear teeth remain clean, allowing periodic inspections. However, the system is a once-through one, so there is still waste lubricant to deal with.

An alternative solution is to convert the open gear system into an enclosed lubricant system. This is possible by adding suitable shrouding around the gears so that the oil returns to a sump via gravity. It is necessary to include a filter to remove the cement dust and foreign material. This has the benefit of removing the root cause for gear wear in these applications, abrasion, and allows the viscosity of the oil to be reduced to the point where it is possible to re-circulate it. There is an investment cost for the conversion to an enclosed system, but this approach has been found to be the cleanest and has proven to be very reliable.

9. ADVANCED LUBRICANTS CURRENTLY IN COMMERCIAL USE

Advanced lubricants formulated with high performance base stocks and specialty additives are available for almost all automobile, industrial, aviation and marine applications. The formulations of these products are continuously upgraded to keep the performance ahead of commodity and mid-tier products. The following is a list of commercially available synthetic lubricants:

- Automotive lubricants
  - Passenger car engine oils
  - Commercial vehicle engine oils
  - Manual transmission fluids
  - Automatic transmission fluids
  - Rear axle oils

  - Greases
  - Traction fluids for IVTs
- Industrial lubricants
  - Circulating oils
  - Gear oils
  - Open gear oils
  - Hydraulic fluids
  - Fire resistant hydraulic oils
  - Biodegradable hydraulic oils
  - Paper machine oils
  - Gas compressor lubricants
  - Compressor lubricants for HFC refrigerants
  - Greases
  - Biodegradable greases
  - Stationary gas engine oils
  - Stationary gas turbine oils
  - Metalworking fluids
- Aviation lubricants
  - Gas turbine oils
  - Hydraulic fluids
  - Greases
- Synthetic marine oils
  - High speed diesel engine oils
  - Hydraulic oils
  - Gear oils

10. CONCLUSIONS

The following lists some conclusions from this presentation.

- Lubricant formulation is a complex business
  - Many applications, changing needs, multiple components to choose from.
- Lubricant formulations have two main functions:
  - Preservation - maintain lubricant performance over its life, prevent corrosion and keep systems clean and sludge free.
  - Tribology - enhance efficiency, reduce wear, slow fatigue crack growth rate, eliminate seizure and increase productivity.
- Tribology and lubricant component research enables the development of advanced lubricants that provide an economic or operational benefit:
  - Improved energy efficiency, reduced CO_2 emissions.
  - Reduced maintenance costs and machine downtime.
  - Increased design ratings, operation over a wider temperature range.
  - Enhanced environmental performance and safety.
- Equipment and customer demands continue to become more severe:
  - New lubricants can enable new designs and novel solutions.
  - Equipment and lubricant designers need to work together.
- Today's advanced lubricant will be tomorrow's commodity product:
  - Continuous improvement in lubricant technology is needed.

11. ACKNOWLEDGEMENT

The authors would like to thank ExxonMobil Research and Engineering Company for permission to prepare, present and publish this paper.

REFERENCES

1. McCabe, C., Cui, S., Cummings, P.,T., Gordon, P.,A. and Saeger, R.,B., "Examining the Rheology of 9-octylheptadecane to Giga-Pascal Pressures." Journal of Chemical Physics Volume 114 (4) p. 1887-1891 (2001)
2. Cardis, A. B. and Webster, M. N.,"Gear Oil Micropitting Evaluation" AGMA Technical Paper 99FTM4, presented at the October 1999 AGMA Fall Technical Meeting, Denver Colorado.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Ultrasonic Characterisation of an Interference Fit

R. Lewis, M.B. Marshall, R.S. Dwyer-Joyce

Department of Mechanical Engineering, University of Sheffield, UK

When components such as bearings or gears are pressed onto a shaft the resulting interference induces a pressure at the interface. The size of this pressure is important as many components fail because fatigue initiates from press-fit stress concentrations. The aim of this work was to develop ultrasound as a tool to non-destructively determine press-fit contact pressures.

A press-fit interface is essentially a rough surface contact and consequently behaves as a spring. If the pressure is high there are few air gaps; so it is very stiff and allows transmission of an ultrasonic wave. If the pressure is low then interface stiffness is lower and most ultrasound is reflected.

A *spring model* was used to determine maps of contact stiffness from interference fit ultrasonic reflection data. A calibration procedure was then used to determine the pressure.

In the central part of the fits, the pressures agreed closely with predictions obtained by the Lamé equations. At the edges, however, stress concentrations were observed.

1. INTRODUCTION

Interference fits, for example where a bush is shrunk or pressed onto a shaft, are commonly used to join machine components. They represent a flexible and cost effective method for joining components having cylindrical symmetry.

The strength of the interference fit assembly depends on the shaft and bush dimensions and must be sufficient to withstand the force or torque reached during normal service.

For example, railway wheels are press or shrink fitted onto axles. Railway axles occasionally fail by fatigue resulting from the alternating rotating bending loading. The site of the fatigue crack initiation is commonly at the location of the press fit. Clearly a method for measuring the contact stress associated with press-fits would prove useful both in component inspection and product development. This paper describes such a method based on interpreting the reflection of ultrasound at the interface to determine the contact pressure.

2. BACKGROUND

2.1.Interface Pressure Calculations

The established analytical solution relating the geometrical overlap in an interference fit to the resulting contact pressure is that derived by Lamé (see for example [1]). The analysis is based on equilibrium of stress and displacement in thick infinitely long cylinders. Interface pressure, p, is given by:

$$p = \frac{E\delta\left(r_2^2 - r_1^2\right)}{4r_1r_2^2} \tag{1}$$

where r_1 and r_2 are the internal and external diameters of the bush respectively, δ is the diametral interference (as defined in Figure 1) and E is Young's modulus.

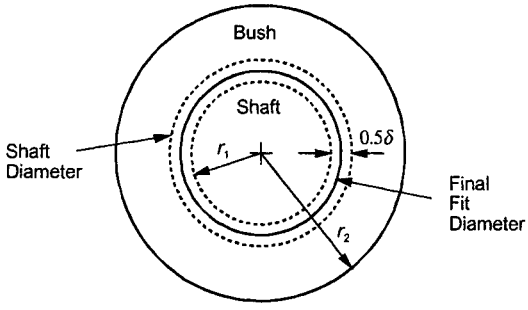


Figure 1. A Simple Interference Fit (not to scale)

Due to its assumptions, however, the Lamé solution is unable to predict the stress concentrations at the ends of the joint nor is it suitable for a bush of changing section. For more complex geometries finite or boundary element analysis, are used to determine contact stress distribution (see for example [2] and [3]).

2.2.Ultrasonic Reflection from an Interface

When a wave of ultrasound is incident on an incomplete interface between two materials it will be partially reflected (as shown in Figure 2). The proportion of the amplitude of the incident wave that is reflected is known as the reflection coefficient.

If the wavelength of the ultrasound is large compared with the size of the air gaps then the response of the interface can be determined using a quasi-static spring model [4]:

$$|R| = \frac{1}{\sqrt{1+(2K/\omega z)}} \qquad (2)$$

where ω is the angular frequency (= $2\pi f$) of the ultrasound wave, z is the acoustic impedance (the product of the wave speed and density) of the bulk material and K is the stiffness of the interface.

It has been demonstrated that typical engineering rough surface contacts were sensitive to ultrasound frequencies in the regime 1 – 50MHz; and that the reflection of ultrasound could be used to determine information about the nature of the contact at an interface [5].

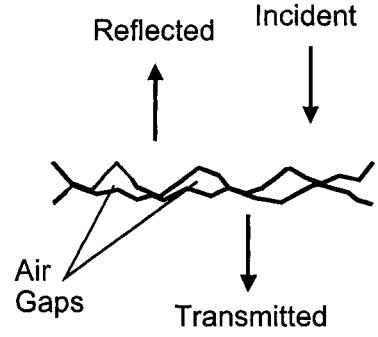


Figure 2. Partial Reflection of Ultrasound at a Rough Surface Contact

The stiffness of a dry interface, K (expressed per unit area), is the load required to cause unit approach of the mean lines of the surface roughness [6]:

$$K = -\frac{\mathrm{d}p}{\mathrm{d}u} \qquad (3)$$

where p is the nominal contact pressure at the interface (i.e. the applied load divided by the apparent geometrical contact area) and u is the separation of the mean lines of the roughness.

Thus, as the load on an interface is increased, the surfaces are pressed into closer conformity and the stiffness increases. The stiffness will vary from zero, when there is very small contact, to infinity, when the

surfaces are completely conformed.

The stiffness of a rough surface interface depends on the number, size and spacing of the regions of contact. A surface with a certain number and size of distributed contact regions would have a higher stiffness if those regions were closely packed. Therefore, there is no unique relationship between stiffness and contact pressure [7]. However, an independent calibration experiment can be performed to determine the relationship between the contact pressure and interfacial stiffness (for a given pair of rough surfaces). Measurement of the reflection coefficient therefore provides a method to determine the contact pressure [8].

When a bush is pressed or shrunk onto a shaft, the surfaces either side of the joint will be loaded together. Because these surfaces are inherently rough the interface will reflect ultrasound to some extent and the will be amenable to the above analysis.

3. EXPERIMENTAL DETAILS

3.1.Ultrasonic Scanning Apparatus

A spherical focusing 10MHz transducer is mounted in a water bath (in this case a scanning tank) such that the ultrasonic wave is focused at the shaft/bush interface. Ultrasonic waves are generated and received by an ultrasonic pulse-receiver (UPR). The reflected signals are captured on a digital oscilloscope and passed to a PC for processing. A schematic illustrating the set-up of the equipment is shown in Figure 3.

The transducer is positioned using an x, y stepper controlled by the PC. This enables ultrasonic readings to be taken using the transducer over an area at prescribed intervals (i.e. a C-scan). The dimensions and resolution of the scan can be varied and selected according to specimen geometry and the accuracy required. Focusing height adjustment of the transducer is achieved manually.

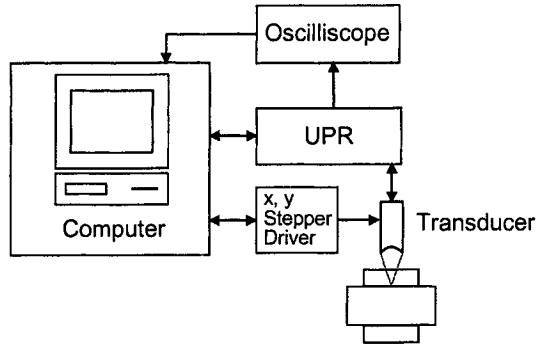


Figure 3. Ultrasonic Apparatus

3.2.Interference Fit Specimens

Each interference fit specimen consisted of a shaft in a sleeve, both of which were manufactured from EN24 steel (see Figure 4).

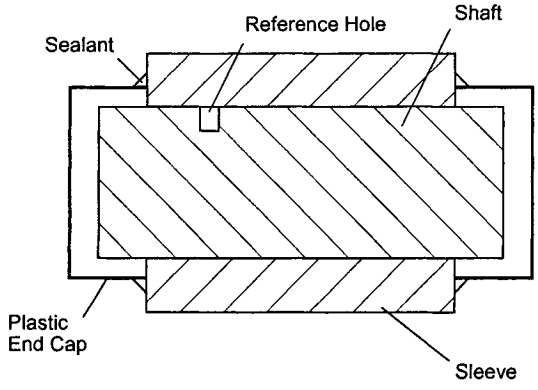


Figure 4. Interference Fit Specimen

Components were lathe finished to a mean surface roughness of 1.5 microns. The hole drilled in the shaft was employed as a reference point for the ultrasonic signal, as explained later. The specimens were assembled by press fitting using a hydraulic

press. While the ultrasonic measurements were performed the specimens were submerged in water, which acts as a couplant for the ultrasound, so plastic end caps were bonded to the specimen to prevent contamination of the interface.

The dimensions of the interference fit specimens examined are shown in Table 1. Dimensions of shafts and sleeves are given as well as interferences. All the specimens were assembled without lubrication.

3.3.Experimental Procedure

The interference fit specimens were placed on a vee block within the scanning tank and the transducer height set to scan the sleeve/shaft interface. The ultrasonic wave is refracted at the water/steel interface. The height of the transducer required to focus the wave on the interference interface was calculated using Snell's law of refraction (see Figure 5).

Line scans were then taken along the length of each specimen at 10 degree intervals and reflected pulses recorded in order to build up a map of the fit. The specimen was rotated in the vee block after each line scan.

Before assembly of the interference fit specimens a reference line scan was taken using the sleeve alone. As shown in Figure 6, away from the edges of the sleeve the reflected pulse is equal to the incident signal, A_i, as it is reflected from a steel-air interface. The decrease in the reflected signal at the edges of the sleeve occurs because as the probe travels over the edge of the sleeve some of the signal is lost (as shown in Figure 7).

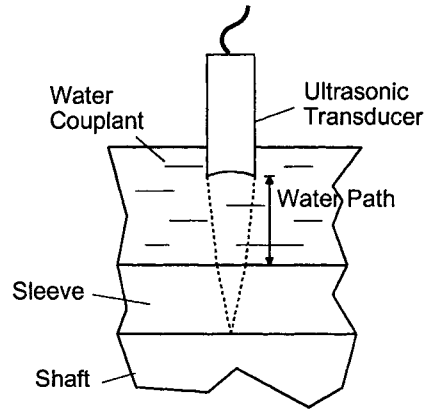


Figure 5. An Ultrasonic Pulse Focused on an Interface

This leads to a reduced ultrasonic signal incident on the interface in this region giving a decreased reflected voltage. The amount of signal lost is dependent on the sleeve geometry so another reference scan was carried out for the larger specimens.

Table 1
Interference Fit Specimen Geometries

		Shaft		Sleeve		
Diametral Interference (mm)	Assembly Method	O.D. (mm)	Length (mm)	I.D. (mm)	O.D. (mm)	Length (mm)
0.025	Press Fit	50	120	49.975	80	90
0.050	Press Fit	50	120	49.950	80	90
0.075	Press Fit	50	120	49.925	80	90

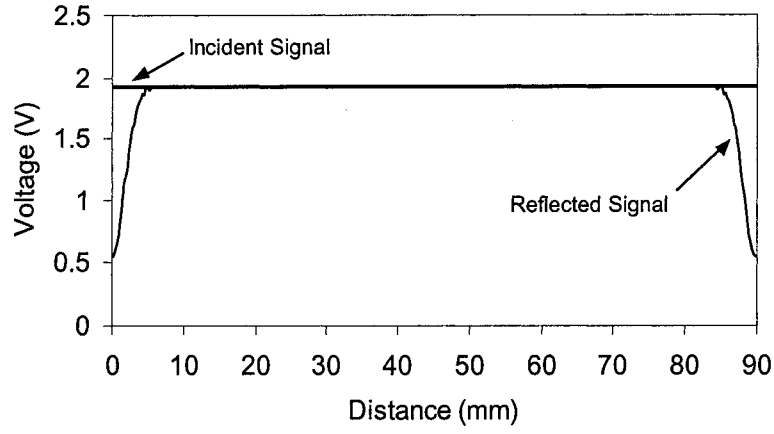


Figure 6. Reference Line Scan

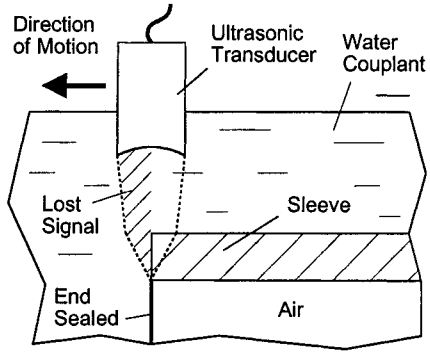


Figure 7. Incident Signal Loss at the Edge of the Sleeve

All subsequent reflected pulses, A_r, from the line scans taken on the assembled interference fit specimens, are divided by the reference scan to give a reflection coefficient, R. The reference holes drilled in each shaft were used to check the reference value as the assembled specimens were scanned.

3.4.Independent Calibration Experiment

In order to determine the relationship between interfacial stiffness and interface contact pressure, calibration tests were carried out using the apparatus shown in Figure 8.

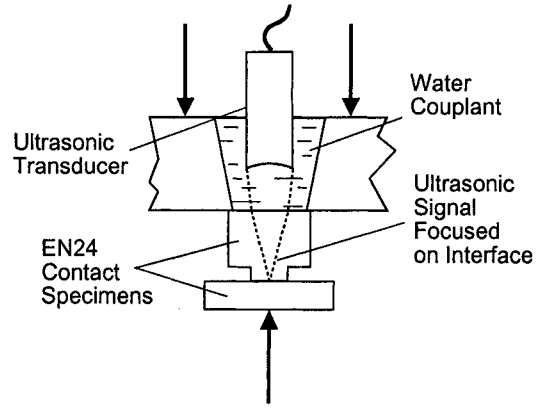


Figure 8. Calibration Apparatus

Two flat faced specimens, manufactured from EN24, with the same roughness on the contact surfaces as the sleeves and shafts from the interference fit specimens, were hydraulically loaded together. An ultrasonic

signal was then focused on the interface between the two specimens. The reflected signal was divided by a reference signal, obtained by recording the reflected pulse when the interface was unloaded, to obtain a reflection coefficient. This was then converted to an interfacial stiffness using the spring model (Equation 4). Stiffness was calculated for a number of different loads to generate a calibration curve (see Figure 9).

The empirical equation for the best fit line through the experimental data is given by $p = 51.72K$. It relates contact pressure, p, to interfacial stiffness, K, and has been used to process stiffness data for the interference fit specimens in this series of experiments.

4. RESULTS

4.1.Interface Pressure

Reflection scans were recorded for each of the interference fit specimens. Figure 10 shows scans for the 0.025mm, 0.05mm and 0.075mm fits.

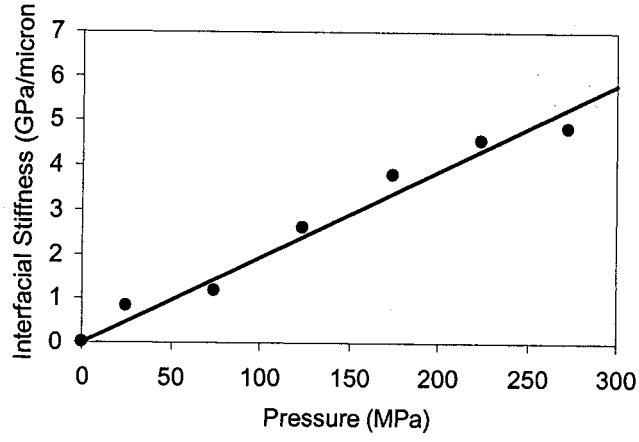


Figure 9. Interfacial Stiffness/Contact Pressure Calibration

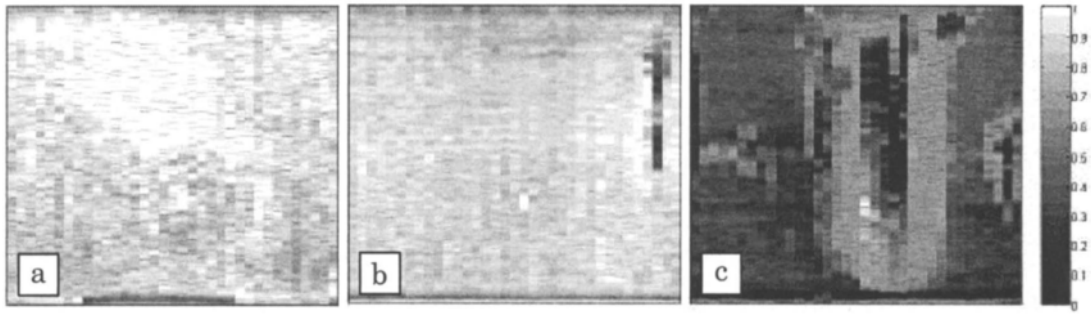
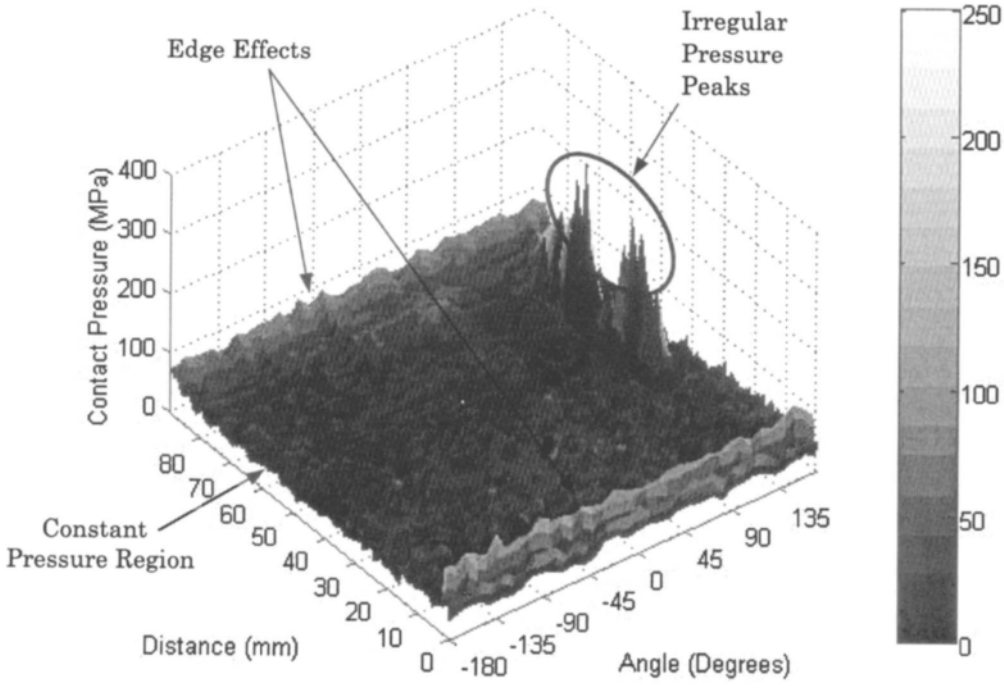


Figure 10. Reflection Coefficient Maps for (a) the 0.025mm Fit; (b) the 0.05mm Fit and (c) the 0.075mm Fit (axial direction top to bottom and circumferential direction left to right)

Each reflection coefficient was converted to a stiffness using the spring model. Then the calibration plot (Figure 9) was used to produce pressure maps. Figure 11 shows the pressure map for the 0.05 mm fit. As shown, the contact pressure along the length of the interface was not constant. However, as may be expected there was a high degree of radial symmetry. A central region is evident where pressure is approximately constant with sharp increases at the edges of the interface. In the 0.05 and 0.075mm fits anomalies occurred (such as that highlighted in Figure 11) where higher than expected pressures were evident.

Figure 12 shows pressures profiles along the length of the interference fit specimens. As the interference increases the contact pressure increases. Variation in the pressure is apparent within the plateau regions. This is probably caused by local variation in the roughness of the two surfaces.

The increase in contact pressure at the edges of the interference fits occurred for all the specimens. At the edges of such an interface, elastic stress analysis predicts an infinite stress discontinuity so this is not unexpected. As can be seen in Figure 12 the contact pressure decreases again before the edge of each fit. This was thought to be due to rounding of the sleeve edges that occurred during the manufacture of the specimens which would reduce the stress raising effect.



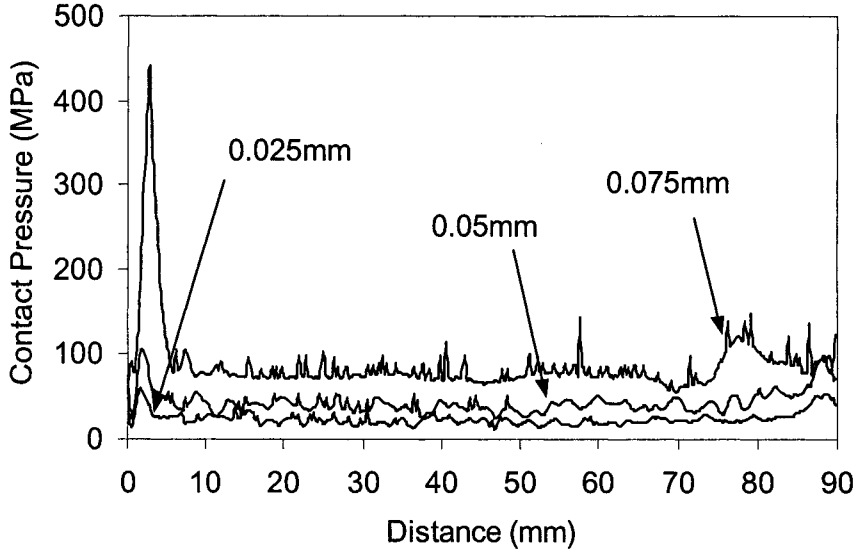


Figure 12. Line Scans of the 0.025, 0.05 and 0.075mm Fits (reflection measurements have been used to determine interface stiffness and hence contact pressure)

The contact pressures recorded for the specimens can be compared to the Lamé predictions given by Equation 1 (see Figure 13). The Lamé theory is not able to predict stress concentration effects at the edges of an interference fit so to provide an appropriate comparison between experiment and theory, the average pressures for the interference fits away from the edge effects were used. As shown in Figure 13, there is good correlation between the experimental results and Lamé predictions.

4.2.Surface Damage

After the experimental readings had been taken, the interference fit specimens were cut open. It was found from visual inspection of the interfacing surfaces that the anomalies seen in the pressure and reflection coefficient maps corresponded to damage. Figure 14a shows a photograph of the surface of the shaft from the 0.075mm interference fit specimen (the photo has been recorded by assembling a series of adjacent pictures of the shaft surface). Comparison of this image with the corresponding reflection coefficient map (Figure 14b) indicates that the surface damage clearly corresponds to the anomalies.

The anomalies were not caused by pressure increases, but attributable to surface damage scattering the ultrasonic signal. Similar correspondence between reflection coefficient disturbances and surface damage were observed in the 0.05mm fit, but where only minor localised scuffing had occurred. A relatively high load was required to fit the 0.05m and 0.075mm specimens, which was clearly high enough to lead to localised seizure and subsequent damage on the shafts and sleeves.

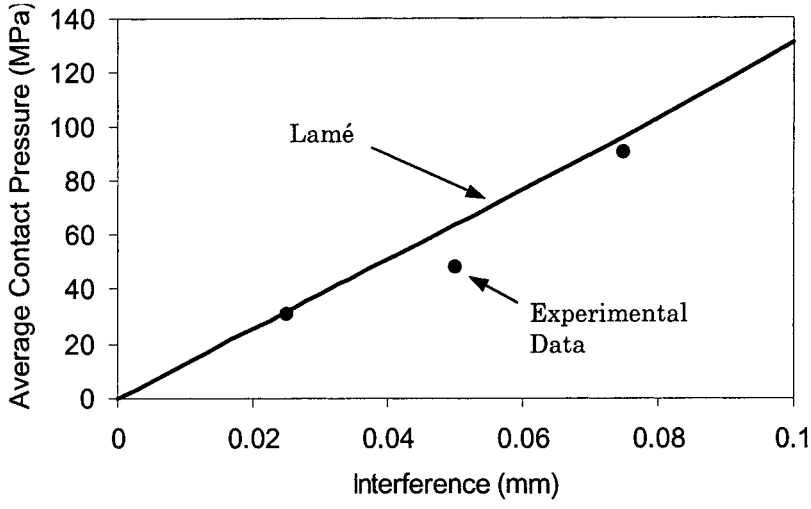


Figure 13. Comparison of Average Measured and Theoretical Interface Pressure (Equation 1)

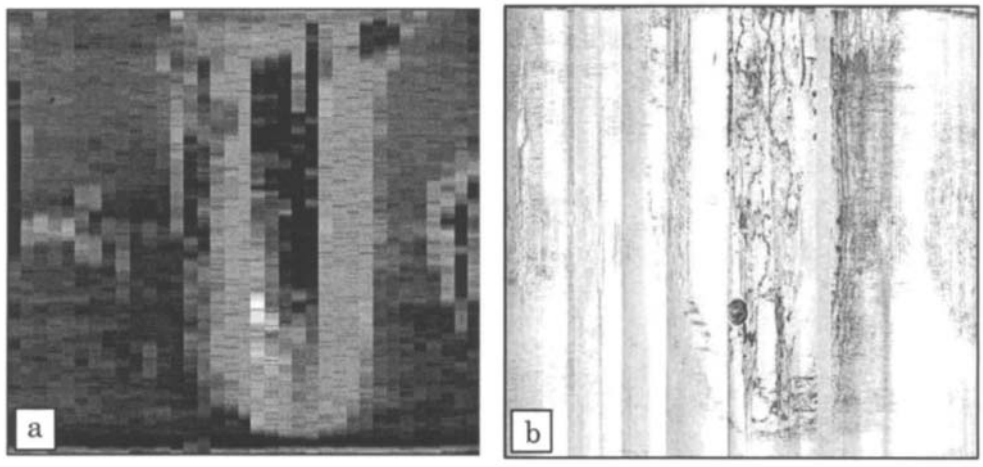


Figure 14. 0.075mm Fit (a) Reflection Coefficient Map; (b) Photograph of Shaft Surface

5. CONCLUSIONS

- A method has been established to determine interface pressures in a press-fit component. The method relies on the measurement of reflected ultrasonic wave and a parallel calibration procedure.
- The interface contact pressure has been determined for a number of different press fit cases. The results show a central

region of approximately uniform pressure with edge stress at the contact sides. The magnitude of the pressure in the central region agrees well with the elastic Lamé analysis.

- In the more severe fits the surface scuffed which resulted in anomalies in the reflected ultrasound. These anomalies were associated with regions of surface damage at the interface.
- This approach is limited by the necessity for a separate calibration to obtain contact pressure directly from reflection and the spatial resolution (at the frequencies used herein is relatively coarse at 0.9mm spot diameter).

REFERENCES

1. Boresi, A.P., Schmidt, R.J., Sidebottom, O.M., 1952, *Advanced Mechanics of Materials*, Wiley.
2. Karami, G., Oskooei, S.G., 1994, "Boundary Element Analysis of Shrink-Fit Type Problems", *Proceedings of the 2nd Biennial European Joint Conference on Engineering Systems Design and Analysis*, Vol. 8A, London, England, 4-7 July, pp17-21.
3. Zhang, Y., McClain, B., Fang, X.D., 2000, "Design of Interference Fits Via Finite Element Method", *International Journal of Mechanical Sciences*, Vol. 42, pp1835-1850.
4. Tattersall, A.G., 1973, "The Ultrasonic Pulse-Echo Technique as Applied to Adhesion Testing", *J. Phys. D: Appl. Phys*, Vol. 6, pp819-832.
5. Drinkwater, B.W., Dwyer-Joyce, R.S., Cawley, P., 1996, "A Study of the Interaction between Ultrasound and a Partially Contacting Solid-Solid Interface", *Proceedings of the Royal Society, Series A*, Vol. 452, No. 1955, pp2613-2628.
6. Thomas, T.R. Sayles, R.S., 1977, "Stiffness of Machine Tool Joints: A Random-Process Approach", *Trans. ASME, J. Eng. Ind.*, Paper No. 76-WA/Prod-23.
7. Kendall, K. and Tabor, D., (1971), An Ultrasonic Study of the Area of Contact between Stationary and Sliding Surfaces", *Proceedings of the Royal Society, Series A*, Vol. 323, pp321-340.
8. Dwyer-Joyce, R.S., Drinkwater, B.W., "Analysis of Contact Pressure Using Ultrasonic Reflection", *Experimental Mechanics, Proceedings of the 11th International Conference on Experimental Mechanics*, Balkema, Rotterdam, pp747-754.

SESSION XIII
COMPONENT TRIBOLOGY (2) - BEARINGS AND MECHANISMS

Chairman : Professor G Poll

Paper XIII (i) **Tribologically-based Design of Precise Positioning Stage**

Paper XIII (ii) **Journal Bearing Oil Film Measurement Using Ultrasonic Reflection**

Paper XIII (iii) **THD Lubrication of Mechanical Face Seals During Transient Period After Start-up : 2D Modelisation**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribologically-based Design of Precise Positioning Stage

K. Adachi [a], T. Yamaguchi [a], Y. Ishimine [b] and K. Kato [a]

[a] Laboratory of Tribology, Tohoku University, Sendai 980-8579, Japan

[b] Kyocera Corporation, Kokubu 899-4396, Japan

To produce modern engineering system in which high specification is required, tribology becomes an essential design input to achieve ultimate performance. A new concept of machine design from a viewpoint of tribology is therefore necessary: "Tribologically-based machine design, defined as design of the overall system based on to achieve ultimate tribological performance" in addition to the more conventional "design of a reliable mechanical element based on tribological understanding."

In this paper, a compact and highly-precise linear positioning stage driven by friction with an ultrasonic motor is proposed for use in an electron-beam direct-writing system, and is realized by the optimum design of the overall system, including tribological materials and operating conditions, using the concept described above.

It provides a size reduction to one half of the conventional stage, and a positioning precision which is three times as great.

1. INTRODUCTION

Machine design is usually based on machine dynamics, material mechanics, fluid mechanics and thermodynamics. Since 75% of its causes of defects, breakdowns and lifetime limitation in the machines are associated with the interaction of contacting surfaces [1], tribology should be regarded not only as a fundamental basis for the design of machinery alongside the four main fields listed above, but as a central aspect of the design of reliable and durable machines.

Understanding of elastic contacts, based on Hertzian theory, enables the fundamental design of well-performed ball bearings, and hydrodynamic lubrication theory makes the design of advanced sliding bearings possible. With the improvement of materials and/or the introduction of new materials based on tribological understandings, the lifetimes of tribological elements such as ball bearings [2], sliding bearings [3] and mechanical seals [4] can be substantially extended.

These examples illustrate the successful application of tribological knowledge to the design of mechanical elements for the efficient and reliable operation of machinery.

The design of a mechanical element based on tribological knowledge is sometimes called "Tribodesign"[5].

The limit of the performance of modern high-tech machinery is often governed by characteristics of the contact interface [6]. An example is the hard disk drive system, in which the storage capacity is decided by the head-disk spacing, which is controlled by surface design. In such high specification machinery, tribology is increasingly essential as the key design input to achieve ultimate performance.

A new concept of machine design from the viewpoint of tribology is therefore necessary: "Tribologically-based machine design, defined as design of the overall system based on to achieve ultimate tribological performance" in addition to the more conventional "design of a reliable mechanical element based on tribological understanding."

In this paper, a new precise positioning stage is described, designed using the concept stated above.

2. PRECISE POSITIONING STAGE DRIVEN BY FRICTION WITH AN ULTRASONIC MOTOR

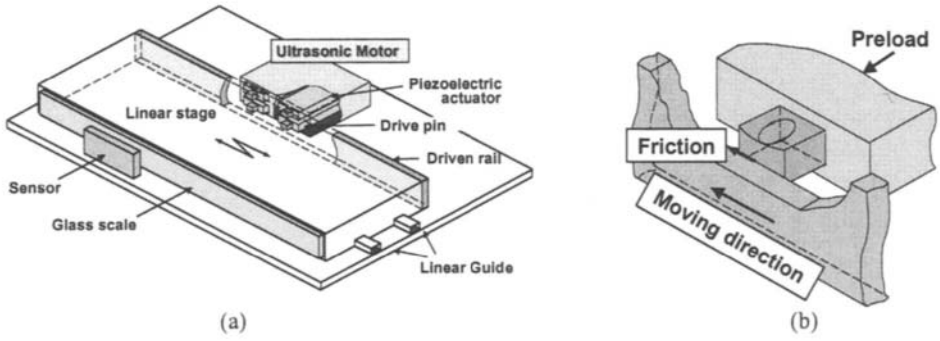


Fig. 1 A schematic image of precise positioning linear stage driven by friction with an ultrasonic motor (a), and its driving mechanism (b).

High performance Large Scale Integrated circuits (LSI) with high processing speed and low power consumption have been realized by integration and miniaturization. The semiconductor road map for DRAM (Dynamic Random Access Memory) predicts the capacity of the memory in 2013 will reach 64Gbit, which requires a minimum line width of 32nm [7].

High-resolution manufacturing for such small-scale integration is achieved by a micro-machining technology such as lithography. The electron beam (EB) direct-writing system is one of the key methods to manufacture a high-resolution mask for lithography, and a precise positioning stage is an essential element for high-resolution manufacturing using an EB direct-writing system. The stage is required to operate in vacuum and to have no magnetic fields since these would influence the path of the electron beam. In addition, high precision and high speed positioning is required to achieve high accuracy and high productivity [8].

Conventional stages for EB direct-writing systems are driven by electromagnetic motors that generates magnetic fields. The motor is therefore placed outside the vacuum chamber, and the driving force and motion are transmitted through ball screws. This causes the stage system to be large and complex, and moreover limits the performance in terms of precise positioning and high speed

The linear ultrasonic motor (USM) is a tribological device, in which the driving force is frictional. It provides a promising actuator for small and precise positioning stages, because of its

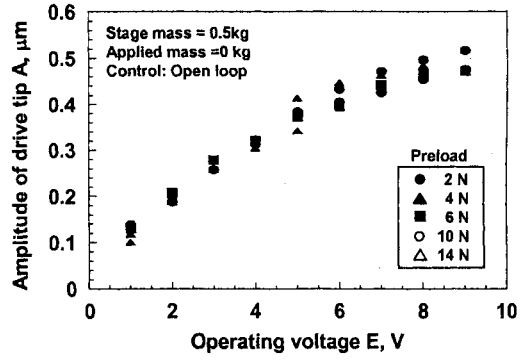


Fig. 2 Amplitude of the drive tip's oval vibration that is controlled by the operating voltage applied to the piezoelectric actuator of the USM.

compact size, high torque, superior controllability and responsiveness [9], [10].

In addition, the USM generates no magnetic fields, and can therefore be used beside the stage within the vacuum chamber, even for EB direct-writing systems.

Fig. 1 (a) shows a schematic diagram of the precise positioning stage driven by USM for EB direct-writing system developed in this work, and driving mechanism is shown in Fig. 1 (b).

The use of a USM resulted in a size reduction to one half of the previous design, and a positioning precision which was three times higher.

The stage is driven by the friction force generated in the contact interface between a drive tip of USM and a driven rail on the stage. The drive tip of the USM is oscillated elliptically (i.e. with an oval

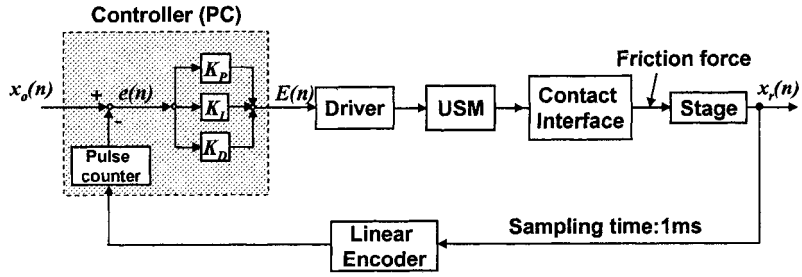


Fig. 3 Block diagram of PID control for precise positioning.

vibration path) at an ultrasonic frequency of 40 kHz by a controlled piezoelectric actuator. The amplitude of vibration is controlled by the operating voltage applied to the piezoelectric actuator, as shown in Fig. 2.

Through the vibration amplitude change, the stage speed increases with the applied voltage.

The positioning of the stage is controlled by a closed loop (feedback) algorithm of PID controller in the PTP (point to point) positioning method. The design demands the movement of the stage by 650μm ± 1μm within 200ms. The value of the operating voltage applied to the USM is determined from the deviation *e(n)* between the true and ideal positions of the stage every 1 ms. The position control method is summarized in the block diagram in Fig. 3. The operating voltage *E(n)* applied from the PID controller to the motor is determined from the following equation.

$$E(n) = K_P e(n) + K_I \left(\sum_{i=1}^{n} e(i) \right) + K_D (e(n+1) - e(n)) \quad (1)$$

Where K_P, K_I and K_D are the proportional gain, integral gain and differential gain respectively, and *e(n)* is the position error at nth position detecting.

3. DESIGN OF THE PRECISE POSITIONING STAGE FROM A TRIBOLOGICAL VIEWPOINT

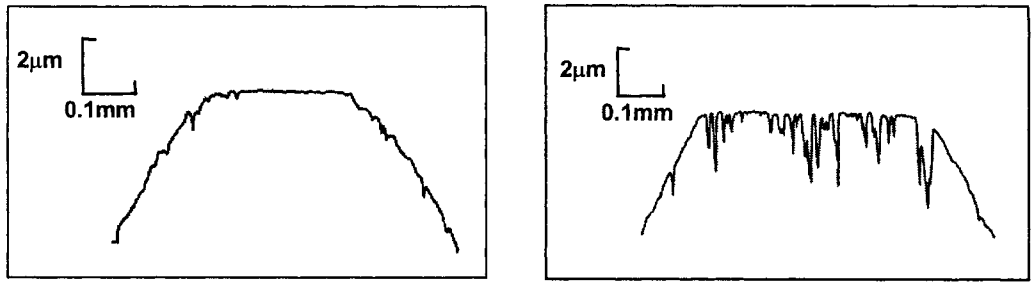
The stage is driven and controlled by the friction force between the drive tip of the USM and the driven rail fixed to the linear stage. The positioning precision and the durability of the stage are therefore determined by tribological properties such as friction and wear at the contact interface between the drive tip and the driven rail.

For accurate transmission of the motion, stable and relatively high friction at the contact interface is required. Moreover high wear resistance of both tribological elements is essential to maintain accurate performance, since wear causes a decrease in the mechanical accuracy of the system.

3.1. Material selection of tribological elements

From previous work on the wear of materials in unlubricated sliding conditions [11]-[14], it was shown that advanced ceramics exhibit two wear modes which can be classified as "mild wear" and "severe wear" [11]. In the mild wear mode, the surface became relatively smooth (polishing behavior), and a low wear rate of less than 10^{-6} mm^3/Nm was obtained even though the friction coefficient was not small. This wear rate corresponds to a higher wear resistance, than that seen in any metal.

This suggests that the use of advanced ceramics in the mild wear region for the tribological elements is essential to achieve the design performance, the initial contact condition between the drive tip and the driven rail will be maintained because of the small change in surface roughness and geometry of the tip and the rail. Based on this tribological understanding, alumina ceramics were selected for both the drive tip and the driven rail.



(a) (b)

Fig. 4 Surface profiles of two representative wear surfaces of the drive tip observed after operation of the stage.

3.2. Design to achieve precise positioning and mild wear

Figures 4 (a) and (b) show surface profiles of two representative wear surfaces of the drive tip with spherical shape (radius: 7mm) observed after operation of the stage. They clearly show the results of mild and severe wear modes as mentioned above. It is obvious that the optimum design of the whole system, including the operating condition, is the key to maintaining low wear in the mild wear regime, even if a so-called wear-resistant material is used.

For ceramics, both of the following criteria must be satisfied to achieve mild wear [11].

$$S_{c,m} = \frac{(1+10\mu)P_{max}\sqrt{d}}{K_{IC}} \leq 6 \qquad (2)$$

$$S_{c,t} = \frac{\gamma\mu}{\Delta T_s}\sqrt{\frac{vWH_v}{k\rho c}} \leq 0.04 \qquad (3)$$

where μ is the friction coefficient (tangential coefficient), P_{max} the maximum Hertzian contact pressure, d the grain size, K_{Ic} the fracture toughness, γ the heat partition ratio, ΔT_s the thermal shock resistance, v sliding velocity, W the normal load, H_v the Vickers hardness, k thermal conductivity, ρ density and c specific heat. $S_{c,m}$ and $S_{c,t}$ is named as mechanical severity of contact and thermal severity contact, respectively, which describes an index of crack propagation caused by mechanical stress or thermal one

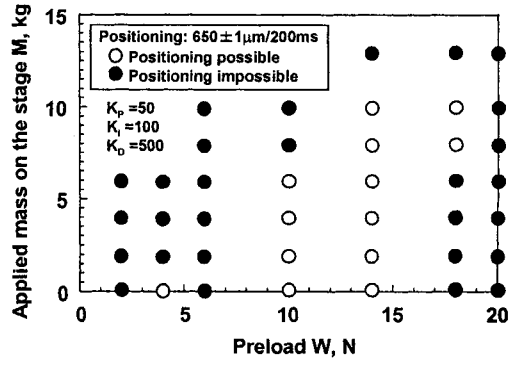


Fig. 5 Distribution of preload and driven stage mass for precise positioning possible.

These necessary conditions for the mild wear of ceramics stress the importance of the design of contact pressure between the drive tip and the driven rail, and control of the amount of slip generated at the contact interface to transmit the force and the motion, related to the tangential coefficient in above criteria.

3.2.1 Selection of preload between the drive tip and the driven rail

Figure 5 shows for various combinations of preload and driven stage mass whether precise positioning (i.e. movement by 650μm±1μm within 200ms) is possible or not under constant PID gains.

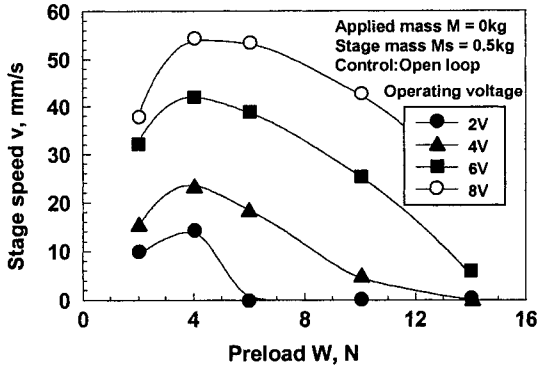


Fig. 6 Effect of preload on the speed of the stage with four different operating voltage applied to the USM in open loop control.

It can be seen that the optimum preload, at which precise positioning is possible, depends on the mass of the driven stage.

Figure 6 shows the effect of preload on the speed of the stage for the same driven mass with four different operating voltage applied to the USM in open loop control. For all operating voltages, there exists a preload W_c which gives the maximum stage speed. For preload W smaller than W_c, it would appear that macroscopic slip between the drive tip and the driven rail introduces with a decrease in preload. Increase in the amount of macroscopic slip decreases transmission efficiency between the drive tip and the driven rail, and therefore the stage speed decreases. For preload W greater than W_c, stage speed decreases with increasing preload. The stage speed is zero under conditions E=2V, W=8N because the drive tip is in contact with the driven rail in at all times and the stage oscillates about a fixed position. As shown here, for preload $W > W_c$, the transmission efficiency of the driving force decreases with an increase in preload because the time for which the drive tip and the driven rail are out of contact decreases, but macroscopic slip is avoided. This shows that the stage should be driven with a preload $W>W_c$ to avoid macroscopic slip between the drive tip and the driven rail.

From these considerations of transmission efficiency, including slip characteristics at the contact interface, which is related the driven stage

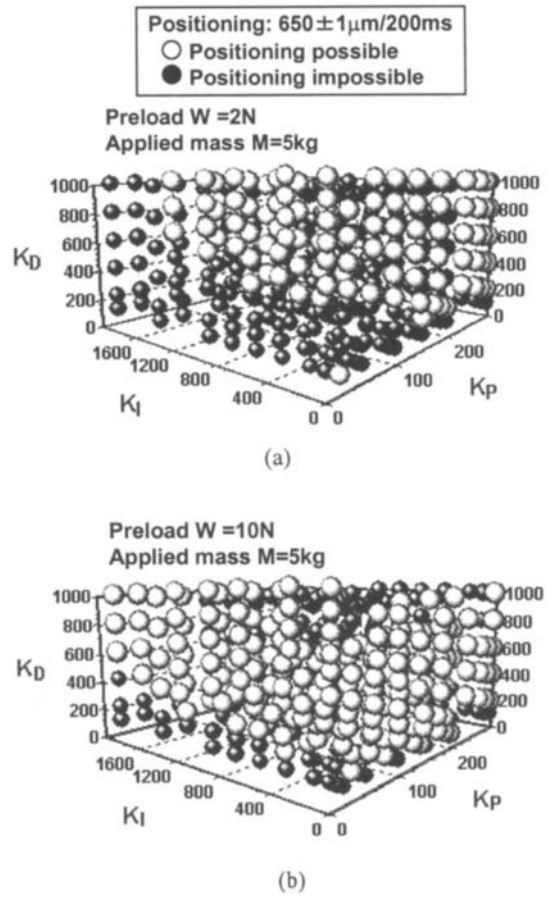
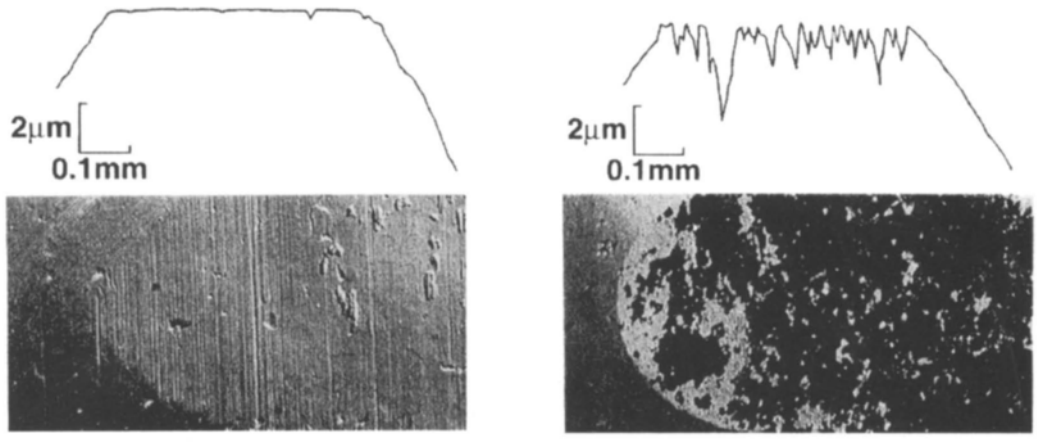


Fig. 7 Combinations of PID gains for precise positioning possible for two different combination of preload and driven stage mass.

mass, the optimum preload for attaching the USM can be selected.

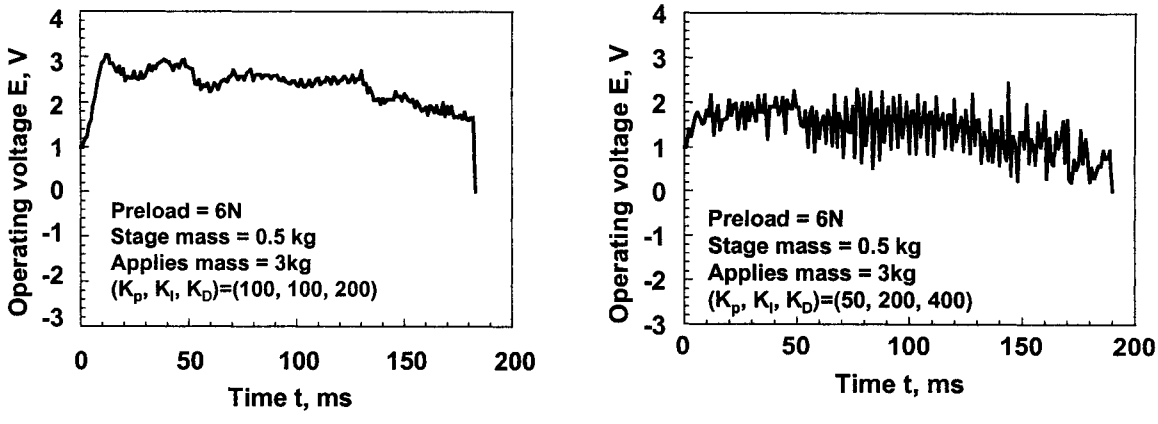
3.2.2 Design of PID gains for feedback control of positioning

Figures 7 (a) and (b) shows for various combinations of PID gains whether precise positioning (i.e. movement by 650μm±1μm within 200ms) is possible, for two different combination of preload and driven stage mass. It can be seen that setting of the optimum PID gains is required to achieve precise positioning, which are related to the combination of preload and driven stage mass.



(a) (b)

Fig. 8 Wear surfaces of the drive tip observed after operation with the same preload and driven stage mass with two different PID gains for which precise positioning is possible.



(a) (b)

Fig. 9 Operating voltage applied to the USM to control precise positioning with PID gains (K_P, K_I, K_D)=(100, 100, 200) (a) and (50, 200, 400)(b).

Figures 8 (a) and (b) show representative wear surfaces of the drive tip observed with the same preload and driven stage mass with two different PID gains for which precise positioning is possible.

It can be seen that severe wear can be generated even under conditions where precise positioning is possible.

This is caused by differences in the amount of slip generated in the contact interface to transmit the force and the motion to the driven rail, for different PID gains. Selection of PID gains to achieve mild wear is therefore more important than that needed for precise positioning.

These observations demonstrate that tribologically-based design for the control algorithm

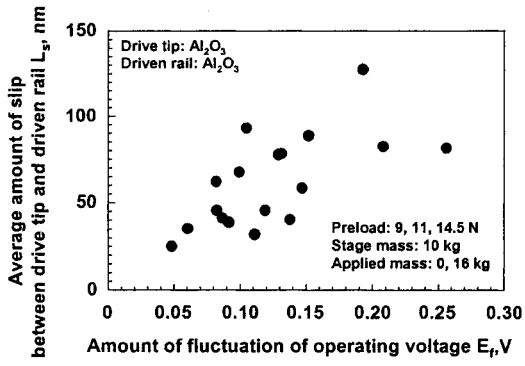


Fig. 10 Relation between the amount of fluctuation of the operating voltage Ef and the average amount of slip between the drive tip and the driven rail.

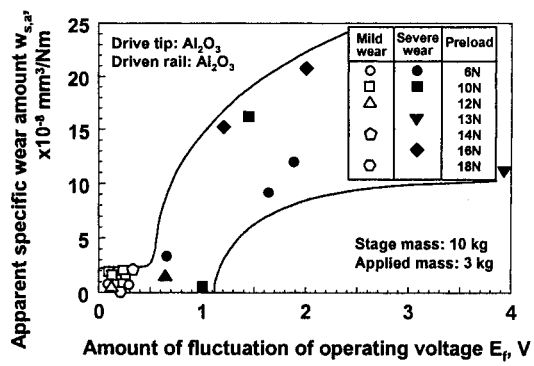


Fig. 11 Specific wear amount and wear mode of the drive tip as a function of the amount of fluctuation of operating voltage Ef.

is required in addition to design of the mechanical system and selection of materials based on tribological knowledge.

The PID gains directly affect the operating voltage applied to the USM. The operating voltage controls the amplitude of the drive tip vibration shown in Fig. 2. Two examples of operating voltage applied to the piezoelectric actuator to control precise positioning, with different combinations of PID gains, are shown in Figs. 9 (a) and (b). The difference in operating voltage caused by the different PID gains causes a different motion of drive tip to achieve precise positioning. As a result, the amount of slip at the contact interface between the drive tip and the driven rail is different: the amount of slip is related to the operating voltage applied to the USM.

The extent of fluctuation of the operating voltage E_f, which means the average variation of operating voltage during the sampling time, can be used to express the difference in the waveforms of the operating voltage.

$$E_f = \frac{\sum_{n=0}^{N} (E_{n+1} - E_n)}{N} \qquad (4)$$

where n is the number of sampling time, N is the number of the sampling time when the required positioning is completed, E_n is the operating voltage applied to the USM at nth sampling .

Figure 10 shows the relation between the fluctuation of the operating voltage E_f and the average amount of slip between the drive tip and the driven rail. It is clear that the amount of slip can be evaluated indirectly from the value of E_f expressed by eqn. (4). Figure 11 shows the specific wear amount of the drive tip, which is related to the wear mode, as a function of the E_f. Fig.11 shows a critical value of E_f that determines the transition from mild wear to severe wear. This point would correspond to the critical tangential coefficient in eqns. (2) and (3).

The fluctuation of the operating voltage E_f has been shown to be a good indirect parameter for monitoring and controlling the wear mode of tribological elements. Based on this understanding, optimum PID gains can then be designed to achieve minimum value of E_f.

The precise positioning stage described above is an example of new ultimate performance can be realized by tribologists through tribogically-based design of overall system including selection of material, pre-setting as a hard system and PID gains as a soft system.

4. CONCLUSIONS

"Tribologically-based machine design" has been proposed as a concept to produce engineering systems in which the ultimate performance is required.

A precise positioning stage driven by friction with an ultrasonic motor provides a promising solution for an electron-beam direct-writing system for the next generation LSI, and has been realized by the optimum design of overall system using the concept described above.

The use of ceramics, exhibiting mild wear, as contacting surfaces in this design leads to substantially enhanced performance: three times the positioning accuracy and one half of the physical size of a conventional stage driven by an electromagnetic motor.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Professor I. M. Hutchings at the University of Cambridge for his kind and helpful discussion. We also are grateful to Mr. M. Suzuki and Mr. C. Yamashita at Tohoku University for their assistance in experiments.

REFERENCES

[1] I. V. Kragelskii, M. N. Dobychin and V. S. Kombalov, Friction and Wear, Pergamon press, (1987) 237.

[2] J. R. Miner, W. A. Grace and R. Valori, A demonstration of high-speed gas turbine bearing using silicon nitraide rolling elements, Lubrication Engineering, 37, 8 (1981) 462-464, 473-478.

[3] K. Tanba, Recent sliding bearing and tribology, Nonthly Tribology, Shinju-sha press, 6 (1992) 24-26.

[4] S. Asanabe, Applications of ceramics for tribological components, Tribology International, 20, 6 (1987) 355-364.

[5] T. A. Stolarski, Tribology in machine design, Heinemann Newnes, Oxford, (1990) 1-12.

[6] B. Bhushan, Tribology and mechanics of magnetic storage devices, Second Ed., Springer, (1996).

[7] ITRS (International Technology Roadmap for Semiconductors) Executive Summary 2001 Edition, Downloaded from http://public.itrs.net/Files/2001ITRS/Home.htm

[8] T. Saito, *Precision positioning technology for next generation*, Fuji Techno System, T. (2000) 102-108.

[9] K. Adachi, K. Kato, Y. Sasatani, The micro-mechanism of friction drive with ultrasonic wave, Wear, 194 (1996), 137-142.

[10] H. V. Barth, Ultrasonic Driven Motor, IBM Tech. Disclosure Bull., 16, 7, (1973) 2263.

[11] K. Adachi, K. Kato and N. Chen, Wear map of ceramics, Wear, 203-204 (1997) 291-301.

[12] K. Kato and K. Adachi, Wear Mechanisms, Chapter 7 of Modern Tribology Handbook, B. Bhushan Ed., CRC press, (2000) 273-300.

[13] K. Adachi and K. Kato, Formation of smooth wear surfaces on alumina ceramics by embedding and tribo-sintering of fine wear particles, Wear, 245 (2000) 84-91.

[14] K. Adachi, K. Hokkirigawa and K. Kato, The wear mechanism of silicon nitride in rolling-sliding contact, Wear, 151 (1991) 291-300.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Journal Bearing Oil Film Measurement Using Ultrasonic Reflection

P. Harper[a], B. Hollingsworth[b], R. Dwyer-Joyce[a], B. Drinkwater[c]

[a]Department of Mechanical Engineering,
Mappin Street, University of Sheffield, Sheffield S1 3JD, UK

[b]Department of Mechanical Engineering,
Brown University, Providence, Rhode Island, USA.

[c]Department of Mechanical Engineering,
University Walk, University of Bristol, Bristol BS8 1TR, UK

Abstract

A wave of ultrasound will reflect from a lubricant film between two bearing components. The response of the film to a sound wave depends on the acoustic and geometrical properties of the oil layer. For thin lubricant films the layer behaves as a spring and the proportion of the wave reflected (or the 'reflection coefficient') is related to the stiffness of the oil layer. This provides a method for determining the film thickness.

An ultrasonic transducer has been mounted on the outside of a hydrodynamic journal bearing bush. Pulses are emitted through the bearing bush and reflected back from the oil film. The reflected pulses are picked up by the same transducer and processed to determine the proportion of the wave amplitude reflected. The film thickness has then been calculated from the reflection spectra.

The film thickness has been determined for a range of steady state bearing operating conditions and the results compared, qualitatively, with published numerical solutions. The method has been shown to be robust and flexible and demonstrates the potential of an ultrasonic based lubricant film monitoring device.

1. INTRODUCTION

Full film hydrodynamic journal bearings are widely used components in industrial machinery. They are relatively cheap and flexible to incorporate into an item of plant. Under the right conditions of lubrication and assembly they have long lives. Failures usually occur by overload or when the lubricant supply is interrupted. These failures can lead to damage to the bearing surfaces (including 'wipes' on the white metal bearing layer). At start-up and shut-down the lubricant film collapses (unless it is augmented from an externally pressurised source) and again surface contact and potential damage can occur. The failure and repair of these failures can result in expensive plant down time.

Generally film thickness monitoring of journal bearings is not routinely carried out. In larger bearings lubricant film monitoring is sometimes performed using proximity probes to determine the relative position of the shaft and bush. In development apparatus film thickness is sometimes monitored by measuring the capacitance of the separating lubricant layer. This has the disadvantage of giving a reading based in an aggregate of the film thickness (and not a local value at the minimum point) and a requirement to know the lubricant capacitance. This work is aimed at developing an absolute method of monitoring the lubricant film using ultrasonic wave reflection.

2. BACKGROUND

An ultrasonic wave will be partially reflected at an interface between two materials. The greater the difference in acoustic impedance (the product of wave speed and density) between the two materials, the greater the proportion of the wave reflected. So when a wave travelling through steel strikes an air interface it is initially all reflected. But if it strikes an oil interface, for example, some 55% of the amplitude is reflected.

This phenomenon has been widely used to study the rough surface contacts [1, 2, 3] where the wave passes through regions of contact but is reflected at regions of non-contact. Interpretation of the reflected wave can give information about the extent of the contact region, the real area of contact, and the contact pressure.

When a wave strikes an oil layer between two solid surfaces, a series of reflections will occur at each interface (shown schematically in figure 1). If the oil film is sufficiently thick then the first two reflected waves can be spatially separated and the thickness readily determined from the time between their arrival. A typical reflected pulse in shown in figure 2a.

This *time of flight* method is commonly used in thickness gauging. Typically it will measure oil films greater than about 100 μm, depending on the frequency of the wave used. Whilst this is a useful technique many hydrodynamic lubricant films are thinner than this (and in any event proximity probes have a good accuracy at this scale).

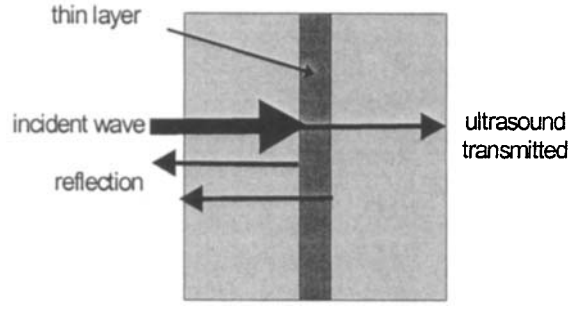


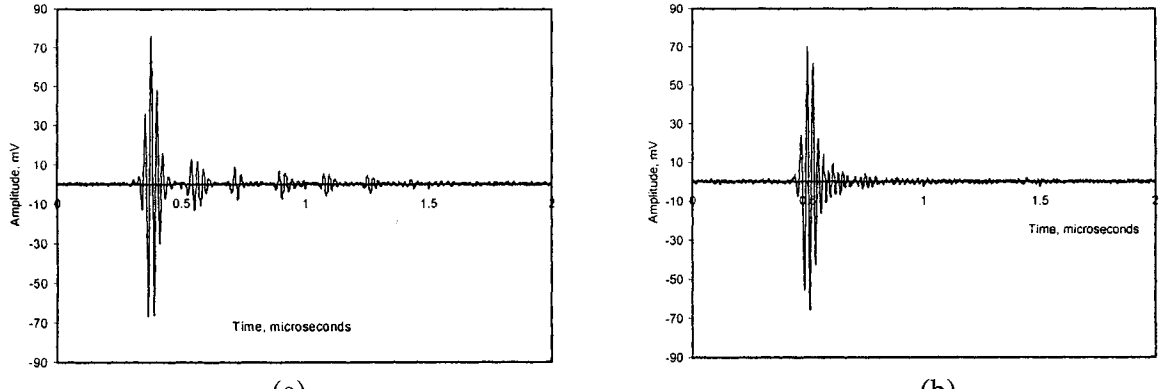
Figure 1. Schematic representation of an ultrasonic wave striking an oil film.

If the film is thin, then the reflections arrive close together and cannot be distinguished in time. The reflected pulse appears with a series of reverberations as shown in the example of figure 2b. Extracting film thickness data from this pulse requires more effort.

The response of the thin film to the ultrasound wave is analogous to that of a spring at the interface [4]. The reflection coefficient depends on the stiffness per unit area K of the layer, according to a spring model.

$$R = \frac{z_1 - z_2 + i\omega(z_1 z_2 / K)}{z_1 + z_2 + i\omega(z_1 z_2 / K)} \qquad (1)$$

where z_1 and z_2 are the acoustic impedances of the materials either side of the layer.



(a) (b)

Figure 2. Reflected ultrasonic pulses from (a) thick oil film and (b) thin oil film.

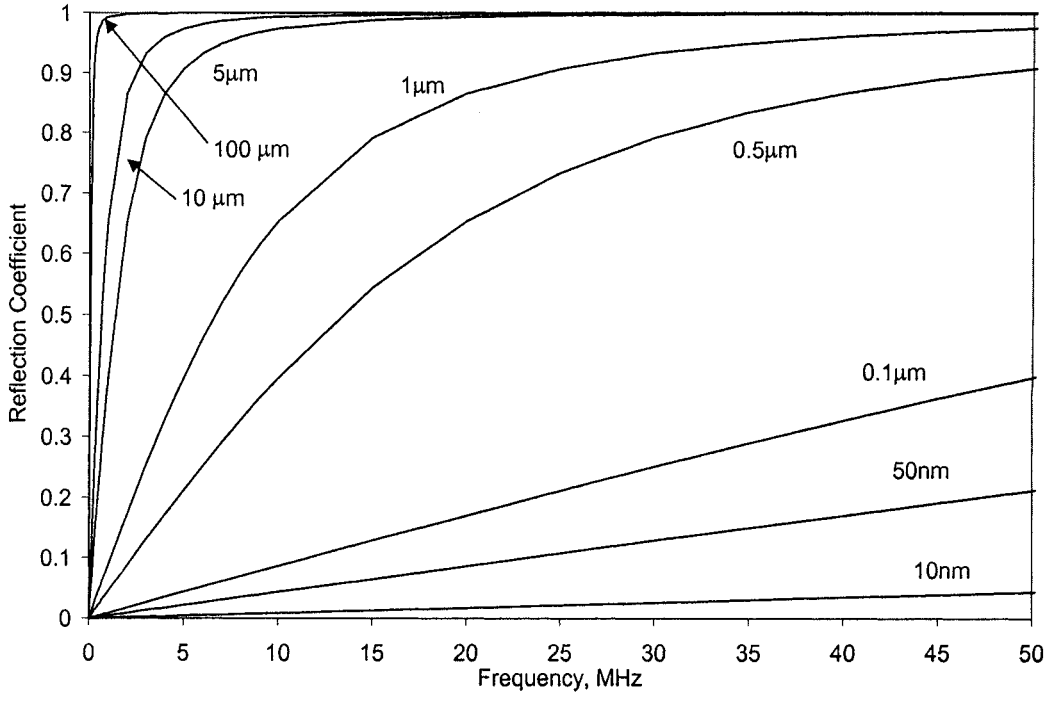


Figure 3. Predicted reflection coefficient variation with frequency for waves striking films of various thickness.

The stiffness of the liquid layer depends on its thickness, h and mechanical properties according to [5],

$$K = \frac{\rho c^2}{h} \tag{2}$$

where ρ is the density of the lubricant and c is the speed of sound passing through it. Combining equations 1 and 2 gives a relationship between the film thickness and reflection coefficient. Figure 3 shows the spring model predictions (equation 1) for reflection coefficient against ultrasonic frequency for a series of lubricant film thickness. In this example the material either side of the oil layer is steel. In practice measurements in the range $0.1 < R < 0.9$ are feasible, which means that to measure the hydrodynamic film of the order of 10 – 50 μm ultrasonic frequencies of 0.5-2 MHz are needed. To measure the thinner films in the elastohydrodynamic regime, higher frequencies are used [Dwyer-Joyce et al, 2002]. This method is known as the *film spring* method.

If either or both the ultrasonic frequency and the film thickness are increased, it is possible to cause the liquid layer to resonate. When this occurs all the amplitude of the resonant frequency is transmitted, i.e. $R(f_{res}) = 0$. The resonant frequency is again a function of the film thickness and mechanical properties. The *film resonance* approach can also be used to determine the film thickness [6] for appropriate combinations of layer thickness and measurement frequency.

In this work experiments on a simple hydrodynamic journal bearing rig have been performed. The maximum radial clearance in the bearing is 50 μm and lubricant films in the range 5 to 20 μm are expected. The *film spring* method is chosen for measurement in this regime.

3. APPARATUS & PROCEDURES

3.1. Journal Bearing Test Apparatus

Experiments were carried out on a purpose built simple hydrodynamic journal bearing test rig (Figure 4 shows a photograph of the apparatus and figure 5 shows a schematic of the loading arm).

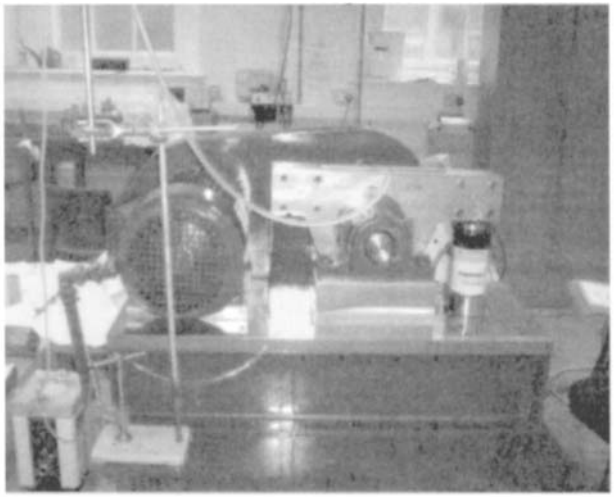


Figure 4. Photograph of the bearing test rig.

The shaft is rotated by means of a 3kW variable speed motor (0-710 rpm) and is loaded by means of a hydraulic actuator (0-8.8 kN). The shaft is constructed from steel (EN 24) and the bearing bush is constructed from 60-40 brass. The diametrical clearance is 50um with the surfaces being ground to a high surface finish.

A mineral oil (Shell Turbo T68) is fed into a lubricating hole out of the loaded region. Gravity feed was found to be sufficient to maintain a continuous film of oil in the bearing. A thermocouple was located in the oil flow immediately as it drained from the bearing.

3.1.1 Ultrasonic Signal Generator

A planar piezo-electric ultrasonic transducer with a 1.1MHz centre frequency was used. The transducer is positioned inside a drilled recess in the bearing bush in the region directly opposite the point of load application (figure 6 shows the location of the transducer).

In the apparatus design of the apparatus it is possible to move the measurement point to other circumferential locations around the bearing (although in this study this work was not done). The active face of the transducer is coupled to the bearing bush by means of a gel couplant. The transducer is held in place by a backing plate and spacer, ensuring a consistent joint.

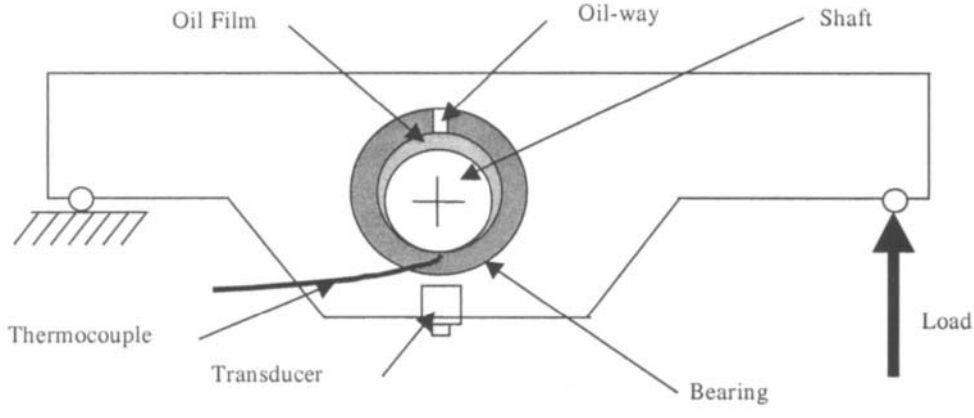


Figure 5. Schematic diagram of the journal bearing and loading arm.

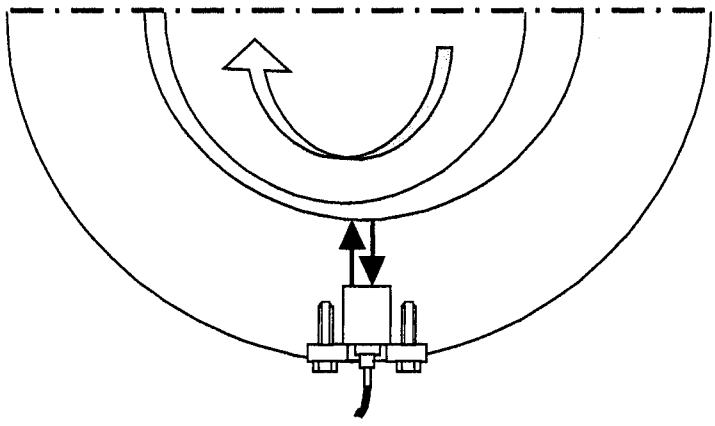


Figure 6. Schematic diagram of the location of the ultrasonic transducer.

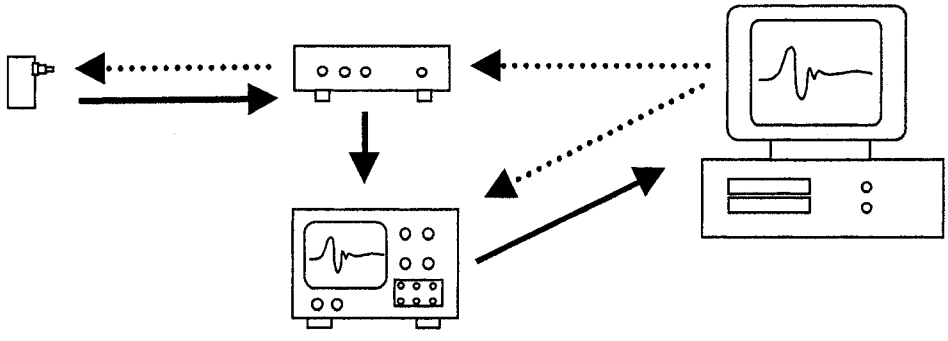


Figure 7. Schematic diagram of the ultrasonic pulsing, acquisition, and processing apparatus.

The measuring equipment consisted of a PC running Labview software, with a data acquisition card (GPIB), a LeCroy Waverunner Oscilloscope, and a Ultrasonic Pulser/Receiver (see figure 7).

The PC controls the Ultrasonic Pulser/receiver (UPR) via a R232 port, setting the pulsing rate, pulsing frequency and the amplification of the received signal. The UPR pulses the transducer that creates the ultrasonic signal. The UPR then receives the reflected pulse and transfers it to the oscilloscope. The oscilloscope then digitises the information before it is downloaded to the computer. The computer also records the temperature of the oil as it escapes from the oil film.

3.1.2 Signal Processing

Initially a pulse is recorded from the back face of the bearing bush when the shaft is lifted away so that there is no oil film present. This received pulse has then been reflected from an interface between steel and bulk oil. This is equivalent to the incident signal when divided by the reflection coefficient for a steel-oil interface. This signal becomes a reference signal to compare with all further reflected signals.

Reflected pulse from an oil film are captured, digitised, and passed to the PC. A fast Fourier transform (FFT) is performed on the pulse to give an amplitude against frequency plot. This is divided by the FFT of the reference signal to give a reflection coefficient spectra. Equations (1) and (2) are then

used to calculate the film thickness for each measurement frequency.

The full control, acquisition, and calculation algorithms are performed by bespoke Labview software.

4. RESULTS

Figure 8 shows a sample graph from the above experiment as it appears in the Labview software. It can be seen that the reflection coefficient is increasing with frequency. This frequency dependence is predicted by equation (1). When the film thickness decreases (as the load is applied) more of the signal is transmitted through the layer and the reflection coefficient drops.

Figure 9 shows the reflection data of Figure 8 processed according to equations (1) and (2) to give the film thickness.

Whilst the reflection spectra is frequency dependent the film thickness does not, of course, depend on the measurement of frequency. The data shows that the film thickness is indeed virtually constant with measurement frequency. This indicates that the spring model approach is valid in this case. Typically the value determined at the probe centre frequency (in this case 1.1 MHz) is presented as the calculated film thickness because this corresponds to the frequency possessing the maximum energy. Although the whole spectra is examined to check for consistency across all frequencies.

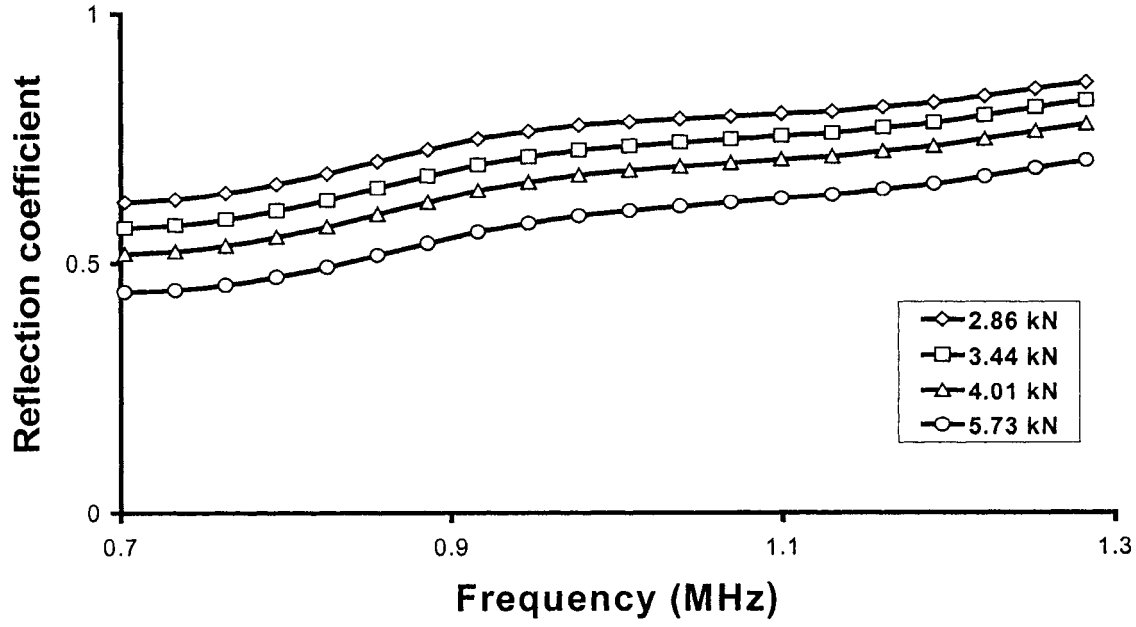


Figure 8. Reflection coefficient spectra measured from increasingly loaded hydrodynamic fluid films.

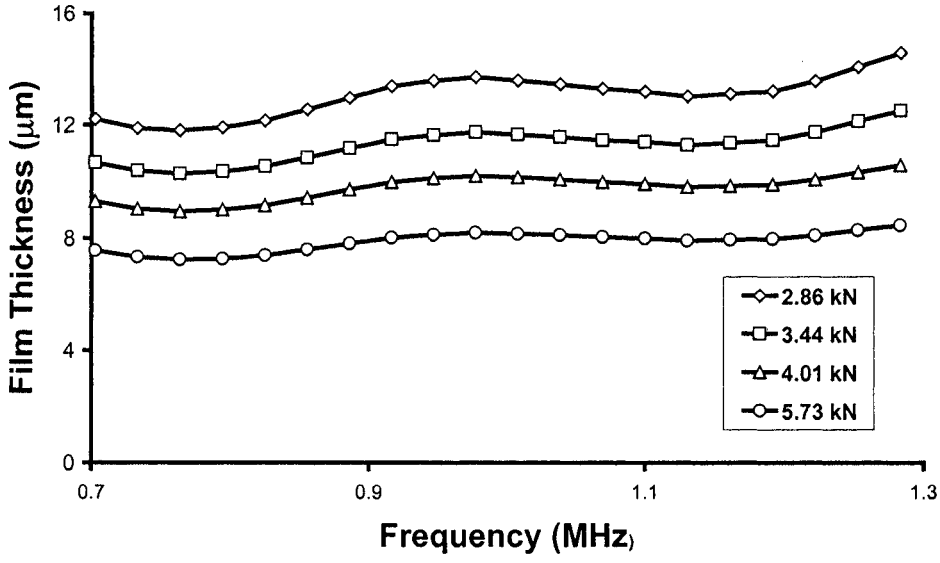


Figure 9. Lubricant film thickness spectra determined by the processing of reflection coefficient measurements.

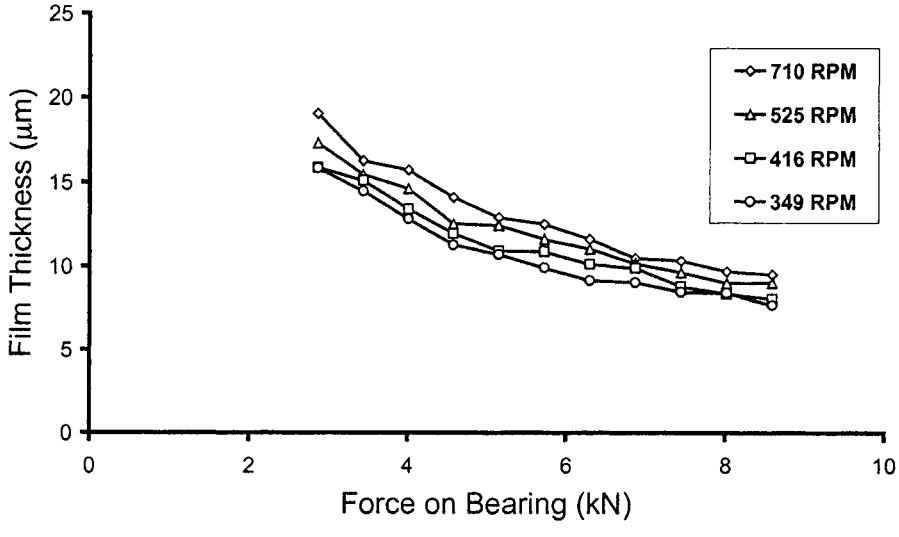


Figure 10. Measured lubricant film thickness from a journal bearing for a range of operating speeds and loads.

A series of such measurements has been performed on the journal bearing for a range of speeds and loads. The output is an array of film thickness values for each measurement frequency (similar to that shown in Figure 9). Software written in Labview control the signal acquisition and processing so that film thickness results are presented automatically. Figure 10 shows the measured film thickness results.

It can be clearly seen that the lubrication film is thinnest at lowest speed and highest load. With film thickness increasing with speed and decreasing with load. This accords with conventional hydrodynamic theory where increasing the rotational speed increases causes more oil to be drawn into the loaded region.

At this stage only measurements at the maximum load point have been recorded. The minimum film thickness in the bearing occurs at some position around the circumference in the direction of rotation. For this kind of steady state hydrodynamic bearing, lubricated with a Newtonian fluid, it is possible to determine the film thickness (and bearing attitude angle) for a numerical solution of Reynolds' equation (see for example 7). Look-up tables and charts are available relating the Sommerfeld parameter to the bearing eccentricity ratio.

For this geometry of bearing (L/D = ½) the Sommerfeld parameter varied from approximately 0.74 at high load and low speed to 5.2 at low load, high speed. The lubricant viscosity has been calculated at the appropriate bearing operating temperature. The theoretical minimum film thickness (from the published solutions in [7]) varies from 8 μm to 28 μm. The measured values in the range 8 to 19 mm are in good agreement (bearing in mind that they are not measures of the minimum but rather a point some distance away from the minimum).

5. CONCLUSIONS

Ultrasound reflects from a layer of oil between two bearing surfaces. The response of the oil film can be modelled using a quasi-static spring model. The proportion of the wave reflected depends on the stiffness of the oil film and its acoustic properties. Experiments have shown this spring model approach to be valid for the thickness of oil films typically encountered in the operation of hydrodynamic bearings.

An ultrasonic procedure has been developed to emit pulses of mixed frequencies and record reflection from an oil film. The signals are processed to give reflection coefficient spectra and lubricant film thickness.

An ultrasonic transducer has been mounted on a hydrodynamic journal bearing and the film thickness has been measured for a range of bearing operating conditions. The results are stable, repeatable, and compare well with published numerical solutions for film thickness.

6. REFERENCES

[1] Kendall, K. and Tabor, D., (1971), An ultrasonic study of the area of contact between stationary and sliding surfaces, *Proc. R. Soc. Lond.* A 323, 321-340.

[2] Drinkwater, B.W., Dwyer-Joyce, R.S. and Cawley, P., 1996, A study of the interaction between ultrasound and a partially contacting solid-solid interface, *Proc. R. Soc. Lond. A*, **452**, 2613-2628.

[3] Krolikowski, J. and J. Szczepek, (1991), Prediction of contact parameters using ultrasonic method, *Wear, 148*, 181-195.

[4] Tattersall, H.G., 1973, The Ultrasonic Pulse-Echo Technique as Applied to Adhesion Testing, *J. Appl. Phys.D.*, **6**, 819-832.

[5] Hosten, B., 1991, Bulk heterogeneous plane-wave propagation through viscoelastic plates and stratified media with large values of frequency-domain, *Ultrasonics*, **29**, 445-450.

[6] Dwyer-Joyce, R.S., Drinkwater, B.W., and Donohoe, C.J., (2002), *"The Measurement of Lubricant Film Thickness using Ultrasound"*, Proc. Roy. Soc. A., in press.

[7] Hamrock, B.J., 1994, Fundamentals of Fluid Film Lubrication, McGraw-Hill. Page 2 onwards.....

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

THD LUBRICATION OF MECHANICAL FACE SEALS DURING TRANSIENT PERIOD AFTER START-UP : 2D MODELISATION

Danos J.C.[a] Tournerie B.[a], Frêne J.[a]

[a] Laboratoire de Mécanique des Solides – Université de Poitiers – France

Abstract

A numerical model of the THD behaviour of an aligned face seal in transient regime was developed in order to study the viscous heat dissipation in the interface and the thermal transfers with the seal environment during start-up.

The Reynolds equation and the energy equation in the interfacial film and the conduction equation in the seal rings are solved numerically by the finite difference technique. These equations are coupled by the heat exchange conditions on the boundaries of their domains. They are integrated by using the Crank Nicholson scheme.

This study also compares the results of the calculation code and those of existing 1D models. Good correlation is found and some simplifying assumptions are validated. The results show the presence of a peak of the dissipated power just after start-up. The transient evolution of the temperature field in the seal is described and the influence of dissipation and heat transfer parameters is examined.

1. INTRODUCTION

Mechanical face seal operation is greatly affected by thermal phenomena that provoke fluid viscosity variation, friction face distortions, and, eventually vaporisation in the interfacial fluid film. These effects also interfere with the seal dynamic behaviour. More particularly, during transient periods such as start-up, these phenomena can lead to irreversible instabilities.

The simultaneous analysis of all these interacting effects is complex. Firstly, we intend to examine only the effect of the lubricant viscosity variation due to the temperature evolution versus time, i.e. the transient thermohydrodynamic (THD) behaviour.

Most of the THD theoretical studies and models deal with the steady state regime. Dumbrava and Morariu [1] developed a model based on limiting assumptions but one which lead to an analytical solution. This model was improved by Pascovici et al. [2,3] and then by Cicone [4]. However, the analytical solution proposed by Buck [5] produces more realistic results as shown by Etsion and Groper [6].

Other methods consist in formulating analytically the heat dissipated in the film and to solve numerically, generally by using commercially available codes, the heat conduction equation in the rings. Most of these studies also consider the seal ring distortions in thermoelastohydrodynamic (TEHD) analyses (Metcalfe; [7], Li [8], Salant and al. [9,10], Zeus [11], Zhu [12]).

Some authors solve numerically the 3D governing equations of the THD problem in the film and in the ring by using finite difference or finite element techniques (Knoll et al. [13], Person et al.[14], Tournerie et al. [15])

Studies concerning the transient regime are less numerous. Pascovici et al. [16, 17] have developed a 1D model leading to an analytical solution when the heat flux dissipated in the film varies linearly versus time. Cicone [4] has solved numerically the 1D equation of conduction in the rings and has taken into account the fluid viscosity variation versus temperature.

There have been very few attempts to develop more elaborate models. These are studies on TEHD seal behaviour using commercial finite element codes to determine the field of temperature and displacements in the rings. The heat dissipated in the interface is obtained experimentally or calculated as a function of the film thickness and face taper by Doust and Parmar [18, 19]. The model of Blasbag, Harp and Salant [20, 21] is similar but it also takes into account asperity contact and phase change.

This paper presents a numerical model of the THD behaviour of an aligned face seal in transient regime developed by Danos [22]. The study

examines the evolution of the field of temperature, in the film and in the rings, in relation to the conditions of viscous heat dissipation in the interface and of thermal transfers with the seal environment during start-up.

2. THEORY

2.1. Geometric and kinematic model

The model is shown on figure 1. The faces of the stator 1 and of the rotor 2 are aligned with the axis of rotation and are tapered.

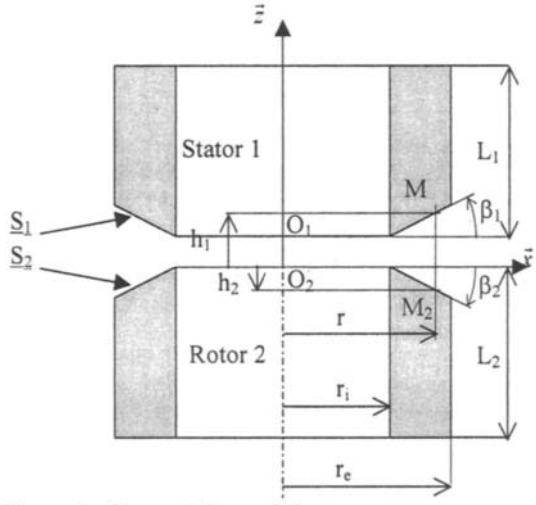


Figure 1: Geometric model

The taper angles are noted respectively β_1 and β_2. Points M_1 and M_2 are placed respectively on the stator and on the rotor face. In a reference *R* fixed on the frame, the co-ordinates and the velocities of M_1 and M_2 are in a local system:

$$\left(\vec{O_2M_1}\right)_R\left|\begin{matrix} r \\ h_1 = \ell + (r - r_i)\beta_1 \end{matrix}\right., \left(\vec{V}_{(M_1)}\right)_R\left|\begin{matrix} V_{r1} = 0 \\ V_{\theta 1} = 0 \\ V_{z1} = 0 \end{matrix}\right.$$
$$\left(\vec{O_2M_2}\right)_R\left|\begin{matrix} r \\ h_2 = (r - r_i)\beta_2 \end{matrix}\right., \left(\vec{V}_{(M_2)}\right)_R\left|\begin{matrix} V_{r2} = 0 \\ V_{\theta 2} = \omega r \\ V_{z2} = 0 \end{matrix}\right. \quad (1)$$

with $\ell = \left\| \vec{O_2O_1} \right\|$ and ω is the shaft angular velocity.

2.2. Governing equations and boundary conditions

2.2.1 Velocity field

The film is not isothermal. It is incompressible, its density does not vary with the temperature, but the dynamic viscosity $\mu(r,z,t)$ is dependant on the local temperature $T(r,z,t)$ at each time.

For the above face velocities (1) the field of velocities in the lubricant film is described by the following equations:

$$\begin{cases} v_r = \dfrac{\partial p}{\partial r}\left(I - \dfrac{I_1}{J_1}J\right) \\ v_\theta = \omega r\left(1 - \dfrac{J}{J_1}\right) \\ v_z = -\dfrac{1}{r}\displaystyle\int_{h_2}^{z}\left\{\frac{\partial(r\,v_r)}{\partial r}\right\}d\xi + cste \end{cases} \quad (2)$$

with :

$$I = \int_{h_2}^{z}\frac{\xi d\xi}{\mu(r,\theta,\xi,t)} \text{ and } J = \int_{h_2}^{z}\frac{d\xi}{\mu(r,\theta,\xi,t)} \quad (3)$$
$$I_1 = I(h_1) \text{ and } J_1 = J(h_1)$$

2.2.2 Reynolds equation

For aligned faces the Reynolds equation is of a simple form:

$$\frac{\partial}{\partial r}\left(r\,G_1\frac{\partial p}{\partial r}\right) = 0 \quad (4)$$

with $$G_1 = \int_{h_2}^{h_1}\frac{\rho(z - H_2)}{\mu}\left(z - \frac{I_1}{J_1}\right)dz \quad (5)$$

2.2.3 Energy equation

The energy equation in the lubricant film is as follows:

$$\rho_f C_f\left\{\frac{\partial T}{\partial t} + v_r\frac{\partial T}{\partial r} + v_z\frac{\partial T}{\partial z}\right\} = \mu\left\{\left(\frac{\partial v_r}{\partial z}\right)^2 + \left(\frac{\partial v_\theta}{\partial z}\right)^2\right\} + K_f\left\{\frac{\partial^2 T}{\partial z^2}\right\} \quad (6)$$

where the LHS term corresponds to the heat transfer by convection and the RHS terms correspond respectively to the heat viscous dissipation and of the heat transfer by conduction across the film thickness.

The equation of conduction in the seal rings is of the form:

$$\left(\frac{\partial T}{\partial t}\right)=\frac{K_i}{\rho_i C_i}\left\{\frac{1}{r}\frac{\partial T}{\partial r}+\frac{\partial^2 T}{\partial r^2}+\frac{\partial^2 T}{\partial z^2}\right\} \qquad (7)$$

where $d_i=\frac{K_i}{\rho_i C_i}$ is the thermal diffusivity.

2.2.4 Boundary and initial conditions

On the inner and outer radial boundaries of the interfacial film, pressure is imposed.

On friction faces S_1 and S_2, there is conservation of the temperature and of the heat flux transmitted by conduction from the fluid film to the rings.

On the other surfaces of the rings, different types of boundary conditions may be considered. In this study only condition of heat transfer by convection with the inner fluid is considered:

$$-K_i\left.\frac{\partial T}{\partial r}\right|_{r_i}=H_i\{T(r_i)-T_0\}\ ,\ i=1,2 \qquad (8)$$

where K_i and H_i are respectively the thermal conductivity and the convection coefficient related to the *i-th* ring and T_0 is the feeding temperature chosen as the reference value.

The other surfaces are considered adiabatic.

At the initial time t = 0 the temperature in the film and in the rings is equal to T_0.

2.3. Shaft velocity variation

The time law of shaft velocity during start-up is of the form:

$$\omega(t)=\omega_0(1-e^{-7t}) \qquad (9)$$

That gives $\frac{\omega}{\omega_0}>0.99$ for t = 0.66s and $\frac{\omega}{\omega_0}>0.9999$ for t = 1s.

2.4. Numerical solution

The governing equations are discretized by using the finite difference method. The Reynolds equation solution leads to solving a tri-diagonal system of linear equations. Using the semi-implicit scheme of Crank-Nicholson solves the energy equations in the film and in the rings.

3. RESULTS

3.1. Comparison with 1D model

In order to validate the model, a comparison with one-dimensional (1D) transient models, chosen for the completeness of the published results, has been made. Comparison will be made between the calculated values of temperature and heat flux of 1D models and those calculated with the 2D model on the mean radius.

3.1.1 Review of 1D model

Pascovici et al. [16, 17] and Cicone [4] have presented different versions of a basic transient model for an externally pressurised, aligned, liquid film face seal. The main assumptions are:

- the heat flux on the seal faces is uniformly distributed,
- all the heat dissipated by friction drag in the seal dam is transferred by conduction through the rings and then by convection with the surrounding fluid inside the rings,
- the heat transfer in the rings is axially predominant so the rings are modelled as fins,
- the maximum shaft rotational speed is reached instantaneously at time t = 0.

Three solutions corresponding to different time variations of the heat flux at the rotor face were considered.

In the first analytical solution the heat flux is assumed to be constant and equal to the flux in a steady state regime.

In the second analytical solution, during the transient period, the flux is assumed to vary linearly between two values corresponding to the film temperature at the initial time and at the steady state respectively.

In the 1D numerical model, the heat flux variation is due to the exponential law of viscosity variation versus film temperature identical to the one used in the 2D numerical model (10).

3.1.2 Data of the studied cases

The constructive and the functioning parameters of the studied seals are those of Cicone [4] which are reported in the tables 1 and 2.

Except for the results in figure 4, all results presented in this paper were obtained with oil as lubricant.

It must be recalled that heat transfer by convection is restricted to the inner surfaces.

The viscosity-temperature law is of exponential form:

$$\mu(T) = \mu(T_0)e^{-B(T-T_0)} \qquad (10)$$

Parameters	Rotor	Stator
Outer radius r_e (m)	$50{,}5.10^{-3}$	
Inner radius r_i (m)	$46{,}5.10^{-3}$	
Material	Steel	Graphite
Thermal conductivity K_i (W.m^{-1}.K^{-1})	50	15
Density ρ (Kg.m^{-3})	7800	1800
Specific heat C_i (J.Kg^{-1}.K^{-1})	470	670
Thermal diffusivity $K_i/\rho_i C_i$ (m^2.s^{-1})	$13{,}6.10^{-6}$	$12{,}4.10^{-6}$
Length L_i (m)	50.10^{-3}	20.10^{-3}

Table 1: Seal ring parameters

Parameters	Water	Oil
Density ρ (Kg.m^{-3})	1000	890
Specific heat C_f (J.Kg^{-1}.K^{-1})	4180	1980
Thermal conductivity K_f (W.m^{-1}.K^{-1}) at T_0	0,65	0,14
Feeding Temperature T_0 (°C)	50	
Convection coefficient (on the inner surfaces): - stator H_{1i} (W.m^{-2}.K^{-1}) - rotor H_{2i} (W.m^{-2}.K^{-1})	500 1000	
Dynamic viscosity $\mu(T_0)$ (Pa.s)	$562{,}7.10^{-6}$	$141{,}4.10^{-6}$
Temperature-viscosity coefficient B (K^{-1})	$4{,}7.10^{-3}$	45.10^{-3}
Shaft angular speed ω_0 (rad.s^{-1})	400	100
Minimum film thickness $h(r_e)$ (m)	1	

Table 2: Lubrication parameters

3.1.3 The stator is insulating

Initially the model may be simplified assuming an adiabatic condition on the stator face S_1. This is equivalent to considering a stator made of heat insulating material.

Figures 2 and 3 show the evolution during start-up of the heat flow and the temperature on the rotor face respectively, when the lubricant is oil, from the three solutions of the 1D model and from the 2D numerical model. As is illustrated on figure 2, the discrepancies in the results during start-up are due to the differences of the law of heat flux variation versus time.

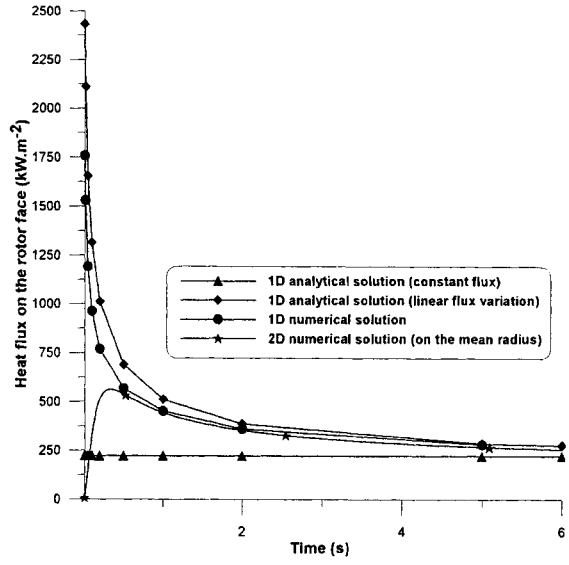


Figure 2: Heat flux variation on rotor face during start-up.

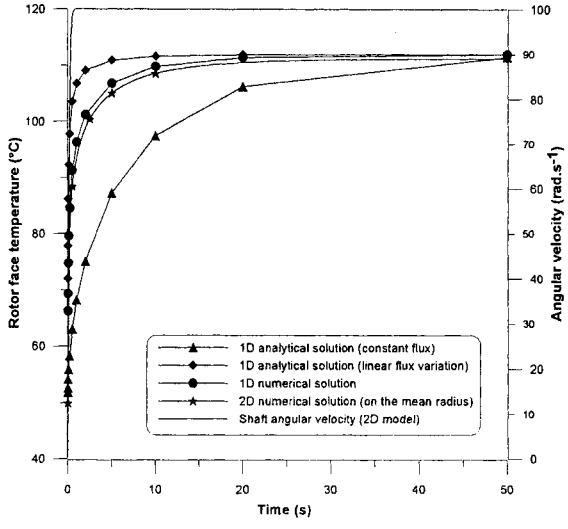


Figure 3: Rotor face temperature variation, on the mean radius, during start-up.

In particular a peak of heat flux only appears for the 2D model where the shaft speed continuously increases (9). The power dissipated by shear stress first increases with the increasing rotational speed, then, when the speed becomes constant, the temperature begins to grow, leading to a fall off in viscosity and heat dissipation.

The 1D analytical solutions with linearly variable flux or constant flux lead respectively to the

overestimation and the underestimation of both flux and temperature.

1D and 2D numerical solutions are very similar. The only difference is due to the time law variation of the shaft speed during the first second.

3.1.4 The stator is conducting

In thispart of the study, the stator is made of graphite and the heat dissipated in the interfacial film is transferred both through the rotor and the stator. Figure 4 shows the temperature variation of the rotor and stator face, which have been established by Cicone [4] only for water. The values obtained by analytical and numerical solutions are in a good agreement. From 1D and 2D numerical solutions the face temperature of rotor and stator are practically equal. This implies that the calculation of integrals I and J (3) may be simplified by considering a constant dynamic viscosity across the film thickness as assumed by Pascovici and Etsion [2, 3].

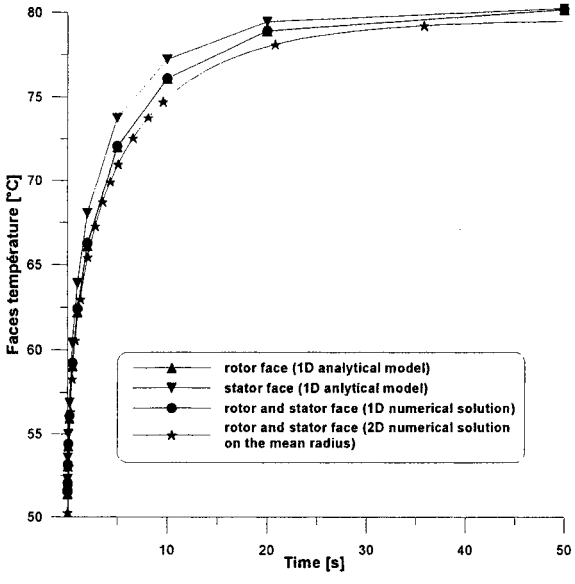


Figure 4: Temperature variation of the ring faces (Water)

The heat dissipated in the film is shared out between the two rings. One defines a partition ratio $\lambda_i = \frac{Q_i}{Q}$ where Q_i is the heat transmitted to the ring i and Q is the total heat generated in the interface. The curves of figure 5 give the values of these ratios calculated by means of the 1D and 2D solutions during start-up when the fluid is oil. The results are well correlated; the discrepancy is mainly due to the time law of the heat flux during the first second. As shown in the analysis of Cicone [4] the repartition coefficients do not depend on the nature of the lubricant fluid but on the conditions of transient heat transfer by conduction in the rings and by convection with the surroundings.

3.2. Study of the transient THD behaviour

The 2D numerical model enables temperature mapping versus time both in the film and in the rings. It also permits one to analyse the influence of the parameters governing the heat dissipation in the film and the heat transfers through the rings to the surrounding fluids.

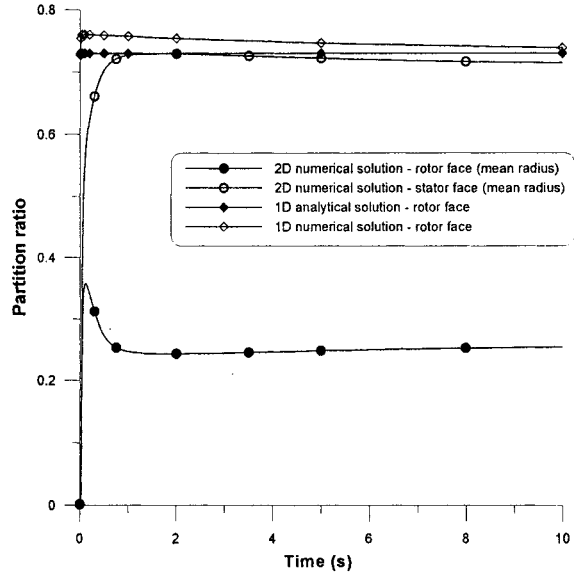


Figure 5: Heat flux partition ratio.

3.2.1 Time evolution of the transient temperature fields

Figure 6 shows a temperature mapping in a cross-section of the rings and of the lubricant film at time t = 0,5 s when steady state is reached. The heat transfer by convection through the feeding fluid is localised on the zero value of the normalised radius. The faces are plane and parallel. The temperature in the fluid film increases along the radial width in the direction of the flow. The higher temperature value is at the outer circumference of the dam, where there is the maximum sliding speed. The temperatures of the ring faces are practically equal, however there is

a slight variation in the middle of the film thickness. In the rings the isothermal lines are almost straight and perpendicular to the axis. In this case, the temperature field is practically one-dimensional.

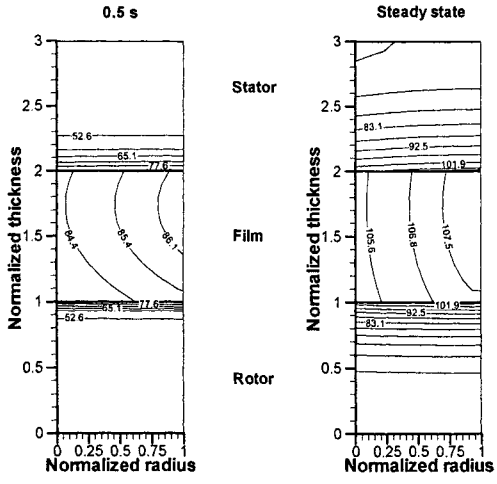


Figure 6: Temperature field variation in the film and in the rings for plane faces β = 0

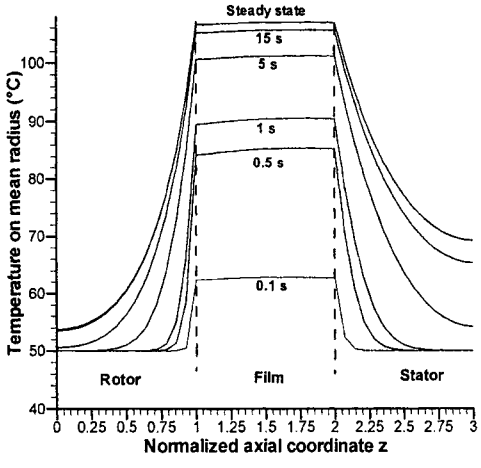


Figure 7: Evolution of the axial variation of the temperature on the mean radius during start-up.

Figure 7 shows the evolution during start-up of the temperature along an axis placed on the mean radius.

In the followin, this case will be considered as the reference case.

3.2.2 Heat dissipation in the lubricant film

In the 2D model, the heat flux dissipated by friction drag is given by the equation:

$$q(r) = \omega^2 \frac{r^2}{I_1(r)} \qquad \text{with } I_1(r) = \int_{h_2(r)}^{h_1(r)} \frac{d\xi}{\mu(r,\xi)} \# \frac{h(r)}{\mu_m(r)} \tag{11}$$

Obviously, the heat flux is an increasing function of the viscosity μ and of the angular velocity ω. But the heat flux q is also proportional to r^2 and the integral I_1 depends on r through the variation of the film thickness $h(r) = h_1(r) - h_2(r)$ due to the face taper β_1 and β_2. The film thickness variation along the radial width is characterised by the ratio

$$\frac{\Delta h}{\ell} = \frac{|h(r_e) - \ell|}{\ell} = \frac{|(r_e - r_i)\beta|}{\ell} \quad \text{with } \beta = \beta_1 - \beta_2 \tag{12}$$

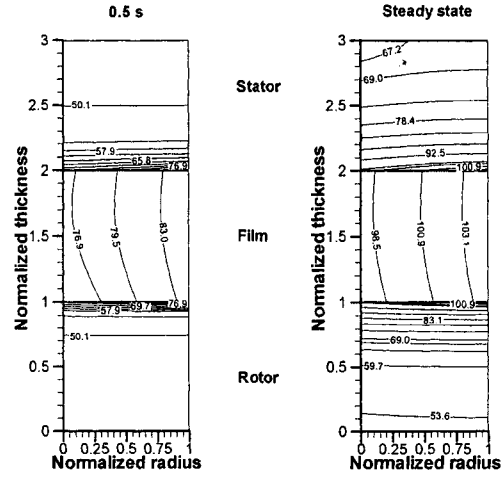


Figure 8: Temperature field variation in the film and in the rings for β = 0.25 10^{-3}rad.

Figures 8 and 9 show the temperature maps obtained respectively for β = 0.25 10^{-3}rad. and β = 0.75 10^{-3}rad. which correspond to a thickness variation ratio of $\frac{\Delta h}{\ell} = 1$ and $\frac{\Delta h}{\ell} = 3$ respectively.

By comparison with the case of flat faces (figure 6) it is shown that the more the taper is accentuated, the more the temperature varies along the radial width. However, it is interesting to note that although the heat flux is far from uniform on the dam width, the isothermal lines in the rings remain very close to the pattern of the 1D fin model.

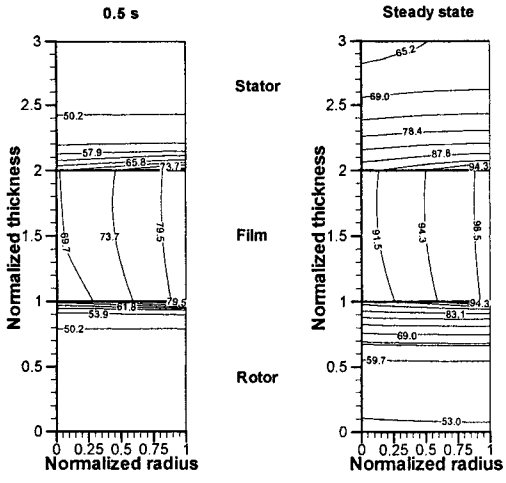


Figure 9: Temperature field variation in the film and in the rings for $\beta = 0.75\ 10^{-3}$rad.

3.2.3 Heat transfer through the rings

As shown in the dimensional analysis of the 1D models [2-4] transient heat transfer through the rings is characterised by two non-dimensional factors:

- the time parameter

$$a = \frac{d}{L^2} = \frac{1}{L^2}\left(\frac{k}{\rho C}\right) \tag{13}$$

- the heat transfer parameter

$$m = \sqrt{\frac{2H}{\frac{k}{r_i}} \frac{L^2}{(r_e^2 - r_i^2)}} \tag{14}$$

The time parameter a depends on thermal diffusivity and on the length of the ring. The larger a is, the better the ability of the materials to propagate heat flux.

When the ability to transfer the heat by convection with the surrounding fluid increases, m increases. When the ability to transfer heat by conduction in the ring increases, m decreases. Factor m contains two terms:

- one is the ratio of the heat transfer coefficients by convection and by conduction:

$$b = \frac{2H}{\frac{k}{r_i}} \tag{15}$$

- the other is the ratio of the respective characteristic areas of the transfer by convection, and of the transfer by convection:

$$c = \frac{L^2}{(r_e^2 - r_i^2)} \tag{16}$$

In the following, three typical cases will be compared to the reference case.

The first case deals with new materials in order to examine the influence of thermal diffusivity (parameter a) and the thermal conductivity (parameter m and ratio b).

The second case consists of the reference case materials but with half ring length in order to examine the influence of the convection transfer area (parameter m and ratio c).

In the third case the rings are those of the reference case but the convection coefficient are three time greater (parameter m and ratio b)

Figures 10, 11, 14 and 16 show the temperature mapping in a cross-section of the oil film and of the rings respectively for the three cases.

Figure 12 shows the variation of face temperature versus time.

This study is limited in THD aspects, but it is well known (Doust and Parmar [23]) that the main thermoelastic distortion of the face is a toroïdal rotation due to the temperature gradient along the ring. The authors refer to this as "primary thermal rotation". It is therefore particularly pertinent to examine, in the following, the time variation of the difference in temperature between the two end-faces of the ring. Figure 13 and 15 show the difference in temperature, on the mean radius, at the end-faces of the rings.

3.2.3.1 The Influence of the thermal diffusivity of the materials

Table 3 presents the characteristics of ceramics that seal rings are usually made of. In comparison with the reference materials (table 1), the tungsten carbide density is near twice the steel density, moreover, the thermal diffusivity d_2 of the rotor is divided by 2.5. Silicon carbide has a conductivity that is ten time that of graphite; its density is greater as well, and the thermal diffusivity d_1 of the stator is multiplied by 4.

Parameters	Rotor	Stator
Material	Tungsten carbide	Silicon carbide
Thermal conductivity K_i (W.m^{-1}.K^{-1})	65	150
Density ρ (Kg.m^{-3})	14800	3100
Specific heat C_i (J.Kg^{-1}.K^{-1})	800	1000
Thermal diffusivity $K_i/\rho_i C_i$ (m^2.s^{-1})	5.5 10^{-6}	48.4 10^{-6}

Table 3: Characteristics of a new couple of ring materials.

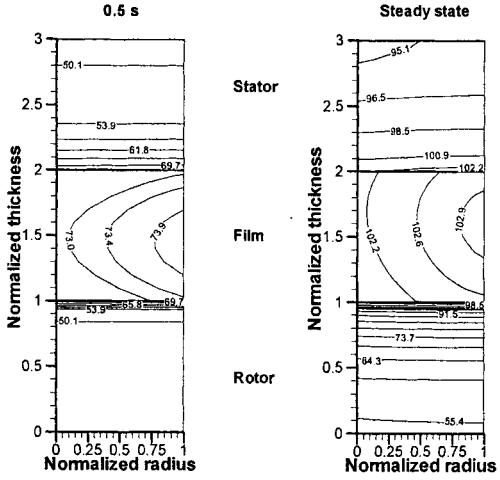


Figure 10: Temperature field variation in the oil film, Tungsten carbide rotor and silicon carbide stator.

By comparison with the reference case (figure 6), it is found on figures 10, 12 and 13, that the larger heat conductivity of the silicon carbide reduces the axial temperature gradient in the stator and increases the mean value of the temperature along the whole ring length. This enables to evacuate more heat resulting for smaller temperature level in the film.

The main effect of the four-fold thermal diffusivity is to accentuate the peak in axial temperature gradient during start up. This is a cause of a temporary distortion, but possibly leading to an irreversible unwanted reaction and will have to be thoroughly examined.

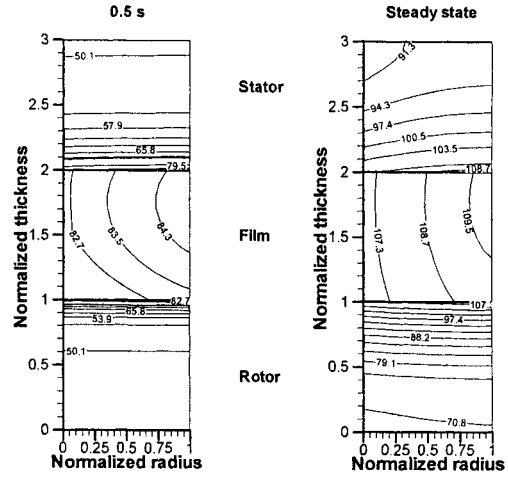


Figure 11: Temperature field variation in the oil film, steel rotor and graphite stator but with half length.

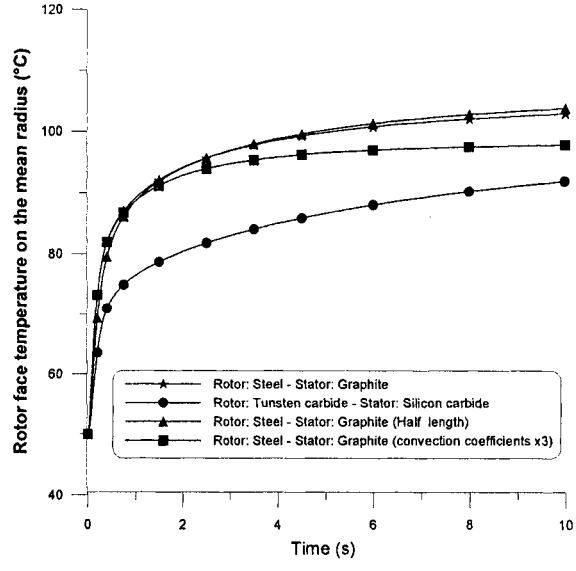


Figure 12: Face temperature variation versus time.

3.2.3.2 Influence of rings length and area ratio

In the reference case the stator and rotor length are L_1 = 20mm and L_2 = 50mm respectively, the ratio of effective areas of convection transfer and conduction transfer are $c_1 = 1.031$ and $c_2 = 6.443$. Halving the ring length also reduces by half the effective areas of heat transfer by convection, and

also divides the area ratios by four; $c_1 = 0.256$ and $c_2 = 1{,}611$.

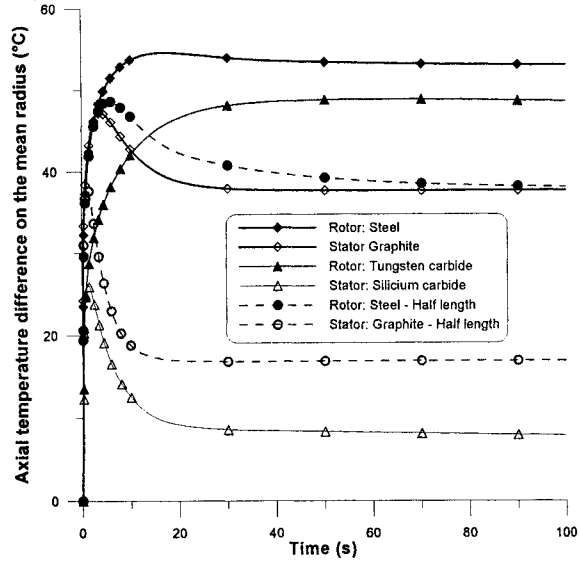


Figure 13: Axial temperature variation versus time

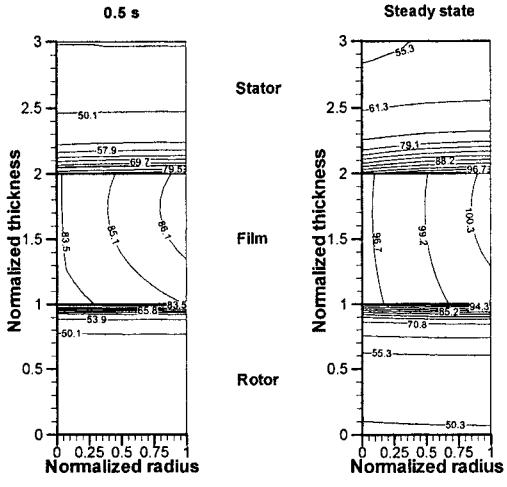


Figure 14: Temperature field variation in the oil film, steel rotor and graphite stator, for three time larger convection coefficient.

This leads to a higher mean temperature level in the rings and a sensibly reduced temperature difference between the two end faces (figures 13 and 14). This produces an increased convection heat flux.

At the same time the isothermal lines in the rings become slightly inclined which produces an increased radial conduction heat flux.

The effect of reducing the ring length by half, is to produce a peak of axial temperature gradient (figure 13), which is all the more pronounced since the length is reduced, and thus susceptible to produce an undesired consequence.

3.2.3.3 Influence of the heat transfer coefficient ratio

The thermal mapping of figure 6 is obtained for convection coefficients on the stator and rotor inner surfaces of $H_1 = 500\text{W/m}^2\text{K}$ and $H_2 = 1000\text{W/m}^2\text{K}$ respectively. The graphs of figure 14 and 16 are obtained for a convection coefficient which is three times greater; $H_1 = 1500\text{W/m}^2\text{K}$ and $H_2 = 3000\text{W/m}^2\text{K}$. As expected, it is found that the level of temperature in the film and in the rings decreases. But as shown on figures 12 and 15, the effect of a strong surrounding fluid cooling is to shorten sensibly the the transient period and above all to avoid the transient peak of the axial temperature gradient.

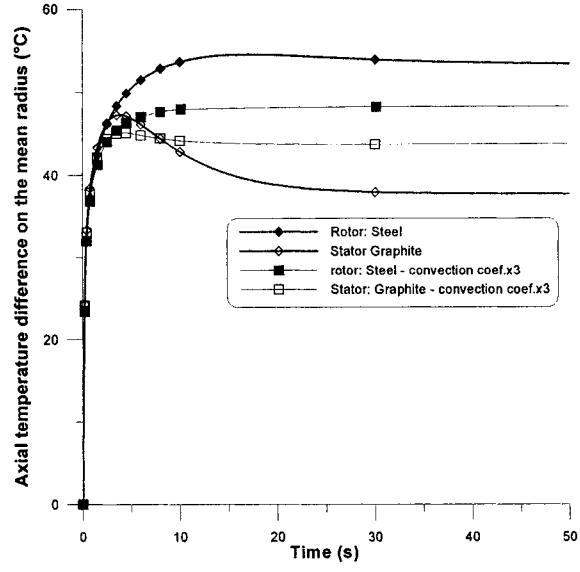


Figure 15: Axial temperature variation versus time

The figure 16 shows a thermal mapping in a cross section of the rings, drawn at the real scale. In this case, the heat flux transferred by convection is quite large resulting in a particularly important radial temperature gradient. However it is found that the

isothermal line pattern is not very far from a one dimensional variation along the axial direction. That justifies the fin model for determining the face rotation which is not very much influenced by the radial temperature gradient in the ring.

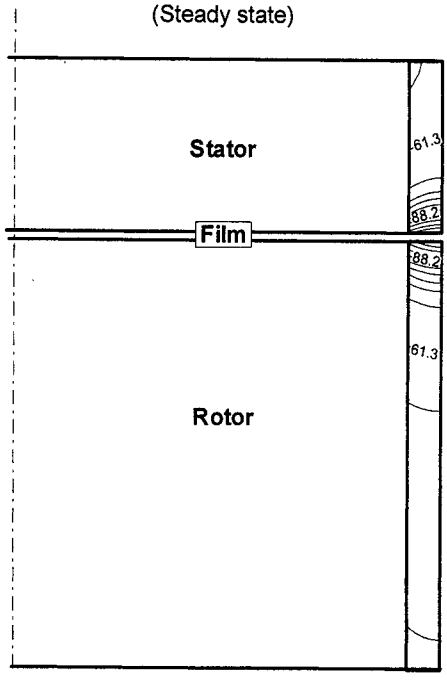


Figure 16: Example of isothermal lines patern in rings for a strong convection heat transfer

4. CONCLUSION

This study is solely an analysis of the thermal behaviour of face seals during start up independently of ring distortions. This is quite unrealistic, since it supposes thermal expansion coefficient of the materials small enough to be neglected. However this work is an exploratory analysis and the THD model is preliminary to a more complete TEHD-2D model which will take into account the thermoelastic distortions of the rings.

The 2D-THD study has pointed out the presence of a peak in dissipated heat flow and in ring axial temperature gradient. This behaviour, which may lead to irreversible thermal instabilities, will have to be examined with attention in the TEHD study.

The 2D-THD study has also permitted us to show that the temperature distribution in the rings is practically one dimensional in most cases, either in transient period or in steady state, which validate the fin model.

The influences of parameters governing viscous heat dissipation, heat transfer by conduction and by convection, while considering the time-dependant aspects during transient period, have been examined. The time evolution of the axial difference in temperature, which is responsible for the "primary rotation", has been given particular attention. Possible causes of thermal instabilities have been underlined.

NOMENCLATURE

C	specific heat [J.Kg^{-1}.K^{-1}]
L_i	length of the rings (i = 1,2) [m]
H_i	convection coefficient (i = 1,2) [W.m^{-2}.K^{-1}]
I_1, J_1,G_1	integrals of functions of the viscosity
K	thermal conductivity [W.m^{-1}.K^{-1}]
ℓ	centerline clearance [m]
M_i	points on the seal face (i = 1,2).
r, h_i	coordinates of M_i in the inertial reference system (i = 1,2) [m,m]
Q	thermal power [W]
p	pressure in the lubricant film [Pa]
S_i	boundary surfaces separating the film and the rings (i = 1,2) (Figure 1.)
T	temperature [K]
t	time [s]
$\vec{V}(M_i)$	velocity of the seal faces (i = 1,2) [m.s^{-1}]
$V_{ri}, V_{\theta i}, V_{zi}$	face velocity components in the local system (i = 1,2) [m.s^{-1}]
v_r, v_θ, v_z	fluid velocity components in the local system [m.s^{-1}]
$\vec{z}$	shaft rotation axis
β_i	taper angles of the seal faces, (i = 1,2) [rad]
λ_i	heat partition ratio
μ	dynamic viscosity [Pa.s]
ρ	fluid density [Kg.m^{-3}]
ω	shaft angular velocity [rad.s^{-1}]
$(\vec{e}_r, \vec{e}_\theta, \vec{e}_z)$	local system associated with the cylindrical coordinates

$(O_2, \vec{x}, \vec{y}, \vec{z})$ system axis of the inertial reference R

Subscripts

0	reference value
1	stator
2	rotor
i	on the inner circumference of the sealing dam
e	on the outer circumference of the sealing dam
f	fluid film
m	mean value

REFERENCES

1. Dumbravă M.A., Morariu Z. – "Thermo-Hydrodynamic Aspects of the Double Mechanical Seals", Proc. Of the 11-th Int. Conf. on Fluid Sealing (BHRA), Paper F4, Cannes, France, April 1987.
2. Pascovici M. D., Etsion I. – "A Thermo-hydrodynamic Analysis of a Mechanical Face Seal, Trans. ASME – Journal of Tribology, Vol. 114, N° 4, pp. 639-645, 1992.
3. Etsion I., Pascovici M.D. – "A Thermo-hydrodynamic Analysis of a Misaligned Mechanical face Seal, STLE Tribology Transaction, Vol. 36, N°4, pp. 589-596, 1993.
4. Cicone, T. - Study of thermohydrodynamic problems in fluid film mechanical face seals", PhD. Thesis, University Polytehnica of Bucarest, 1997.
5. Buck G.S. – "Heat Transfer in Mechanical Seals", 6th Int. Pump Users Symp., April 1989, Huston, Texas.
6. Etsion I., Groper M. – "The Accuracy of Analytical Solutions for the Temperature Distribution in Mechanical Seals", Proc. Of the 14th Int. Conf. on Fluid Sealing, BHRA, pp.341-350 Firenze, Italy, April 1994.
7. Metcalfe R. – "Performance Analysis of Axisymmetric Flat Face Mechanical Seals", AECL-4432, Report for Chalk River Nuclear Laboratories, 1973.
8. Li C-H. - "Thermal deformation in a mechanical face seal", ASLE Transactions, 19(2), pp146-152, 1976.
9. Salant R.F., Key W.E. – "Development of an Analytical Model for Use in Mechanical Seal Design, Proc. of the 10th Int. Conf. on Fluid Sealing, BHRA, Paper G3, Insbruck, Austria, April, 1984.
10. Salant R.F., Giles O., Key W.E. – "Design of Controllable Mechanical Seals", 15th Leeds-Lyon Symp. on Tribology, Paper III(i), 1988
11. Zeus D. – "Viscous Friction in Small Gaps - Calculations for Non-Contacting Liquid or Gas Lubricated End Face Seals", STLE Tribology Trans., Vol. 33, No 3, pp. 454-462, 1990.
12. Zhu G. - "Computer Modeling of Mechanical Seal Performance and Experimental Validation", I Mech, Vol. 213, Part J, 1999.
13. Knoll G., Peeken, Höft H-W. - "Thermo-hydrodynamic calculation of end face seals", Proc. of the 14th International Conference on Fluid Sealing, BHRA, pp. 367-383, Firenze, Italy, April 1994.
14. Person V., Tournerie B., Frêne J. - "THD aspects in misaligned and wavy face seals", 15th Int. Conf. On Fluid Sealing, BHRG, pp. 505-519, Maastricht, The Netherlands. September 1997
15. Tournerie B., Danos J.C., Frêne J. - "Three-dimensional modeling of THD lubrication in face-seals", ASME Journal of Tribology, Vol. 123, pp. 196-204, 2001.
16. Pascovici, M.D., Olariu, V. and Cicone, T. – "Transient Thermohydrodynamic Analysis of a Mechanical Face Seal. Part I:The Case of Insulated Stator"., 6th Intern. Conf. VAREHD-6, Suceava, Romania, June 1992.
17. Pascovici, M.D., Cicone, T. et Kucinschi, B. – "Transient Thermohydrodynamic Analysis of a Mechanical Face Seal. Part II: The Apportionment of the Heat Flux", Proceedings of Micro-CAD '94 - International Computer Science Conference -Vol. 2, Section K2, p. 6-13, March 3, 1994, Miskolc, Ungaria.
18. Doust T.G., Parmar A. - "Transient thermoelastic effects in a mechanical face seal", Proc. of the 14th International Conference on Fluid Sealing, BHRA, paper F4, pp. 407-422, Cannes, France. April 1987.
19. Parmar A. – "Thermal cycling in mechanical seals – Causes, prediction, prevention", Proc. of the 13th International Conference on Fluid Sealing, BHRA, pp. 507-526, Brugge, April 1992.
20. Blasbalg D.A., Salant R.D. – "Numerical Study of Two-Phase Mechanical Seal Stability", STLE Tribology Trans., Vol. 38, No 4, pp. 791-800, 1995.
21. Harp S.R., Salant R.F. - 1997 - Analysis of Mechanical Seal Behavior During Transient Operation, ASME Paper, No 75-Lub-22.

22. Danos J.C., 2000, "Lubrification thermohydrodynamique dans les joints d'étanchéité à faces radiales", Thèse de Doctorat de l'Université de Poitiers, France.
23. Doust T.G., Parmar A. – "An Experimental and Theoretical Study of Pressure and Thermal Distortions in a Mechanical Seal", ASLE Trans., Vol. 29, No 2, pp. 151-159, 1986.

SESSION XIV
ENGINE TRIBOLOGY (1)

Chairman :	Dr A Jackson
Paper XIV (i)	**Modelling Lubricant Related Fuel Economy in Heavy Duty Diesel Engines**
Paper XIV (ii)	**A New Tribometer for Piston Ring Friction Measurements**
Paper XIV (iii)	**The Wear Response of Ceramic Matrix Nano-composite Coatings in Simulated Piston-Ring/Cylinder Wall Test Environments**
Paper XIV (iv)	**An Initial Investigation of the Lubricant Condition in the Automotive Ring Zone Under Cold Start Conditions**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Modelling Lubricant Related Fuel Economy in Heavy Duty Diesel Engines

R. Castle, S. Arrowsmith

Infineum UK Ltd, Milton Hill Technology Centre, PO Box 1, Abingdon, OX13 6BB, UK

Current and future legislation controlling the exhaust emissions from Heavy Duty Diesel (HDD) vehicles requires an overall improvement in engine efficiency as well as the use of specialised after treatment systems to remove harmful species from the exhaust gas, and prevent these from entering the environment. Engine efficiency can be improved in a number of ways, and a combination of all approaches is required to ensure that the desired levels can be achieved. One approach is to improve the fuel efficiency of crankcase lubricants to optimise fuel consumption and minimise the amount of harmful combustion gas generated. In addition to the environmental benefits, reductions in overall fuel consumption in heavy duty diesel vehicles will represent significant cost savings to transport fleet operators.

It is well known that vehicle fuel economy can be improved significantly by using oils of lower viscosity grades. More recently it has been shown that improvements in HDD fuel economy can also be achieved by optimisation of the lubricant additive package[1], and that the appropriate selection of lubricant viscometrics and additive package can generate improvements in fuel economy without compromising wear performance.

In this study, the impact of both lubricant viscometrics and additive type and treat on HDD fuel economy has been evaluated in a series of laboratory bench tests which measure friction under different lubrication regimes, and also in a full driveline rig. It will be shown that fuel economy improvements of up to 4% can be achieved in a fired engine by moving from a 15W40 crankcase lubricant to an optimised 0W20. In addition, it will be shown that within a single viscosity grade, significant improvements in fuel economy can be achieved by tailoring the viscometrics and also by optimising the additive package. Finally, it will be shown that by using laboratory bench test data to model quantitatively the response in the fired engine, the relative importance of different lubrication regimes in determining overall HDD vehicle fuel economy can be estimated.

1. INTRODUCTION

Proposed legislation controlling exhaust emissions from heavy duty diesel vehicles will require significant reductions in the levels of harmful gases and particulates escaping to the environment. The allowed limits will decrease in some cases by an order of magnitude over the next five to eight years. A number of different approaches will be needed to ensure that these targets can be met, including the use of after treatment devices, design changes to improve engine efficiency and maximising the contribution of the crankcase lubricant to fuel economy. In addition to the legislative requirements for reduced emissions, there is considerable pressure from haulage operators to improve the fuel efficiency of heavy duty diesel vehicles. Fuel costs represent a significant part of the total running costs of any transport fleet. Estimates vary[2,3] between 30% and 50% and will depend upon taxation schedules in different countries, type of vehicle and service requirements. The ability of the engine lubricant to reduce fuel consumption is therefore an important route not only to the reduction of vehicular emissions, but also to enhanced operating profits for the consumer.

The extent to which crankcase lubricants can reduce fuel consumption in heavy duty diesel vehicles is dependent on many factors, but previous estimates[2-4] suggest that 3.5% might be achievable relative to the consumption recorded when a conventionally formulated 15W40 HDD oil is used.

In order to maximise lubricant related fuel efficiency it is necessary to minimise engine friction over a wide range of conditions of temperature, pressure, shear rate and speed. The combinations of these conditions will result in different lubrication regimes operating in the engine, including full film

and mixed hydrodynamic lubrication, full film and mixed elastohydrodynamic lubrication and some boundary lubrication.

In an HDD engine, all of these lubrication regimes will be in operation, although to differing extents. Previous estimates[5] have suggested that in an HDD engine, frictional losses in the bearings and piston assembly (mixed and full film HD lubrication) account for the up to 90% of engine friction, the rest being generated in the valvetrain (mixed/full film EHD and boundary lubrication). Thus the maximisation of fuel economy from the lubricant will require primarily minimisation of friction under mixed and full film HD lubrication conditions, and additionally (although the effects are expected to be less significant) minimisation of friction under mixed/full film EHD and boundary lubrication conditions.

Under full film lubrication conditions, be it HD or EHD, frictional losses are dependent on the dynamic viscosity of the lubricant and upon the effective pressure viscosity coefficient, α. Engine friction can therefore be reduced by minimising the dynamic viscosity of the lubricant (and, under EHD conditions, using a basestock of lower α value).

However, the benefits obtained from reductions in viscosity are not unlimited, since under HD and EHD conditions, both film thickness and friction depend equally on viscosity. Therefore reductions in viscosity lead to a decrease in film thickness until eventually the ratio of film thickness to composite surface roughness, λ, falls below a critical level (~3) and asperity contact occurs. This causes an increase in friction as the lubrication moves from full-film to the mixed regime.

It has been demonstrated[6] previously that surface active additives can delay the transition from full film to mixed lubrication. This is achieved by the formation of discrete chemisorbed layers on the metal surfaces which effectively reduce the composite surface roughness and thereby delay the transition to mixed lubrication.

Under boundary conditions, as might be experienced in the valvetrain (particularly with oils of lower viscosity grades), frictional losses can be reduced by appropriate additive selection, such that metal to metal contact is minimised, and boundary friction controlled, through the action of the lubricant.

In this study, the impact of three variables on lubricant related fuel economy has been studied. The effect of changing viscosity grade, varying VM type and treat and optimising the additive package on friction and fuel economy has been evaluated in laboratory bench tests and a driveline rig respectively. Finally, the complete data set has been analysed statistically in an attempt to understand the role of different lubrication regimes in the generation of total engine friction.

2. LUBRICANT SELECTION

A matrix of oils with differing viscometric properties and formulations has been blended to evaluate the impact of viscosity grade, ratio of newtonian to non-newtonian viscosity and additive package on friction under different lubrication regimes and engine fuel efficiency. A description of the test oils is given in Table 1.

Table 1 Test oil matrix

Visc Grade	Newt. Visc	BS	VM	Addpak
15W40		Group II	Block copolymer	A
0W20	Max	PAO	Block copolymer	A
5W30	Max	PAO	Block copolymer	A
5W30	Min	PAO	Block copolymer	A
5W30	Min	PAO	Star A	A
5W30	Min	PAO	Star B	A
5W30	Min	PAO	Star C	A
5W30	Min	PAO	OCP A	A
5W30	Min	PAO	OCP B	A
5W30	Max	PAO	Block copolymer	B
5W30	Max	PAO	Block copolymer	C

Table 1 shows how the viscometric and formulation parameters were varied in each of the 11 oils. Basestocks (BS) used were either a Group II or

PAO and three different additive packages were studied. Six viscosity modifiers (VM) and three different viscosity grades have been evaluated. Within each viscosity grade it is possible to vary the Newtonian viscosity of the oil (that is the part which is independent of shear rate), and still achieve the same finished oil viscometrics (HTHS and kV100). This is achieved by using different amounts of viscosity modifier to boost the viscosity at low and medium shear rates. This is illustrated schematically in Figure 1.

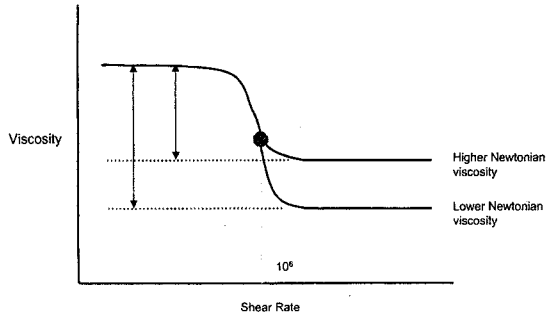


Figure 1 Schematic showing how the same low shear rate viscometrics can be achieved using different ratios of Newtonian (or base viscosity) to non-Newtonian viscosity.

The Newtonian component of the viscosity is provided by the basestock and the additive package, whereas the non-newtonian contribution is provided by the viscosity modifier. At high shear rates, the viscosity modifiers undergo shear thinning and their contribution to viscosity is lost. At these shear rates, the viscosity of the oil will be determined by the Newtonian component of the viscosity which is independent of shear rate.

Each of the 11 oils in Table 1 have been formulated either to maximise the Newtonian component of the viscosity, or to minimise it. These details are given in the second column of Table 1. Two oils formulated to the same low shear rate viscometrics, where the Newtonian viscosity has been maximised in once case and minimised in the other, might well be expected to behave differently in an engine or rig test where shear rates are often in excess of 10^6 s^{-1}.

3. LABORATORY BENCH TESTS

The matrix of oils described in Table 1 has been evaluated using a number of laboratory bench tests, and a full driveline rig at Shell Global Solutions PAE laboratory, Hamburg. The bench tests aim to evaluate either friction or contact film thickness under a number of different lubrication regimes and the driveline rig measures the impact of the crankcase lubricant on vehicle fuel economy under controlled running conditions. The equipment used is described below.

3.1. Journal Bearing Friction Rig

The journal bearing friction rig, as shown in Figure 2, has been developed by Infineum and PCS Instruments. This rig uses automotive production crankshaft bearings and a modified production connecting rod. Load can either be held constant, or varied so as to replicate the load/crank angle profile of a real engine.

Bearing friction is measured as a function of speed at a range of temperatures (up to 150 °C). Stribeck curves generated using this rig show the variation of friction with speed as the lubrication regime changes from full film hydrodynamic to mixed hydrodynamic.

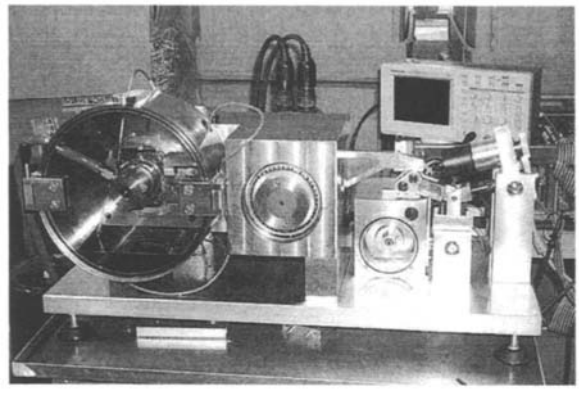


Figure 2 Bearing Friction Rig.

3.2. EHD Optical Film Thickness Rig

EHD film thickness between a steel ball and glass disk is measured as a function of entrainment speed using an optical EHD test rig[6], as shown

schematically in Figure 3. In these tests, entrainment speed was varied between 10^{-3} and 10^{2} ms^{-1}, while test temperatures ranged between 40 °C and 100 °C. Tests were carried out in pure rolling.

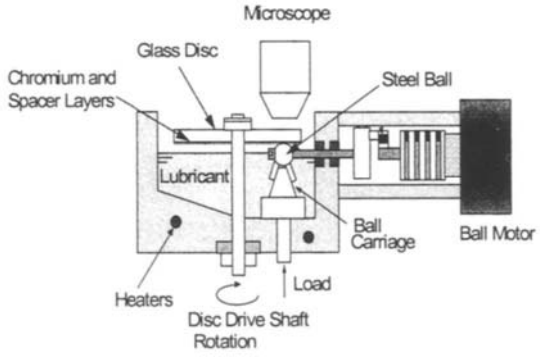


Figure 3 EHD Optical Film Thickness Rig.

3.3. Mini Traction Rig (MTM)

Mixed and full film EHD friction (or traction) is measured using a ball on disc traction rig, as shown in Figure 4. Traction measurements are made between a steel ball and a steel disc at a range of different slide/roll ratios using a fixed disc speed. Measurements were made at a series of temperatures ranging from 40 °C to 135° C.

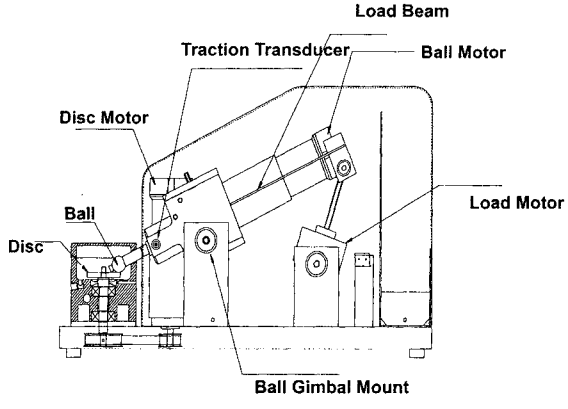


Figure 4: Mini Traction Rig.

3.4. HFRR (Boundary friction rig)

Boundary friction coefficients were measured at temperatures between 40 °C and 140 °C using a High Frequency Reciprocating Rig (HFRR), as shown schematically in Figure 5. The test comprises a steel ball reciprocating against a steel disc at 20 Hz with a stroke length of 1 mm. A 4 N load was used in all tests. Friction coefficient and temperature are measured as a function of time.

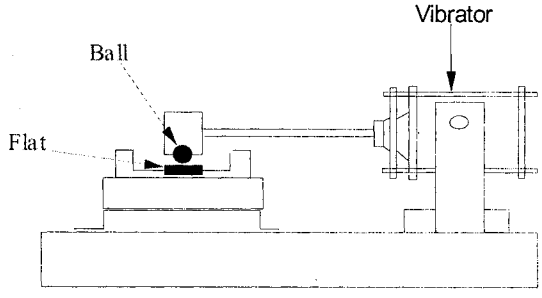


Figure 5 High Frequency Reciprocating Rig.

3.5. Driveline Rig

A Volvo FH12 direct injected 6 cylinder engine, in line with EURO 2 emissions, was installed with its associated gearbox and axle system in the Shell Global Solutions PAE Hamburg test facility. A schematic of this rig is shown in Figure 6 below. Fuel consumption is monitored continuously and differences in performance are calculated relative to a 15W40 reference oil.

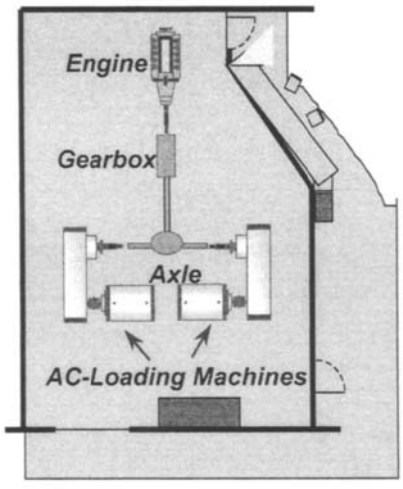


Figure 6 Schematic of driveline rig used to measure fuel economy contribution of test lubricants.

The testing programme was conducted under both steady state and transient operations so as to reflect realistic driving conditions. The gearbox and axle were lubricated with a standard gear and axle oil. Table 2 shows the general loading conditions used for this study.

Table 2 General loading conditions.

Engine Condition	RPM	Torque (Nm)	Power (kW)
Transient City Running	600 (min)	250 (min)	15.7 (min)
	1400 (max)	1100 (max)	160 (max)
Steady State Freeway	1200	750	94

4. RESULTS & DISCUSSION

4.1. Effect of Viscosity Grade and Newtonian Viscosity

The effect of viscosity grade and Newtonian viscosity within a given viscosity grade on friction and fuel economy was evaluated using the rigs described in section 3. Figure 7 shows the variation of bearing friction with journal speed for four oils measured at 100 °C in the journal bearing rig.

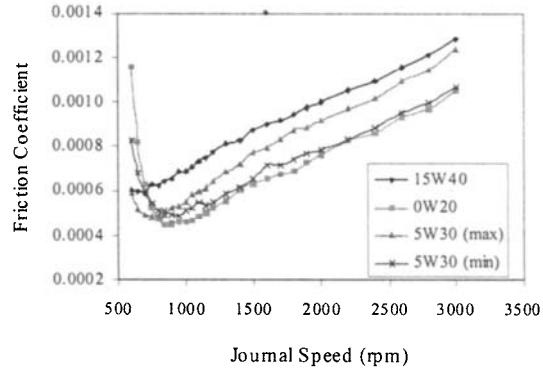


Figure 7 Variation of bearing friction with journal speed at 100 °C for oils with different viscometric properties.

The graph in Figure 7 shows the impact of changing viscometrics on bearing friction. All oils were blended using the same additive package and viscosity modifier. In the 15W40, 0W20 and 5W30(max) oils, the Newtonian viscosity has been maximised. A comparison of these three oils shows clearly that the bearing friction is reduced over the majority of the speed range by moving to a lower viscosity grade. However, it should be noted that at low speeds, the lubrication transitions from full film to mixed hydrodynamic, and that this transition occurs at progressively higher speeds as the viscosity grade is reduced. The result is that at low speeds the 0W20 oil exhibits considerably higher friction than the 15W40 oil. In the 5W30(min) oil, the Newtonian component of the viscosity has been minimised so that a much greater contribution to viscosity is derived from the non-Newtonain viscosity modifier. A comparison of the two 5W30 oils demonstrates that within a given viscosity grade, significant reductions in friction can be achieved through tailoring of the viscometric properties. It should be noted that both these oils have an HTHS of 3.5 cP, and a kV100 of 12 cSt.

Figure 8 shows measured values of EHD friction as a function of speed at 100 °C for the same four oils.

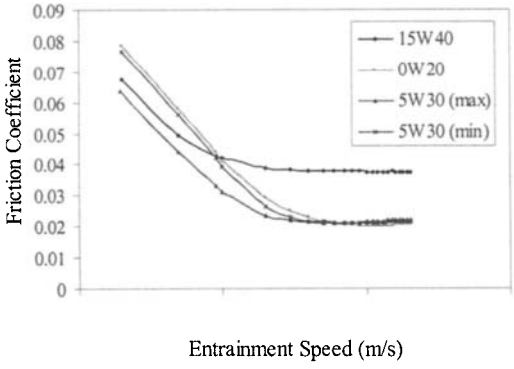


Figure 8 Variation of EHD friction with entrainment speed at 100 °C for oils with different viscometric properties.

The greatest differences in EHD friction are observed at high speed where the lubrication is in the full film regime. Under these conditions, there is a significant increase in measured friction for the 15W40 relative to the other oils. This is a basestock effect, and reflects the difference in pressure viscosity coefficient between the Group II basestock in the 15W40 and the PAO basestock in the other oils. Under the contact pressures experienced in this test, the viscosity of the Group II basestock increases to a much greater extent than does that of the PAO, hence the substantial difference in measured friction. Under low speed conditions, the less viscous oils, 5W30 (min) and 0W20, show an earlier transition to the mixed regime due to an earlier onset of asperity contact. The 5W30(max) oil shows the lowest friction across the speed range.

Figure 9 shows the variation of measured film thickness as a function of entrainment speed for the same oils.

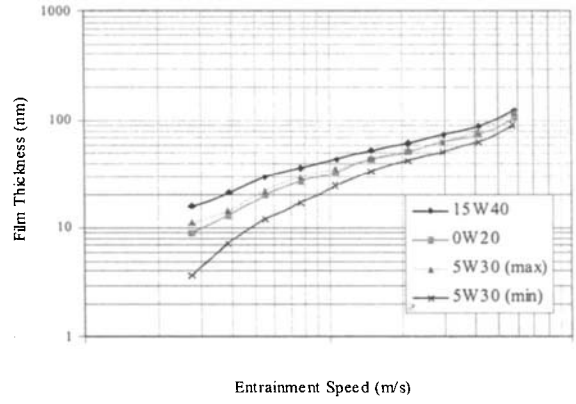


Figure 9 Variation of film thickness with entrainment speed at 100 °C for oils with different viscometric properties.

As expected, the 15W40 oil gives the greatest film thickness at any given speed. The film thickness measured for the 5W30(min) oil is significantly lower at low speeds than that measured for the other oils. This poorer performance might be expected to lead to reduced wear protection in service.

Figure 10 shows the relative performance of these oils in the driveline rig. The improvement in fuel economy relative to the 15W40 reference is plotted for each oil and for each set of running conditions.

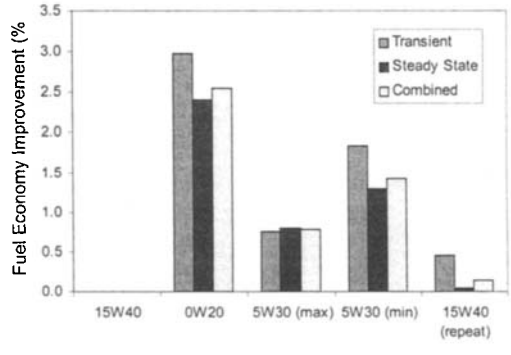


Figure 10 Measured improvements in fuel economy at 90 °C relative to 15W40 under various running conditions in the driveline rig.

The driveline rig data confirms the trends observed in the journal bearing rig, that is, the greatest fuel economy improvement is observed for the 0W20 oil, and the 5W30(min) oil shows a FE credit compared with the 5W30(max) oil. This validates the hypothesis that two oils formulated within the same viscosity grade, and with the same HTHS and kV100 viscosities can have markedly different performance in fuel economy tests. The final result on the graph shows the repeat measurement on the 15W40 oil. This repeat suggests that the error in the measurements is of the order of 0.3%.

4.2. Effect of Additive Package

Three oils containing different additive systems were used to evaluate the impact of additive chemistry on HDD fuel economy. Figure 11 shows the measured bearing friction as a function of journal speed for these oils.

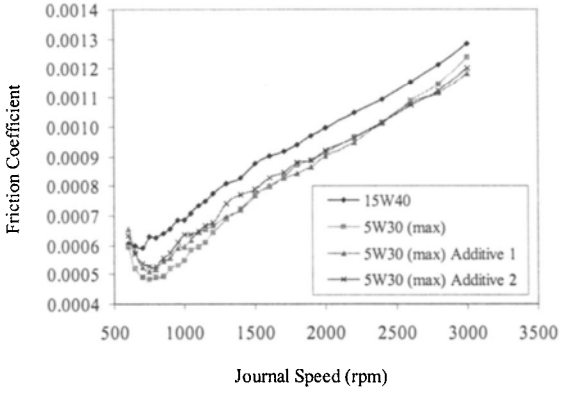


Figure 11 Variation of bearing friction with journal speed at 100 °C for oils containing different additive systems.

The graph in Figure 11 demonstrates that no reduction in friction is achieved under these conditions by optimising the additive package. However, the relative improvement in friction of the 5W30 oils compared with the 15W40 is clearly seen.

Figure 12 shows the effect of optimising the additive package on boundary friction. In this case, the 0W20 oil is added for comparison.

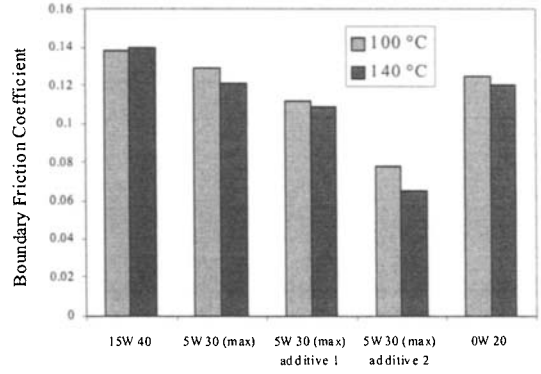


Figure 12 Measured boundary friction coefficients from the HFRR for oils containing different additive systems.

The results in Figure 12 show that under boundary conditions, the addition of additives 1 and 2 yields significant reductions in friction. In this case, changing the viscosity grade from 15W40 down to 0W20 has little impact on friction. This is expected under boundary conditions, where there the oil film in the contact is extremely thin, and surface protection is afforded largely by chemisorbed additive layers alone.

Figure 13 shows the effect of optimising the additive package on fuel economy measured in the driveline rig.

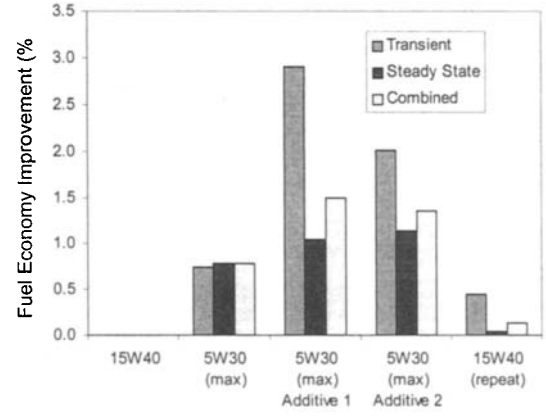


Figure 13 Measured improvement in fuel economy at 90 °C relative to 15W40 for oils containing different additive systems.

The plot in Figure 13 shows that under the conditions in the driveline test, the addition of both additives 1 and 2 gives improvements in fuel economy, and that the majority of these improvements are afforded under transient running conditions. One explanation for this is that more boundary lubrication is found under the transient running conditions than under the steady state conditions. This data suggests that in the combined cycle, there is indeed a significant proportion of boundary lubrication in the engine.

4.3. Effect of Viscosity Modifier Type

Viscosity modifiers with different chemical structures have different shear characteristics, and the effect of these differences on HDD fuel economy and friction has been evaluated. Figure 14 shows the effect of changing VM structure on bearing friction. In each case, the treat of the VM is different, but the oils have all been formulated to meet the same high temperature viscometrics.

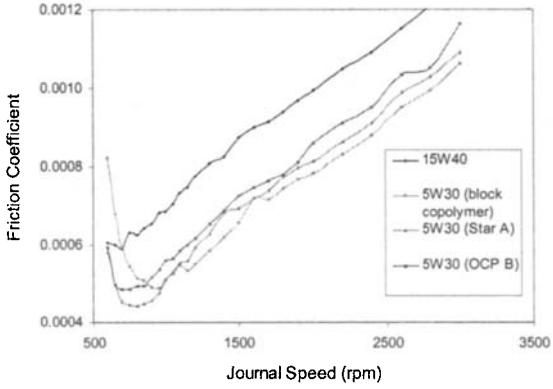


Figure 14 Variation of bearing friction with journal speed at 100 °C for oils formulated with different viscosity modifiers.

The data from the journal bearing rig in Figure 14 shows a small advantage in friction reduction for the Star and block copolymer over the OCP. At very low speeds, the star polymer gives the lowest friction. All 5W30 oils show an advantage over the 15W40. Figure 15 shows the comparative data from the driveline fuel economy rig.

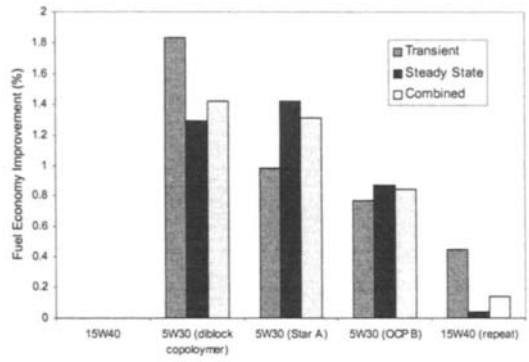


Figure 15 Measured improvement in fuel economy at 90 °C relative to 15W40 for oils formulated with different viscosity modifiers.

The data from the driveline rig shown in Figure 15 suggests the same trend as that observed in the journal bearing rig at higher speeds, that is a fuel economy advantage for the Star polymer and block copolymer over the OCP. The behaviour of these molecules in formulated oils under high shear conditions is not well understood, and an explanation for this behaviour will require further investigation. However, the data from this study suggests strongly that the chemical structure of the viscosity modifier and the treat rate are important factors in controlling the fuel economy in HDD applications.

5. STATISTICAL ANALYSIS

Improving fuel efficiency in any vehicle requires the minimisation of friction in all parts of the engine, and it is well known that different areas of the engine will operate under different lubrication regimes. Since friction reduction under each of these regimes requires a different formulation approach, the minimisation of total engine friction is facilitated by an understanding of the relative importance of each of the lubrication regimes in the engine.

The bench rigs used in this study are designed to measure friction or film thickness under different lubrication regimes. It was anticipated that by

analysing statistically the results from the bench test data and the driveline rig, it might be possible to estimate the relative importance in the engine test of friction generated under the different regimes.

The data from the driveline rig (the "response") has been modelled with a series of predictors using a linear regression technique. The predictors included data from all the bench tests, chosen to represent the behaviour of the lubricants under specific lubrication regimes.

The analysis looked firstly at main effects and secondly at the interactions between those effects. The objective was to determine which of the predictors, if any, were most highly correlated with the response data, and thereby estimate the importance of various lubrication regimes in determining friction in the engine.

Table 3 Statistical models from linear regression analysis using main effects only; JBM = journal bearing rig, FT = film thickness rig; MTM = traction rig

Response	Model	R^2	Associated Lubrication Regimes
Transient 90	$k_t + a_t$[JBM1] + b_t[FT1]	0.97	Mixed hydrodynamic
Steady State 90	$k_s + a_s$[MTM135]	0.35	Mixed EHD
Combined 90	$k_c + a_c$[FT1]	0.41	Mixed hydrodynamic

Table 4 Statistical models from linear regression analysis including interaction terms

Response	Model	R^2	Associated Lubrication Regimes
Combined 90	$K_c + A_c$[FT2*MTM80] + B_c[FT1*FT3]	0.91	Mixed EHD/HD?

Table 3 shows some of the resulting models from this analysis, using main effects only. In each case, the R^2 value is also given to show the extent of the correlation. A strong correlation is observed between the transient data from the driveline rig at 90 °C and a combination of two terms from the journal bearing rig and the film thickness rig respectively. This suggests that mixed hydrodynamic lubrication may be the dominant regime in the engine under these conditions. It was not possible to correlate the other responses at 90 °C to the bench data, as shown by the low R^2 values for "Steady state 90" and "Combined 90".

Table 4 shows that on adding interaction terms to the analysis, it was possible to improve the correlation between "Combined 90" and the bench data significantly. While this process allowed a large improvement in the R^2 value, the addition of interaction terms into the model makes any physical interpretation of the data almost impossible

The validity of this exercise is limited by the small data set (only 11 data points), and the results from this analysis must be understood to be indications at this stage as opposed to well-established predictions.

6. CONCLUSIONS

Fuel economy benefits in heavy duty diesel engine testing have been observed for oils of lower viscosity grade.

Optimising the balance of Newtonian to non-Newtonian viscosity within a grade has been shown to give measurable benefits in fuel economy.

The addition of lubricant additives can provide improvements in fuel economy, particularly under transient running conditions.

The chemical structure of viscosity modifiers has been shown to have a significant effect on both bearing friction and engine fuel economy, with an

advantage observed for star polymers and block copolymers over OCPs.

Statistical analysis of the data set suggests EHD/HD mixed lubrication are the dominant regimes, but engine test data also shows importance of boundary friction control in heavy duty diesel fuel economy.

7. REFERENCES

1. "The Role of Lubricant Additives in Improving Fuel Economy for Heavy Duty Diesel Applications", R. C. Castle, S. Arrowsmith, C. J. Locke, C. H. Bovington, Technische Arbeitstagung Hohenheim, Stuttgart, 2002
2. "European Requirements for a Super High-Performance Diesel Oil", D.C. Colbourne, Truck Technology International, 1988, pp 108-111
3. "Engine Friction Lubricant Sensitivities : A Comparison of Modern Diesel and Gasoline Engines", R. I. Taylor, Proceedings of the 11th International Colloquium, Esslingen, "Industrial and Automotive Lubrication", 1998 (published by the Technische Akademie Esslingen, 1998)
4. "Fuel Economy Improvement by Engine and Gear Oils", W.J. Bartz, Proceedings of the 24th Leeds-Lyon Symposium on Tribology (published in Tribology for Energy Conservation, Tribology Series, pp 13-24, 34, 1998, Elsevier, Editor: D.Dowson)
5. "Heavy Duty Diesel Engine Fuel Economy: Lubricant Sensitivities", R. I. Taylor, SAE Spring Fuels and Lubricants Conference, Paris, 2000
6. "The Importance of the Stribeck curve in the Minimisation of Engine Friction", C. Bovington, S. Korcek and J. Sorab; Leeds-Lyon Tribology Symposium, Lyon, Sept 1998

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

A new tribometer for piston ring friction measurements

Peter Andersson

VTT Technical Research Centre of Finland, VTT Industrial Systems, Machine Operability, P.O.Box 1702, FIN-02044 VTT (Espoo), Finland

The present paper contains a description of a new piston ring tribometer. The new piece of equipment allows instrumented tribological testing of piston ring and cylinder liner samples at temperatures, contact pressures, oscillating frequencies and lubrication conditions typical of diesel engines. The friction force is measured with a transducer that moves with the piston ring sample.

The present paper includes as examples the results of initial tests with a compression ring segment against a cylinder liner specimen at a stroke amplitude of 40 mm, a temperature of 110°C and frequencies up to 20 Hz. The results show that the friction measurement principle with a moving friction force transducer is useful for friction measurements under reciprocating motion at high frequency. The results obtained were in good correlation with similar results found in literature, which encourages to further investigations with the piston ring tribometer.

1. INTRODUCTION

The operational conditions of piston rings in internal combustion engines are highly demanding. The combustion gas temperature can reach temperatures of 1800 to 2600°C, and the temperature at the groove for the top piston ring is in the order of 200°C under normal operation. The maximum gas pressure to be sealed by the piston ring pack is in the excess of 180 bar in turbo diesel engines. The piston reciprocates with an near-sinusoidal velocity distribution, with velocity peaks that may well exceed 10 m/s and accelerations of several thousand m/s^2 arising thereof. [1]

The rings must possess an excellent sealing capability, for suppressing the exhaust emissions by limiting the escape of engine oil to the combustion chamber, and for minimising the blow-by, with the aim of both conserving the power efficiency rate and minimising the risk of cylinder scuffing. Due to emission suppression demands the amount of oil available for the lubrication of the top ring groove is presently rather limited.

The sealing ability of the piston rings relies on a spring force, obtained by ring compression and, particularly regarding the top ring, by a gas pressure on the back face of the ring. Various ring profiles are currently utilised for optimising the ring performance. Mass and friction forces cause deformation of the ring and complicate the sealing performance of the ring. Wear of the ring and the ring groove furthermore causes interference in the sealing ability.

The power losses due to the frictional work of the piston rings are responsible for a significant proportion of the entire frictional power losses of present internal combustion engines [2]. In order to support a high power efficiency rate continuous efforts are made for achieving piston ring solutions that operate under as low coefficient of friction as possible.

In addition to the sealing function, the piston rings are responsible for part of the

transfer of heat from the piston to the cylinder.

In order to ensure a long operational lifetime, the wear of the ring must be low in spite of severe sliding conditions against the cylinder liner. The cylinder liner wear that is caused by the piston rings must remain low, for maintaining the oil retention capacity of the honed texture of the liner and for eliminating the formation of scratches. A reliable operation of the ring pack requires material and coatings with good resistance against mechanical-thermal fatigue and chemical attacks. Ring wear is reduced by the use of various coatings on the ring face. Soft top coatings can be applied on piston ring faces for reducing friction during running-in. Finally, the piston rings must be reliable and cost effective for a significantly long time.

The tribological considerations in the contacts formed by the piston skirt, piston rings and cylinder liner, or bore, have gained a significant amount of attraction over several decades, not least indicated by a large number of articles published on this topic in recent years. Recent studies include modelling, miniaturised experimental work and full-scale engine testing. Features of particular interest are the coefficient of friction, the oil film thickness, the scuffing resistance of the ring/cylinder system, and aspects related to coating and lubricant development. Miniaturised or downscaled tests for the evaluation of the tribological properties of the ring/liner system are less expensive than engine tests, and different phenomena can be investigated more easily without interactions from other simultaneous phenomena.

Previous test rigs used in recent years by other authors for the evaluation of the tribological properties of the piston ring / cylinder liner contact have been built for a piston ring or a ring segment sliding under reciprocation against a part of a cylinder liner or a complete floating cylinder liner [3-12]. A crank or eccentric mechanism or an electromagnetic actuator has been employed for achieving a horizontal or vertical reciprocating motion of a piston ring or cylinder liner specimen. The friction force has been measured either from the ring or from the liner; however in most instances from the stationary specimen.

2. EQUIPMENT DESCRIPTION

2.1. General lay-out of the piston ring tribometer

The VTT piston ring tribometer has been optimised for the experimental simulation of the tribological conditions of piston ring and cylinder liner samples at crankshaft speeds up to 3000 rev/min. The tribometer is designed for tests with samples related to cylinder diameters in the range 100-200 mm. For larger diameters the frequency of the oscillation has to be reduced. The friction force and other parameters of interest can be recorded at a sampling rate of 500 data points each per crankshaft revolution. The fast response and the high sampling rate have been achieved by minimising the mass of the friction force linkage and by employing advanced data acquisition equipment [13]. When fully equipped, the measurements will include oil film thickness, position of the moving assembly, crankshaft speed and temperature of cylinder liner sample. The electrical oil film thickness measurements will utilise the cylinder liner sample and the piston ring sample as the electrodes. In terms of mechanics, the tribometer was built on two cast iron plates mounted on rubber blocks on a welded frame that stands on rubber blocks. One of the cast iron plates supports an actuator unit. The motion is transferred from the actuator unit by a force transmission rod to a friction measurement cell, which is mounted on the other cast iron plate. The use of two separate cast iron plates minimises the level of undesired vibrations at the friction measurement cell. An overview of the tribometer is presented in Fig. 1.

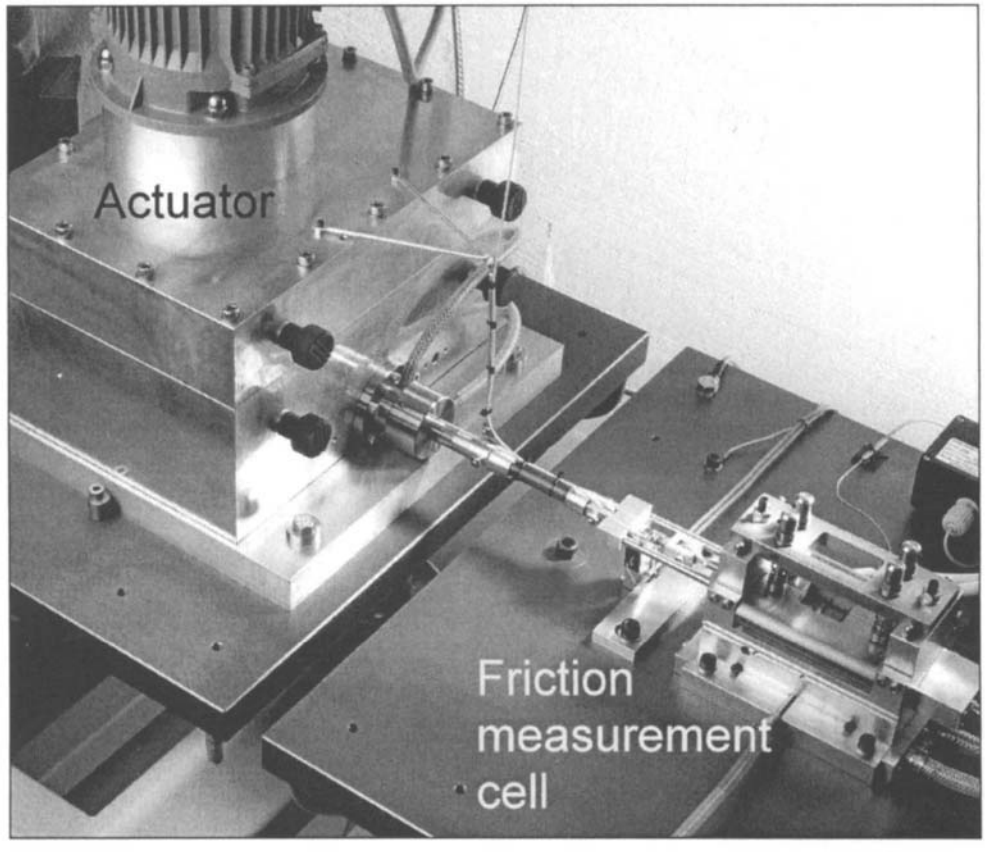


Figure 1. The VTT piston ring tribometer.

2.2. Actuator for near-sinusoidal motion

The actuator is based on a crankshaft, a connecting rod and a piston rod in a linear bearing. The unit is directly driven by an AC motor with a continuously variable speed control. The velocity curve of the piston rod is characteristic of a piston in an internal combustion engine, see Fig. 2. Due to a simple and robust design the function is reliable. Calculation formulae for ring position, velocity and acceleration in a crank mechanism are commonly available [2]. A major proportion of the mass forces are outbalanced by a counterweight at the crankshaft and by two counterbalancing shafts, one of which has a rotational direction opposite to that of the two other shafts.

The force transmission rod, which transfers the motion and force from the piston rod to the moving sample in the friction measurement cell, is a lightweight structure made from carbon fibre reinforced epoxy resin, aluminium and spherical joints.

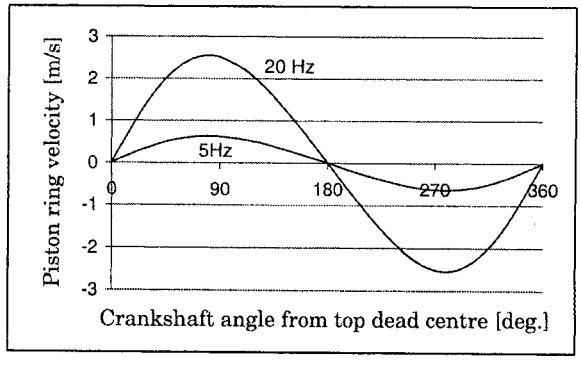


Figure 2. Sliding velocity distribution for the piston ring sample at 5 and 20 Hz frequency.

2.3. The friction measurement cell

The friction measurement cell comprises the novelty of the tribometer. It acts as an independent friction and measurement unit. The cell requires external motion that forces the piston ring sample to slide along the cylinder liner sample. The measurement signals are recorded by commercial data acquisition equipment and software developed in-house.

The friction measurement cell casing was machined from aluminium and steel. The casing consists of a water-cooled base plate, a guide box, two top covers with a hardened rail each, a loading bridge with four coil springs, and a lubricant supply installation, see Fig. 3.

The base plate supports the stationary specimen holder, which consists of a thermal insulation plate made of zirconia ceramic, an electrically heated block and an exchangeable sample holder with an electrically heated yoke for axial attachment of the cylinder liner sample. The heating capacity is 1 kW, which is sufficient for obtaining a cylinder liner temperature of 230°C.

Two parallel linear bearings in the guide box, two rails in the two top covers, and the shape of the cylinder liner govern the path of the reciprocating assembly in the friction measurement cell.

The two top covers are pivoted to the guide box and can be opened to expose the interior of the friction measurement cell, for inspection and test sample exchanges.

During testing, each of the two top covers are loaded by two adjustable coil springs in the loading bridge that give the normal force onto the test specimens.

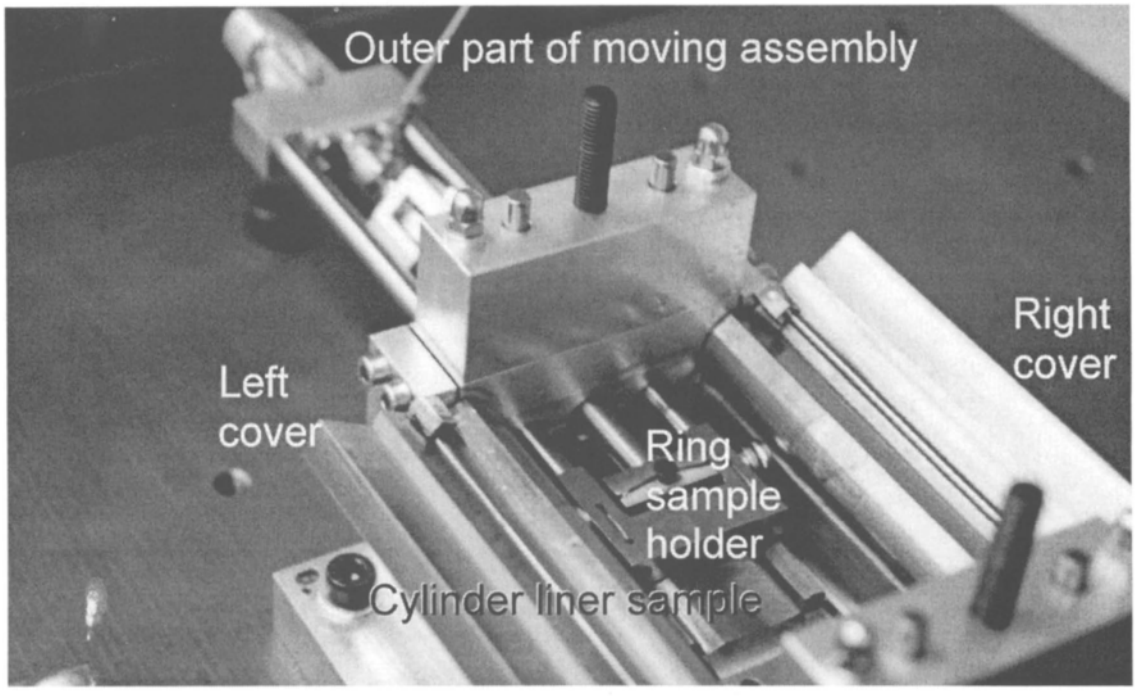


Figure 3. Detail of the friction measurement cell. The load bridge is removed and the two covers are open for exposing the cylinder liner sample and the inner part of the moving assembly including the holder for the piston ring sample. The outer part of the moving assembly including the force transducer is seen on the outer side of the friction measurement cell.

2.4. The moving assembly

The moving assembly (Figs. 4 and 5) of the friction measurement cell maintains a normal force (F_N) on the piston ring sample. The normal force is transferred from the coil springs through the top covers and two rails in the top covers to two grooved rollers at the end of the moving assembly, and further by a U-shaped suspension wire to the piston ring holder. The rollers allow the action point of the normal force to move synchronously with the piston ring sample.

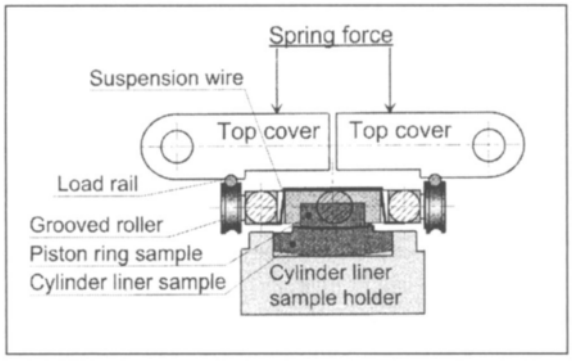


Figure 4. Schematic of the loading mechanism as seen from the actuator side.

The other function of the moving assembly is to allow measurements of the tangential force (F_μ) necessary for moving the piston ring sample under the normal force. The suspension wire allows minute mobility of the piston ring holder, which is necessary for the force measurements. The suspension wire additionally allows for minute lateral motion of the piston ring sample.

The principle of measuring the friction force from the moving piston ring sample was chosen with the aim of achieving fast responding friction recordings with reasonable resolution even at high frequencies. The arrangement requires that the force transducer moves with the moving assembly. Two interconnected guide rods, in linear bearings, are used for moving the assembly and a third, centrally located friction force rod is used for transferring the friction force, see Fig. 3. In the present installation, a piezoelectric force link, capable of measuring tensile and compressive forces up to 2.5 kN, and a charge amplifier are used. The high temperature inside the friction measurement cell casing made it necessary to install the force transducer outside the friction measurement cell.

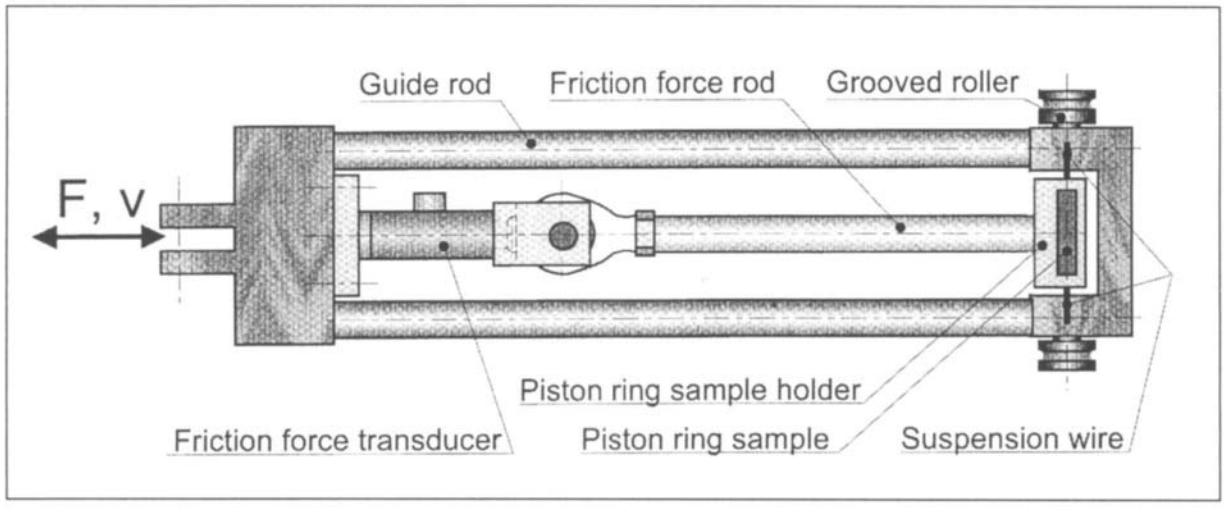


Figure 5. Schematic of the moving assembly of the VTT piston ring tribometer, as seen from the below.

2.5. Mass force compensation

The benefit of using a separate friction force linkage as part of the moving assembly is that the mass forces measured by the force transducer only contain a fraction of the mass forces of the entire moving assembly. The mass forces recorded originate from the accelerations and the mass of the ring sample holder and the friction force linkage. The correlation between the mass force recordings and the calculated values following the calculated acceleration curve of the mechanism and the mass (90 g) of the friction force linkage has been found to be good.

The true friction forces are obtained by mass force compensation in which the mass forces related to zero normal force, determined at a specific frequency, are subtracted from the total force readings. An example of the adopted mass force compensation is given in Fig. 6. The synchronising of the two force recordings is performed by using the simultaneously recorded readings of an optical position sensor, which is located close to the friction measurement cell. The sensor is based on an IR-LED and a position sensitive photodiode detector.

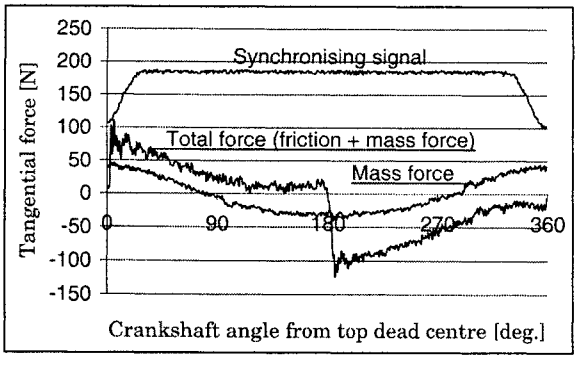


Figure 6. Unfiltered raw data signals for the total tangential (friction + mass) force and the mass force at 20 Hz oscillation, synchronised by the position sensor signals (only one shown), between two top dead centre positions of the crankshaft.

EXPERIMENTAL

3.1. Specimens

In the present study, the piston ring sample was made from an unused commercial ring intended for a cylinder of 108 mm diameter. The sample was a 20 mm long piece of a chrome plated compression ring, cut from the ring end, where the radius of curvature was 54.092 mm. The basic shape of the ring cross-section was rectangular, with a radial depth of 4.5 mm and an axial width of 2.5 mm, and the sliding surface of the ring had been honed to a convex "barrel-shape" profile with radiused edges. The apparent size of the sliding surface of the piston ring sample was thus 2.5×20 mm.

The cylinder liner sample was cut co-axially with the axis of the cylinder, from a commercial, plateau honed, unused cast iron cylinder liner of equally 108 mm diameter. The size of the cut sample was 30×60 mm, with the longer dimension in the direction of sliding.

The surface roughness of the ring and liner samples was measured before and after the test series by using a Taylor-Hobson Form Talysurf PC diamond stylus instrument, see selected parameters in Table 1.

The oil used for the lubrication of the piston ring and cylinder liner specimens was an unused commercial SAE 30 API CD diesel engine oil with a paraphenic base oil. According to the manufacturer of the oil, the viscosity was 102 cSt at 40°C and 12.2 cSt at 100°C temperature.

3.2. Experimental procedure

The friction measurements were carried out at 110°C cylinder liner surface temperature. The lubrication of the specimens was carried out by drip-feed lubrication at a rate of 8 µl/s onto the cylinder liner and the piston ring sample holder.

Table 1
Surface roughness parameters of the test samples before (after) the test series.

Surface roughness parameter	Piston ring sample		Cylinder liner sample	
Ra, arithmetic average roughness	0.17	(0.11 μm)	1.52	(1.09 μm)
RSK, skewness	-0.2	(-1.2)	-1.4	(-1.5)
Rpk, reduced peak height	0.26	(0.09 μm)	1.18	(0.85 μm)
Rvk, reduced trough depth	0.40	(0.35 μm)	7.81	(7.02 μm)

The friction measurements were carried out after a running-in procedure, during which the test was running at 5 Hz frequency and 500 N normal force for 30 minutes at a cylinder liner surface temperature of 110°C.

The friction measurements were carried out at 500 N normal force, at reciprocating frequencies of 5, 10 and 20 Hz. Mass force compensation required additional force measurements to be carried out under zero normal force at the respective frequencies.

4. TEST RESULTS

In the present work the test results are limited to three examples of friction measurement graphs (Fig. 7), piston ring and cylinder liner surface images (Fig. 8) and surface roughness measurement results (Table 1) representing the situation at the end of a test series of approximately one hour duration under 500 N normal force and with frequencies between 5 and 20 Hz.

5. DISCUSSION

The results of the present tests show that the measurement principle employing a moving force transducer is useful for friction measurements on piston ring and cylinder liner samples. The use of a reciprocating friction force transducer and linkage of low mass permits fast responding friction recordings and the calibration of the mass force component of the recorded force signal.

The friction graphs of Fig. 7 indicate kinetic coefficients of friction in the order of μ=0.08-0.12 in the range of reciprocating frequencies studied. The response time of the force measurement circuit is sufficiently short, as indicated by the instantaneous change in the value of the recorded force at the turning points of the motion. The influence of the reciprocating frequency is seen in the graphs, where a higher frequency is reflected by a lower coefficient of friction at mid-stroke of the motion. At mid-stroke the highest velocities and thus the most favourable conditions for hydrodynamic lubrication occur.

The level of the recorded coefficients of friction is in fair correlation with the results presented by Akalin and Newaz [3], Dearlove and Cheng [6], Glidewell and Korcek [7], Tung and Tseregounis [10], and Andersson and Sainio [14], who have presented peak values of μ=0.10-0.15, mid-stroke values in the range μ=0.02-0.10 and weighted average values of μ=0.04-0.11. In some of the works by other authors on piston ring friction, with a variable normal force, no coefficients of friction but instead friction force graphs are presented. The recorded levels of the coefficient of friction are among the larger ones reported in literature, which most likely reflects that a process of running in of the sliding surfaces still was active at the end of the test series.

As obvious from the image of the sliding surface of the piston ring sample (Fig. 8, upper), distributed shining bands had been formed on the piston ring surface during the running-in. Correspondingly, patches of a dark tribofilm had been formed on the cylinder liner surface (Fig. 8, lower) during the running-in process. During the test series the running-in wear had reduced the arithmetic average surface roughness (Ra) of the mating surfaces, mainly through abrasion or plastical deformation of the asperities of

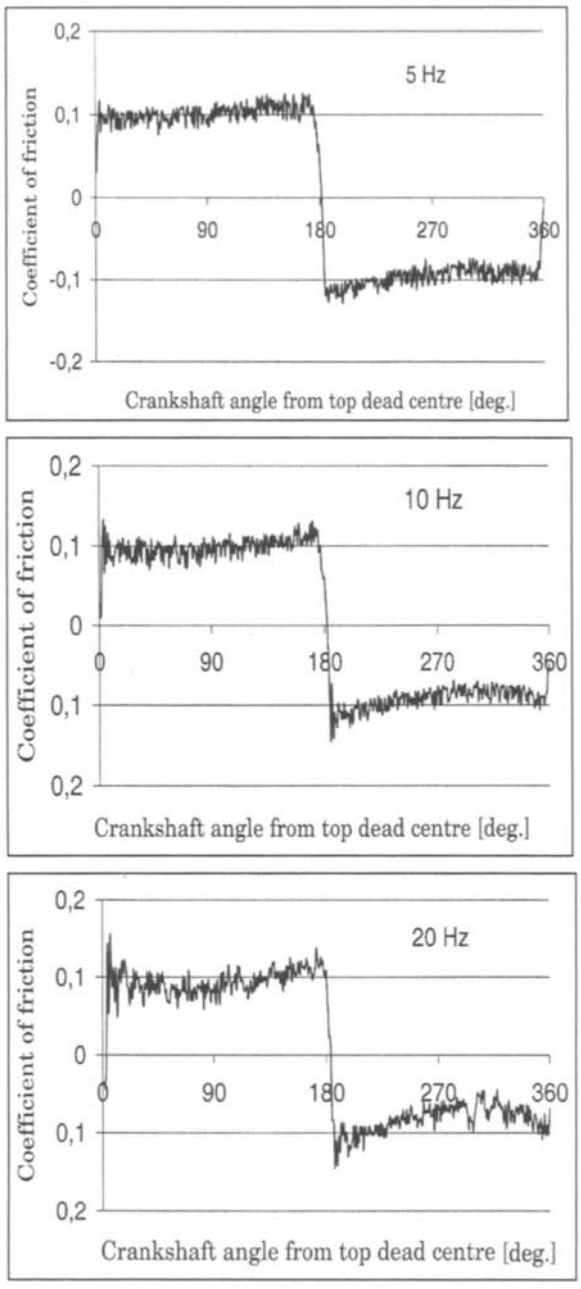


Figure 7. Friction graphs corresponding to 360 degrees of crankshaft angle, for 5, 10 and 20 Hz frequency of the piston ring tribometer. Positive values of the coefficient of friction refer to the downstroke movement of the piston ring and negative values to the upstroke movement of the ring.

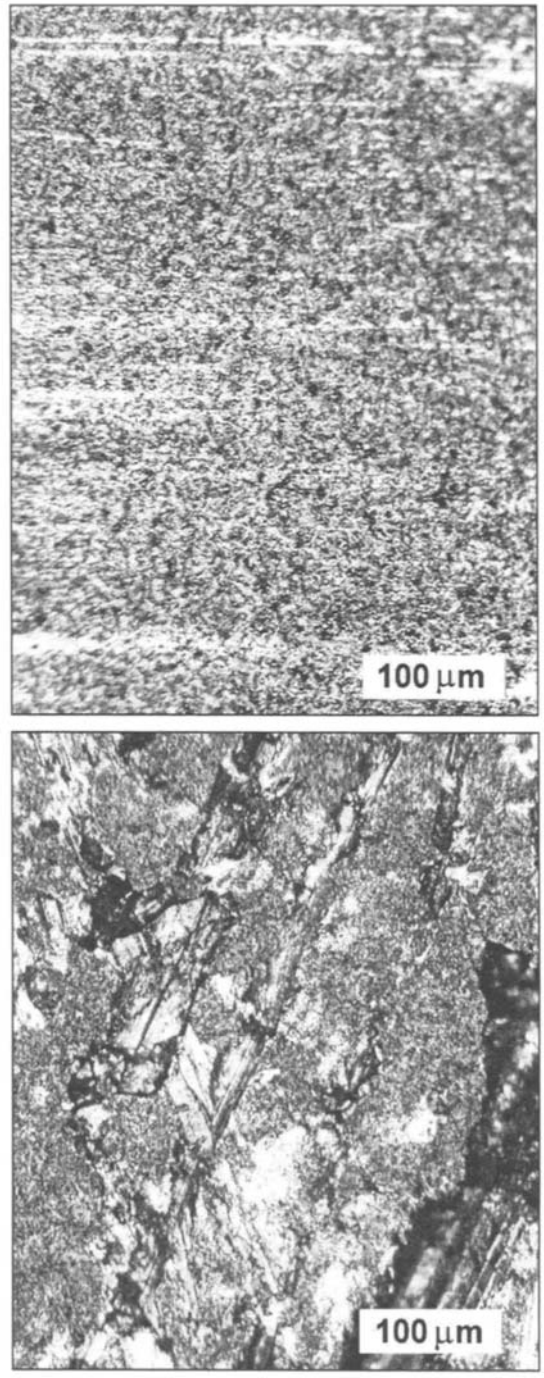


Figure 8. Photomicrographs showing the central location of (top) the piston ring and (bottom) the cylinder liner samples after the tests. Sliding direction is horizontal.

the surface roughness. At the present stage of the running-in process, most of the honing pattern still occurred on the cylinder liner surface (Fig. 8, lower). The wear or deformation at the surface asperities is reflected by larger reductions in the Rpk values than in the Rvk values. The RSK value, which is a dimensionless, relative value, has decreased due to the flattening of the asperities.

As shown in a recent work by Tung and Tseregounis [10], and Glidewell and Korcek [7], the additives in the lubricating oil strongly affect the level of the coefficient of friction between ring and liner. Furthermore it is well-established that the viscosity of the oil affects the coefficient of friction in a sliding contact. Other factors that govern the coefficient of friction between the piston ring and the liner are the surface quality, the piston ring profile, the load and the sliding velocity, hence it is not surprising that friction results presented in different works show variations.

The results of the present tests show an increase in the coefficient of friction at the start and stop locations of the friction graphs. At higher frequency the coefficient of friction at the position related to the start of the motion showed a slight increase. The frequency dependency of the start friction will require further studies.

Similarly to the present tests at, particularly, 10 and 20 Hz frequency, the friction graphs in some works by other authors comprise characteristic peaks at the start and stop locations [3, 5, 7], while other authors [6, 9] present friction graphs that lack any peaks of significance at the start and stop locations. It can be assumed that the difference in shape of the friction graphs from different experimental set-ups is related to the velocity distribution of the respective test rig, the friction force measurement arrangement and the damping characteristics of the entire test set-up, while a concise analysis of the differences would require further studies. The possibility to carry out detailed studies on the transients in the friction coefficient is a feature typical of the present equipment and its fast responding friction measurement arrangement.

6. CONCLUSIONS

The present work, which is a brief introduction to a newly developed piston ring tribometer, shows that the principle of a moving friction force transducer and a light friction force linkage is useful for friction measurements at mechanical oscillation of high frequency. The present results of initial tests with compression ring segments against cylinder liner specimens were in good correlation with similar results found in literature, which encourages to continuous work with more test parameter combinations.

ACKNOWLEDGEMENTS

The work, which was carried out as part of the Finnish national research programme *ProMotor*, was technically and financially supported by the National Technology Agency TEKES and the companies Fortum Oil & Gas Oy, Wärtsilä Corporation and SisuDiesel Inc. in Finland, and Volvo Technological Development Corporation in Sweden, for which the author expresses his gratefulness. Representatives of the above mentioned companies, and colleagues and project co-workers of VTT and Helsinki University of Technology, are acknowledged for fruitful discussions related to piston ring tribology and tribotesting. The author is furthermore grateful to Mr T. Nyström, Mr H. Sainio, Mr S. Varjus and Mr E. Vaajoensuu, for their contributions to the development of the piston ring tribometer.

REFERENCES

1. Röhrle, M.D. Pistons for internal combustion engines - Fundamentals for piston technology. Landsberg/Lech, Germany, 1995, Verlag moderne industrie, 70 p.
2. Dowson, D. Piston assemblies; Background and lubrication analysis. In: Taylor, C.M. (ed.). Engine Tribology.

Amsterdam, The Netherlands, 1993, Elsevier Science Publishers B.V., pp. 213-240. ISBN 0-444-89755-0.

3. Akalin, O. and Newaz, G. M. A new experimental technique for friction simulation in automotive piston ring and cylinder liners. Society of Automotive Engineers, Inc., 1998, SAE Paper 981407, 6 p.
4. Arcoumanis, C. , Duszynski, M., Flora, H. and Ostovar, P. Development of a piston-ring lubrication test-rig and investigation of boundary conditions for modelling lubricant film properties. Society of Automotive Engineers, Inc., 1995, SAE Paper 952468, 19 p.
5. Clarke, D. G., Sherrington, I. and Smith, E. H. The floating liner method applied to measure instantaneous piston assembly friction in a motored engine. In: J. Jakobsen, M. Klarskov, M. Eis (eds.). NORDTRIB '90 Proceedings of the 4th Nordic symposium on tribology, lubrication, friction and wear. 10th - 13th June 1990, Hirthals, Denmark. Copenhagen, Denmark, 1990, Technical University of Denmark. ISBN 87-983538-0-2. pp. 473-483.
6. Dearlove, J. and Cheng, W.K. Simultaneous piston ring friction and oil film thickness measurements in a reciprocating test rig. Society of Automotive Engineers, Inc., 1995, SAE Paper 952470, 11 p.
7. Glidewell, J. and Korcek, S. Piston ring / cylinder bore friction under flooded and starved lubrication using fresh and aged engine oils. Society of Automotive Engineers, Inc., 1998, SAE Paper 982659, 10 p.
8. Ma, Z., Henein, N.A., Bryzik, W. and Glidewell, J. Break-in liner wear and piston ring assembly friction in a spark ignited engine. STLE Tribology Transactions, 41 (1998) 4, pp. 497-504.
9. Noorman, M.T., Assanis, D.N., Patterson, D.J., Tung, S.C. and Tseregounis, S.I. Overview of techniques for measuring friction using bench tests and fired engines. In: Advances in powertrain tribology 2000, SAE/SP-00/1548, SAE Inc., Warrendale, PA, USA, 2000, ISBN 0-7680-0617-1, Paper 2000-01-1780, pp. 1-11.
10. Tung, S.C. and Tseregounis, S.I. An investigation of tribological characteristics of energy-conserving engine oils using a reciprocating bench test. In: Advances in powertrain tribology 2000, SAE/SP-00/1548, SAE Inc., Warrendale, PA, USA, 2000, ISBN 0-7680-0617-1, Paper 2000-01-1781, pp. 13-21.
11. Ye, Z.A. and Cheng, H.S.. Tribological behaviour of advanced material pairs of piston-ring/cylinder-liner. Society of Automotive Engineers, Inc., 1996, SAE Paper 961145, 9 p.
12. Andersson, P., Kytö, M., Mustonen, M., Tamminen, J. and Valkonen, A. Tribology of internal combustion engines - A literature survey. Espoo, 2000, Helsinki University of Technology, Department of Mechanical Engineering, Publications in Machine Design 1/2000, ISBN 951-22-4915-4, ISSN 1456-4955, 133 p.
13. Andersson, P. The VTT piston ring tribometer - A brief description. Espoo, Finland, 2001. VTT Manufacturing Technology, Report no. BVAL73-011148. 22 p.
14. Andersson, P. and Sainio, H. Oscillating piston ring and cylinder liner sliding tests using a pin-on-disc tribometer. Espoo, Finland, 2000. VTT Manufacturing Technology, Report no. BVAL73-001070. 25 p.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

The wear response of ceramic matrix nano-composite coatings in simulated piston-ring/cylinder wall test environments

K.L. Dahm, K. Panagopoulos, and P.A. Dearnley[a]

[a] School of Mechanical Engineering, University of Leeds, Leeds LS2 9JT, UK

A series of piston rings with surface coatings of alumina-matrix nano-composite materials have been evaluated in lubricated high-speed sliding contact with cast-iron cylinder liners at 200°C. Comparisons were made with standard CKS-36 (a commercial coating comprising chromium plating reinforced with alumina particles) coated cast-iron piston rings. Two types of wear mechanism were observed for the nano-composite coatings: (i) micro-scale cohesive failure, and (ii) smooth wear. The latter mechanism resulted in lower wear rates. The predominance of a given wear mechanism was found to depend on the coating composition and type of grinding operation used to finish the nano-composite coated piston rings.

1. BACKGROUND AND RATIONALE

Increasing performance demands placed on the piston ring-cylinder liner couple are increasing demand for highly wear-resistant piston ring coatings. There is additional environmental impetus to replace techniques such as electroplating (of chromium in particular) with environmentally-friendly plasma-based coating techniques.

1.1. Plasma-sprayed ceramic coatings

Plasma spraying represents a versatile technique for deposition of metallic, ceramic and cermet coatings onto metallic substrates. Molybdenum-base coatings are widely used for piston ring coatings due to their resistance to scuffing. These coatings are also nearing the limits of operation [1].

Several authors have investigated plasma sprayed alumina (Al_2O_3) coatings as an alternative to both electrodeposited chromium and thermally sprayed molybdenum coatings [1-3]. Tribological testing of such coatings has shown that while they are resistant to wear and scuffing, long running times leads to coating fracture (or "break-out") [1].

A ceramic coating that combines the inherent tribological advantages of Al_2O_3 but increases its resistance to fracture is therefore highly desirable.

1.2. Alumina-matrix nanocomposites

Ceramic nano-composites of Al_2O_3 and silicon carbide (SiC) have been reported to possess enhanced fracture toughness compared with pure Al_2O_3 [4]. These nano-composites comprise nano-sized (typically <200nm) SiC particles embedded within Al_2O_3 grains. The mechanism by which this enhanced fracture toughness is achieved and the magnitude of the increase are currently unclear [5,6]. Wet erosive wear testing showed that Al_2O_3-SiC nano-composites containing 5 vol% SiC exhibited wear rates at least two times lower than pure Al_2O_3 [7].

To date Al_2O_3-SiC nano-composites have been mainly investigated in bulk form. This paper reports an investigation into the production of wear resistant and fracture-resistant plasma sprayed Al_2O_3-SiC nano-composite coatings.

2. COATING OPTIMISATION

Nano-composite Al_2O_3-SiC powders containing 2.5, 5 and 10 % (by volume) SiC were obtained. Initial x-ray diffraction (XRD) and transmission electron microscopy (TEM) analyses of the powders confirmed the presence of nano-sized (10-20 nm) crystalline SiC particles within the larger α-Al_2O_3 grains.

2.1. Physical Powder Characteristics

Considerable study was devoted to the physical properties (the size distribution and flow characteristics) of the powders. Eventually a set of powders with adequate flowability for plasma spraying was obtained.

2.2. Spray Trials

2.2.1 Procedure

Spray trials were conducted using a commercial Sulzer-Metco spray gun. Spray parameters were based initially on recommended spray parameters for commercial fine alumina powders. For comparison, pure alumina coatings were also deposited. For the optimisation trials all of the coatings were deposited onto plain carbon steel substrates which had been grit-blasted and coated with a commercial Ni-Al bond layer.

Metallographic cross-sections of the coatings were prepared. These were examined using a Leitz Laborlux metallographic optical microscope in conjunction with a JVC camera. Using commercial image analysis software (Image Pro Plus) running on a PC the apparent porosity of the coatings was assessed from at least six micrographs taken from different areas of the coatings at a nominal magnification of 250×. The microhardness of the coatings was measured on the same cross-sections using a Shimadzu HMV2000 microhardness testing machine. At least five hardness indentations were made on each cross-section at a load of 0.98N (0.100 kgf) and a 15 second dwell time.

2.2.2 Results

The initial spray trials yielded highly porous coatings with a low indicated hardness compared to commercial alumina coatings. Subsequent coatings were deposited using a proprietary plasma gun. These coatings showed a lower porosity and higher indicated hardness than the previous coatings but still inferior (higher porosity and lower indicated hardness) than the commercial alumina coatings, Fig 1.

Further optimisation of the spray parameters led to improved nano-composite coatings. At a small working distance (50mm) both the 2.5% and 10% SiC nano-composite coatings showed higher hardnesses and lower porosity than the commercial alumina coating produced at the same distance, Fig 1. The 5% SiC nano-composite coating, although improved by the optimised spray parameters, showed higher porosity and lower hardness than the other nano-composite coatings. The reasons for this difference remain unclear.

XRD of the resultant coatings confirmed the retention of crystalline SiC in the coatings although the alumina matrix for the coatings was γ-Al_2O_3 rather than the α-Al_2O_3 phase present in the feed powders. TEM also showed the presence of 10-20nm SiC crystals within γ-Al_2O_3 coating grains.

3. TRIBOLOGICAL TESTING PROCEDURES

3.1. Sample preparation

Two types of piston ring sample were fabricated: (i) plasma spray coated spheroidal graphite cast iron rings and; (ii) plasma spray coated plain carbon steel rings. The cast iron rings (supplied by AE Goetze)

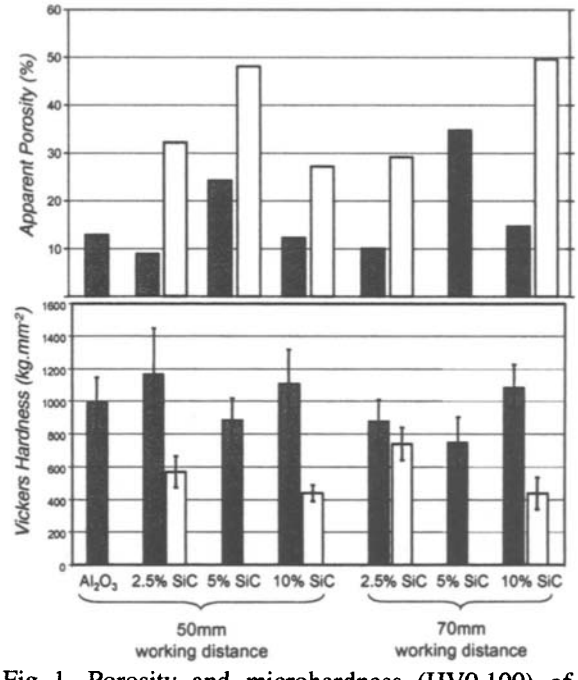


Fig 1. Porosity and microhardness (HV0,100) of plasma sprayed Al_2O_3-SiC nano-composite coatings deposited with trial parameters (open bars) and optimised parameters (shaded bars) using a proprietary plasma torch. For comparison, properties for a plasma sprayed pure alumina coating are included.

were coated whilst mounted on steel mandrels. The coated mandrels were then returned to AE Goetze for finishing using production-scale SiC grinding procedures. The cast iron piston rings were produced in all three nano-composite coating compositions. The plain carbon steel rings were produced by spray coating a solid plain carbon steel cylinder. The coated cylinder surface was then diamond ground and bored out. The ring samples were parted off from the cylinder and a bevel or "land" was applied to the coating edges to prevent premature coating fracture. Only the 2.5% SiC powder was used for these rings. The rings are listed in Table 1.

In all cases the substrates were grit blasted and then plasma sprayed with a commercial Ni-Al bond coat prior to plasma spraying with the nano-composite coating. The optimal spray parameters previously identified were used for all of the coatings. The total coating thickness was approximately 250μm and the outside diameter of all the rings was 90mm.

Table 1. Piston ring samples tested.

Ring	Coating	Subst.	Finish
D02	2.5% SiC	steel	diamond ground
S02	2.5% SiC	cast iron	commercial (SiC)
S05	5% SiC	cast iron	commercial (SiC)
S10	10% SiC	cast iron	commercial (SiC)
CKS	CKS-36	cast iron	commercial (SiC)

A commercial chromium coated ring, CKS-36 (supplied by AE Goetze) was also tested. The CKS-36 comprises a multilayer electroplated chromium coating in which alumina particles are embedded [8].

3.2. Test Parameters

High-speed reciprocating wear tests were performed using a Plint TE-77 test machine. Small sections (approximately 25-30 mm long) of the uncoated and coated piston rings were slid against sections from a 130mm i.d. grey cast iron liner. All tests were conducted using an additive free base oil (Rocol, UK). The wear test parameters are listed in Table 2. Tribological testing of stainless steel piston rings with and without sputter-deposited glassy alumina-based coatings has been reported previously [9].

The use of 90mm diameter rings sliding against a 130mm internal diameter liner produced a small elliptical wear scar. This simplified wear measurement and microscopy considerably. The 40N load gave average contact pressures at the end of the wear tests of between 4 and 8 MPa.

Table 2. Reciprocating wear test parameters.

Applied load (N)	40
Stroke (mm)	15
Frequency (Hz)	40
Average sliding speed (m/sec)	1.2
Temperature (°C)	200
Duration (hours)	6

3.3. Characterisation of worn surfaces

Contacting-2D profilometry was used to quantify the wear of both the rings and the liners. In the case of the liners the profiles were measured parallel to the reciprocation direction while for the rings the profiles were measured perpendicular to the

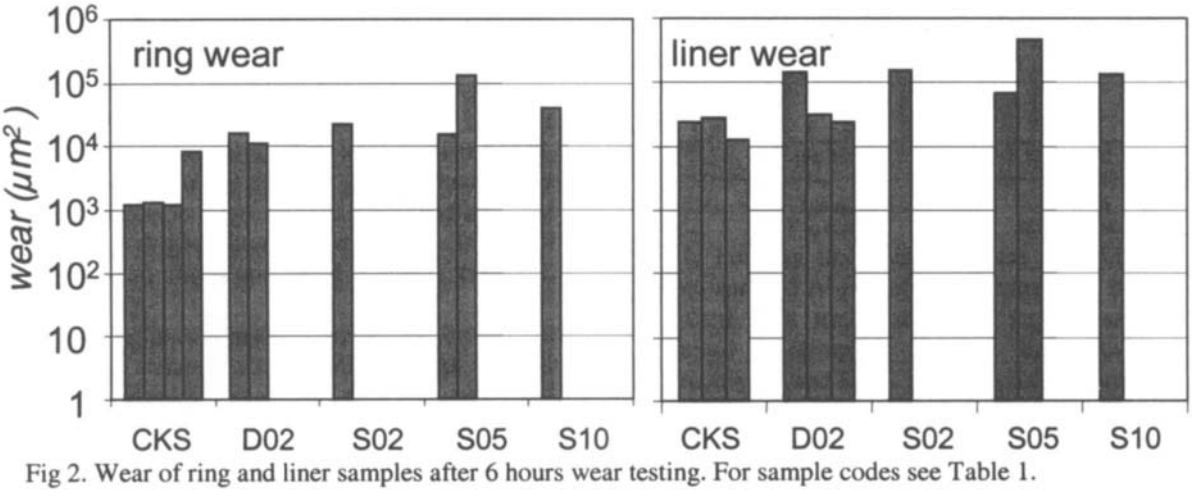


Fig 2. Wear of ring and liner samples after 6 hours wear testing. For sample codes see Table 1.

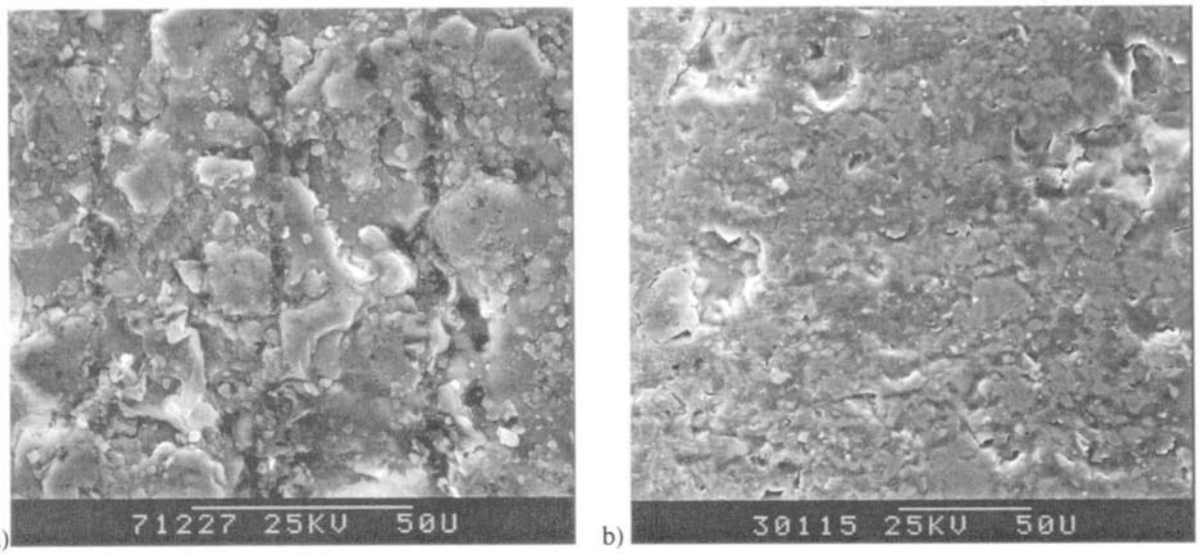


Fig 3. Worn surfaces of plasma sprayed nano-composite coatings after 6 hours testing: a) S05 ring exhibiting pull-out (fine particles) with only a small number of smoothly worn grains; b) D02 ring - smoothly worn. For sample codes see Table 1.

reciprocation direction. The wear (in μm^2) was calculated from the area of the trace after subtracting any background slope and curvature using unworn portions of the profile. The worn surfaces were examined using scanning electron microscopy (SEM).

4. TRIBOLOGICAL TESTING - RESULTS

The wear of both ring and liner samples for the plasma spray coatings was higher than for the CKS-36 coating, Fig 2. While the increase in liner wear due to the adoption of plasma sprayed ceramic coatings on piston rings has been observed previously [1] the high wear of the coatings themselves was not expected since they had a higher hardness than the CKS-36 coating (879±77 HV0,100).

Examination (by SEM) of the worn plasma sprayed coatings revealed two distinct types of behaviour. One gave a rough surface and was attributed to pull out of material from the coating surface, Fig 3a. The second gave a smooth surface due to a much finer-scale wear response of the coating, Fig 3b. The former was more prevalent in coatings that exhibited high wear.

Although there was considerable spread, the S05 rings (5% SiC) exhibited higher wear than the S02 and S10 rings. Additionally there was lower coating wear for the D02 ring compared to the S02 ring.

The high points of the honed liner surface had been worn flat with small grooves parallel to the sliding direction. These latter grooves were more pronounced for the liner samples coupled with the plasma sprayed Al_2O_3-SiC coatings than for the liners paired with the commercial CKS coating. Also visible within the wear tracks were deeper groves remaining from the original (honed) surface topography.

5. DISCUSSION AND FUTURE WORK

The plasma sprayed Al_2O_3-SiC coatings performed poorly in piston-ring simulation tests due

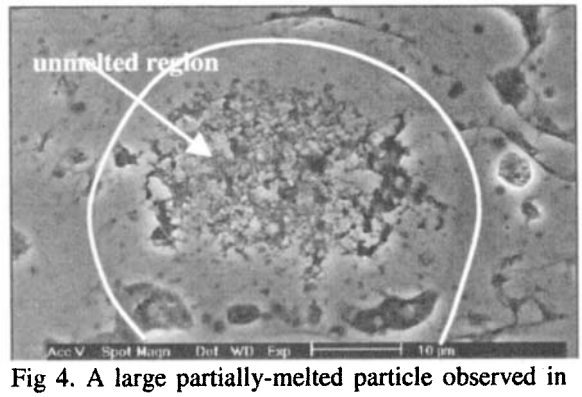


Fig 4. A large partially-melted particle observed in the coating on one of the S02 rings.

to their poor cohesive strength. This allowed large amounts of the coating to be removed during wear testing. This wear mechanism was more prevalent in the higher SiC content coatings and was also sensitive to the surface finishing technique.

Metallographic examination of cross-sections of the coatings revealed the presence of Fe-rich particles (probably formed during the powder production) and unidentified small black inclusions. Most detrimental to the cohesive strength however was the presence of large (up to 30μm in diameter) partially-melted particles within the coatings, Fig 4.

As a result of the present tribological study, significant attention has been paid to improving the coating density (lowering porosity) and in particular to the elimination of unmelted or partially-melted particles within the coatings. Current powders have a much narrower size distribution and the thermal spray parameters have been optimised still further. Tribological testing of this next generation of Al_2O_3-SiC nano-composite coatings is presently underway.

6. CONCLUSIONS

Wear tests have been conducted with plasma sprayed coatings prepared from Al_2O_3-SiC nano-composite powders. The powders contained 2.5, 5 and 10 percent (by volume) SiC. The following conclusions were drawn.

1. Following satisfactory powder characterisation, Al_2O_3-SiC nano-composite coatings were successfully deposited using the nano-composite feed powders. Optimisation of spray parameters gave microhardness behaviour equivalent to or superior to commercial plasma sprayed pure alumina coatings.
2. The plasma sprayed nano-composite coatings showed poor wear resistance in simulated piston-ring tests compared to an existing commercial CKS-36 coated piston ring. Two types of wear mechanism were observed for the sprayed coatings: (i) cohesive failure (grain pull-out due to cohesive failure) and (ii) smooth wear. Cohesive failure was related to the presence of unmelted particles (approximately 30μm in diameter) in the coatings.
3. The wear of the plasma spray coatings was highest for the 5% SiC coatings and lowest for the 2.5% SiC coatings. The use of diamond lapping procedures was also found to give lower coating wear than when the commercial SiC finishing process was used.

REFERENCES

1. U. Buran et al, in *Surface Modifications and Coatings*, R.D. Sisson ed., ASM (1985) 255-264.
2. J.C. Bell & K.M. Delargy, in *Mechanics of Coatings*, Proceedings of the 16th Leeds-Lyon Symposium on Tribology, D. Dowson et al eds, Elsevier, Amsterdam, (1989) 371-378.
3. M. Delaët, in *Surface Modification Technologies XI*, T.S. Sudarshan et al eds, The Institute of Materials, London (1998) 256-273.
4. K. Niihara, in *Advanced Structural Inorganic Composites*, P. Vincenzini Ed., Elsevier Science Publishers, London (1990) 637-664.
5. M. Sternitzke, Journal of the European Ceramic Society 17 (1997) 1061-1082.
6. H.K. Schmid et al, Journal of the European Ceramic Society 18 (1998) 39-49.
7. R.W. Davidge et al, Journal of the European Ceramic Society 16 (1996) 799-802.
8. W.J. Griffiths & F.-G. Cantow, T&N Technical Symposium 1995 (1995), paper 23.
9. K.L. Dahm & P.A. Dearnley in *Boundary and Mixed Lubrication; Science and Applications*, Proceedings of the 28th Leeds-Lyon Symposium on Tribology, D. Dowson et al eds, Elsevier, Amsterdam, (2002) 243-246.

ACKNOWLEDGEMENTS

This research (BRST-CT98-5490) was funded by the European Commission under the Innovative Manufacturing Technologies Programme of Framework IV. The authors are grateful to all partners in the project, especially Federal Mogul Wiesbaden, Plasma Coatings Ltd, the University of Siegen and Zoz GmbH.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

AN INITIAL INVESTIGATION OF THE LUBRICANT CONDITION IN THE AUTOMOTIVE RING ZONE UNDER COLD START CONDITIONS

Asif N G Datoo and Malcolm F Fox

Lubricant Research Centre, School of Molecular Sciences, De Montfort University, Leicester LE1 9BH, UK

Samples of lubricant obtained from the piston ring of a petrol engine under cold start conditions are very different from the sump lubricant condition. Samples from the first operating hour show substantial fuel and water dilution which cause precipitation of additive from the lubrication formulation with multiple phase formation. The physical performance properties of these lubricant samples are drastically changed, e.g., Base Number drops from 10 to less than 2, anti-oxidancy is reduced and viscometric properties are changed. Implications for short vehicle journeys and temperature distributions within the piston assembly are discussed.

1. Introduction

The relevance of studying the engine cold start condition is that most, >90%, vehicle journeys take less than 60 minutes. Most light vehicle engines on necessarily short urban journeys do not reach their design operating temperature. The trend in API[1] and ACEA engine test lubricant procedures, e.g., the SK, SL, series, is to include low power/rpm sequences in the overall test procedure as a partial recognition of the role of urban, short distance, low power, automotive operation. This mode of operation leads to extensive fuel and water condensation within the lubricant. There is increasing interest in the initial, relatively cold, lubricant condition within engine under these conditions.

Lubricants are balanced formulations and introducing fuel and water condensates destabilise the additive/base oil equilibrium. It is therefore of interest to consider the chemical and physical properties of lubricants under these conditions, such as viscosity, anti-oxidant and anti-wear properties.

The ring zone sampling technique first described by Cupples et al[2a], developed for the Petter W-1 engine[2b] and for the Caterpillar 3406B engine[2c], is well suited to investigate the cold start condition as samples may be obtained directly from positions within the ring pack at relatively short time intervals. Indeed, Cupples et al[2a] reported extensive fuel dilution of up to 27% in ring zone samples obtained from a CAT 1Y-540 during the first 15 minutes of cold start engine operation.

The samples obtained from the ring zone may then be quantitatively analysed for fuel and water condensation/dilution and associated effects on the physical and chemical properties of the lubricant formulation.

2. Lubricants, Sampling and Analytical Methods

A range of four lubricant formulations were obtained with essentially common additive packages but with base oil compositions ranging from 100% Poly-Alpha-Olefin (PAO) (Lubricant A) through 90/10 PAO/Ester (Lubricant B), 60/40 PAO/Ester (Lubricant C) to 100% Ester (Lubricant D).

Samples of these lubricants were taken from the first ring groove of a Petter W-1 petrol engine running under the IP 176 procedure[3] for 15, 30, 45, 60, 90, 120 and 180 minutes. Multiple test run samples using a fresh lubricant charge for each test and each lubricant formulation were 'bulked'

together to generate sufficient sample volume for subsequent chemical and physical analyses, due to the small volumes collected.

A separate and parallel set of runs for the same lubricants collected samples at one hour intervals for up to three hours. All of the samples were obtained directly from the first ring groove as a 'cold start' from ambient temperatures of 11-13°C within the engine test cell with a fresh oil charge for each test run. Samples were collected via a flexible tube in a system analogous to that described in the cited examples within reference 2. The sampling system delivers a flow of gas containing fine oil particles which impinge and coalesce on the walls of a glass collection vial.

The percentage mass of condensed water and fuel in the samples was measured using a ThermoGravimetric Analyser (TGA) though the TGA version of the NOACK volatility method[4], a typical TGA analysis for a cold start lubricant sample is given in Fig 1 showing the considerable mass of water and petrol fuel residue lost at relatively low temperatures.

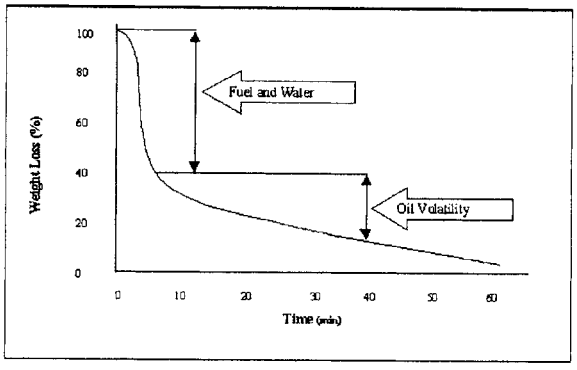


Fig 1 – 'TGA NOACK' Determination for Condensed Water and Fuel from a 'Cold Start' Ring Zone Lubricant Sample.

The concentration of part of the anti-oxidant additive packages present was measured using the 'RULER'[5] ('Residual Useful Lubricant Reserve') analytical method which is a cyclic voltammagram procedure. As an electrochemical method, it directly addresses and measures the overall anti-oxidant reserve within the lubricant. The method is very useful for the very small sample volumes available for analysis arising from this work.

Viscosities were measured using a SETA-VIS[6] microviscometer, requiring 0.23 ml per sample determination, ideal for the small volumes of samples available for analysis.

Sample Base Numbers were measured using the conductimetric IP400 method. This method is ideal for small samples, determinations being made for 0.2 ml of each sample, but reproducible down to 0.1 ml. Some samples readily separated liquid and solids on standing, according to the different base oil compositions. All of the initial, <60 minute samples, were centrifuged at 16000 rpm for 40 minutes in conical tubes to separate the solid phase from the liquid. After the centrifuging the liquid layer was sampled by syringe for FTIR and Base Number analyses.

FTIR spectra were analysed as difference spectra following the procedure of Coates and Setti[7] with integrated absorption areas assigned to the growth of carbonyl oxidation, nitration, water deposition and the degradation of the P=S and P-O-C groups in the anti-oxidant.

3. Sampling Results

The distinctive result of ring zone sampling from operating an engine from a cold start is the extremely degraded nature of the samples. The initial, 60 minute, samples for Lubricants A-D show a brown/buff solid suspended in a two phase fluid, of fuel and water, as shown in the photograph in Fig 2.

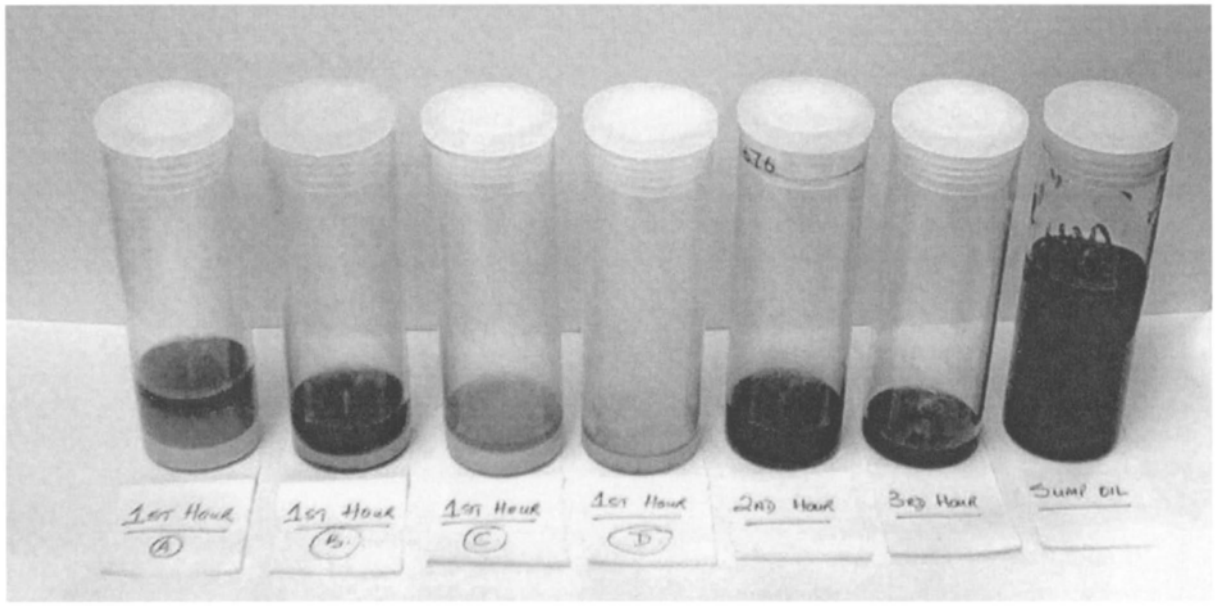


Fig 2 – Degraded One Hour Ring Zone Samples for Different Lubricant Formulations, Showing Multiple Phases for Lubricants A-D as described in Section 2, with Second and Third Hour Zone Samples for Lubricant A and a Three Hour Sump Sample also for Lubricant A on RHS.

The different lubricant formulations studied show a differing mass of samples collected and a different degree of separation, i.e., some sample phases separate readily on standing, others need to be centrifuged to effect separation.

Separation and analysis of the solid phase shows it to be predominantly polymer from the viscosity index improver with traces of other additives present, such as Base Number and anti-oxidant.

The ring zone sample mass varies with time, Fig 3, decreasing sharply with the progress of the test run from 6-9g/hr at 15min, decreasing to ~2-3 g/hr at 60 minutes, according to the base oil formulation.

The interesting point for these samples is that after one hour engine operation the extent of physical degradation gradually declines with time, the samples eventually becoming 'normally degraded', defined as being substantially changed by exposure to the conditions in the ring pack zone but as a homogenous phase.

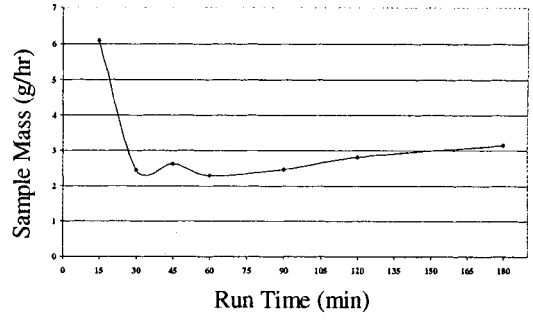


Fig.3 – Lubricant Sample Mass for Formulation 'A' against Run Time.

Different lubricant formulations, primarily in terms of the chemical nature of the base oil, show

different emphases for the same type of effects described. Whilst all lubricant formulations are similarly affected, it is the 60/40 PAO/Ester base oil formulation, Lubricant C, which is affected the most.

The sample viscosities increased by a factor of 4 initially, then declines with time, a typical example being shown in Fig.3. The peak in viscosity is associated with the multi-phase nature, both solid and two liquid phases, of the sample.

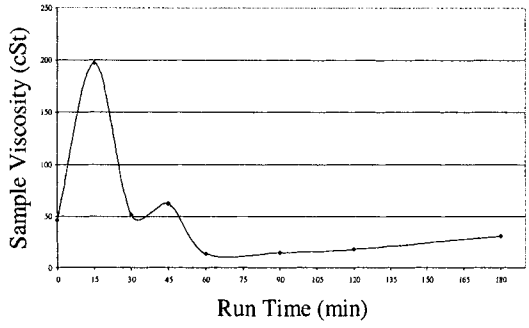


Fig 4 – Variation of Viscosity Against Run Time for Lubricant A.

The decline and subsequent stabilisation at lower values of the sample viscosities at longer run times is associated with the homogenisation of the multiphase samples into a single phase.

For Base Number measurements of the early run time, multi-phase samples, the liquid was separated from the solid using a centrifuge. A 15 minute sample of Lubricant A treated in this way showed that its Base Number value for the liquid phase was sharply reduced from an original value of 10 mg KOH/g down to 5 mg KOH/g, Fig 4. This is followed by a more moderate decline to Base Number values of 2 mg KOH/g, as expected from longer term runs.

It is normally expected that the Base Numbers of lubricants decline with their service life. The Base Numbers of ring zone samples decline more rapidly than for sump samples, usually measured from hour samples. But the short term, 15 and 30 minute, samples show an abnormal decline which is associated with the phase separation of buff/brown deposits. The base additive material would appear to be separating from the base oil together with the other components of the additive pack.

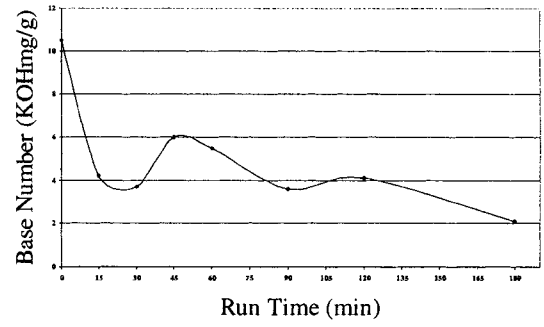


Fig 5 - Variation of Base Number Against Run Time for Lubricant A, Initial Value 10.5 mg KOH/g.

Analysis by the 'RULER' technique[5] (**R**emaining **U**seful **L**ife **E**valuation **R**outine) studies the overall anti-oxidant additive reserve in fresh and used lubricant samples. It is particularly appropriate for the small sample masses available. The RULER voltammogram scan of a 15 minute ring zone sample for lubricant B shows substantial loss of anti-oxidant additive, as evidenced by the relative decline in the peak intensities in the lower curve of Fig 5, relative to the fresh lubricant, the upper curve.

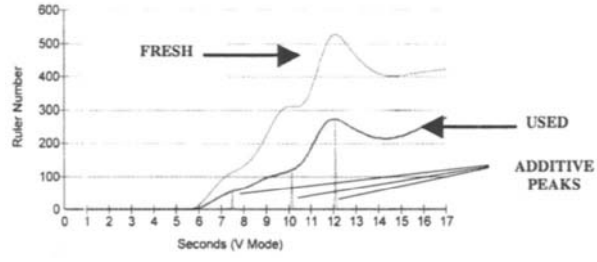


Fig 6 – 'RULER' Plot for 15 Minute Ring Zone Sample of Lubricant B

Of the four lubricant formulations investigated it is the 60/40 mixture of PAO/Ester base oil, Lubricant

C, which is most affected by the deposition of/dilution by water and fuel into the ring zones samples. This is evidenced by its physical separation into additive solid and base oil and also the associated changes in viscosity, Base Number and anti-oxidant presence.

Tribology studies normally assume that there is little or no change of the nature/performance of a lubricant within a contact. Unfortunately, the small volumes of lubricant sample available do not allow such wear tests as the 'pin-on-disc' or '4-ball' tests. But the loss of other components of the additive pack such as the viscosity index improver (VII), Base Number and anti-oxidant, indicates that the anti-wear additive component might also be lost in the early, 15 to 45 minute, ring zone samples.

To investigate the anti-wear properties of the ring zone samples, the contact performance of the cold start samples is being evaluated using High Frequency Reciprocating Rig (HFRR) techniques. The very small samples required for each test are ideal for the small sample volumes available. Initial results from the HFRR test show reduced anti-wear performance for the initial cold start samples. The initial results are in accord with the results from the loss of other additives from the initial run period samples. The further results for the HFRR wear tests will be reported later elsewhere.

4. Discussion

There is strong evidence of extensive lubricant degradation in the piston ring zone of i/c engines under 'cold start' conditions. This leads to a multi-phase formation, with a suspended solid and two liquid phases in the first 30 minutes. This effect decreases with time as the engine temperature increases and an equilibrium single phase system is established after 60 minutes.

There are clear differences between the response of lubricant formulations dependent upon the nature of the base oils used, the most sensitive to fuel and water dilution being the 60/40 PAO/Ester base oil composition, Lubricant C.

It is interesting that the 60/40 PAO/Ester composition of Lubricant C is the most sensitive, one possible explanation is the balance between the low polarity PAO and high polarity Ester being de-stabilised by the introduction of the fuel and water mixture.

The solids deposition, the chemical tests for Base Number and additive presence and also the physical measurement of viscosity indicate that it is the additive package which is separating out from the base oil under the initial conditions in the ring zone.

The implications for short journey, cold start conditions are that, for at least the first hour, the ring groove temperatures require a considerable time to equilibrate at what is normally regarded as 'operating' temperatures. The loss of additives appear to give reduced lubricant performance protection properties for the degraded lubricants under these conditions, compared to the fresh lubricant. Further work is examining the 'cold start condition' samples in greater detail which will be reported later.

We are grateful for research support from De Montfort University, Leicester, BP Castrol Technology Centre, Pangbourne and discussions with John Bullen, Hugh Preston and Chris Warrens of the Castrol Technology Centre.

References

1. API SL, SM type series, ASTM Sequence VG Test.

2. (a) S.D. Cupples, M.F Fox, D. J Picken, F. D. Gainey and S. B. Saville, SAE 881586, (b) D.J Picken, A.L Thomson, M.F Fox and W.H Preston, 'Experimental Methods on Engine Research and Development', IMechE (1991), (c) M.F.Fox, C.J.Jones, D.J.Picken and F.A Kelley, 'Condition Monitoring, 1994', University of Wales (1994), Pineridge Press, Ed. M.H.Jones, pp.470-480.

3. The IP176 Engine Test Procedure for Oil Oxidation and Bearing Corrosion, Institute of Petroleum. (now superseded).

4. TGA/NOACK method CEC-L40-A95 (DIN 51581)

5. The 'RULER' method
http://www.koehlerinstrument.comproducts/K320S-200.htm

6. Stanhope Setavis "Mini" Capillary Viscometer Instruction Manual.

7. J.P. Coates and L.C. Setti, Condition Monitoring of Crankcase Oils Using Computer Aided Infra--Red Spectroscopy, Perkin Elmer Corp.

SESSION XV
ELASTOHYDRODYNAMIC LUBRICATION (2) – EHL – MIXED LUBRICATION

Chairman : Dr H P Evans

Paper XV (i) **Slip at the Wall - Evidence and Tribological Implications**

Paper XV (ii) **Calculation of EHL Contacts in Mixed Lubrication Regime**

Paper XV (iii) **Some Insights into Lubricant Film Thickness in Mixed EHD Point Contacts**

Paper XV (iv) **On the EHL Film Breakdown**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Slip at the wall – evidence and tribological implications

H A Spikes

Tribology Section,
Department of mechanical Engineering
Imperial College
Exhibition Road, London, SW7 2BX, United Kingdom

A standard assumption of fluid mechanics and the underlying principle of all current fluid film bearings is "no-slip at the wall"; i.e. that the immediate layer of liquid next to a solid surface moves with the same velocity as the solid surface itself. However in the last few years it has been shown experimentally that simple, Newtonian liquids, including alkanes, can slip against solid surfaces. This slip occurs when the solid surface is very smooth and the liquid is only "weakly-bonded" to the solid, which in practical terms means that the liquid does not fully wet the solid.

This paper first outlines the main evidence for such slip and then describes some of the possible implications of slip on lubrication. One interesting possibility is the development of a fluid film bearing in which the liquid lubricant is able to slip against one solid surface but not against the other. Analysis shows that such a "half-wetted bearing" combines fluid entrainment, and consequent load support, with very low friction and thus may provide a new way to lubricate very thin film and low-load sliding contacts. A second implication is that conventional organic friction modifiers, which are generally presumed to act in boundary lubrication by reducing asperity-asperity interaction, may also reduce friction in thin film, hydrodynamic or elastohydrodyamic lubrication conditions by promoting local slip of the liquid against the bearing surfaces.

1. INTRODUCTION

A basic tenet of fluid mechanics is "no slip at the wall." As expressed in one standard text;

In other words there is no relative motion between the fluid and the solid. This fact may seem surprising but it is undoubtedly true. No matter how smooth the solid surface or how small the viscosity of the fluid, the particles immediately adjacent to the surface do not move relative to it. It is perhaps not without interest that Newton's term for viscosity was 'defectus lubricitatis' – 'lack of slipperiness'. Even for a fluid that does not 'wet' the surface this rule is not violated. [1]

This principle provides the boundary conditions upon which Reynolds equation and thus fluid film lubrication is based. Because there is no slip at the wall, the motion of one or both of the surfaces of a bearing drags lubricant into the contact at a rate dependent on the mean speed of the two surfaces with respect to the bearing. In consequence, the hydrodynamic film thickness depends upon this so-called "entrainment speed".

Despite the widespread acceptance of this no-slip assumption, there has existed for many years, indirect experimental evidence based on anomalous flow in capillaries and other systems that in some cases, simple liquids can slip against solids. Recently this evidence has become much more compelling and it is now reasonably certain that

Newtonian fluids can slip against solid surfaces if two conditions are satisfied. These are, in direct contradiction to the quotation above; (i) that the fluid does not fully wet the solid surface; (ii) the solid surface is very smooth. As might be expected, such slip has significant implications on lubrication and the aim of the current paper is to explore a few of these.

The paper first outlines evidence for liquid slip at solid surfaces. It then briefly describes the existing models for how such slip occurs. Then the implications of this on both fluid film and boundary lubrication are discussed.

It should be noted that the slip of concern to this paper is that of simple Newtonian liquids against solid surfaces and not the much more widely-accepted slip that can occur when some polymer melt systems are sheared, that is believed to be due to disentanglement of polymer chains attached to the surface from those in the bulk fluid [2].

2. EXPERIMENTAL EVIDENCE

2.1 Capillary Flow Work

Early work on slip at solid surfaces by simple fluids has been reviewed recently by Vinogradova [3] and will not be reiterated in detail here. This was almost all based on capillary flow measurement and essentially showed that the rate of flow of liquids through capillaries can exceed conventional hydrodynamic theory if the liquid and capillary surfaces are chosen so that the solid only marginally wets the solid. (In this paper the term "lyophobic" is used to describe a solid that is not wetted or only partially wetted by a liquid so that the liquid forms a large contact angle against it, as shown schematically in figure 1. The terms "hydrophobic" or "oleophobic" may also be used, though generally for the specific liquids water and oil respectively). Systems that have shown enhanced capillary flow include water in capillaries treated with polar surfactants and silanes, hydrocarbon liquids in fluorosilane-treated capillaries and mercury in glass capillaries.

Figure 2, taken from [3] and based on work of Churaev et al. [4] compares flow rates under pressure of two fluids, tetrachloromethane (CCl_4) and water, through a horizontal quartz capillary. In the test using water the capillary was treated with trimethylchlorosilane to render its surface lyophobic. Despite the two fluids having essentially the same dynamic viscosity (coincident at 26.2°C), the water flows more rapidly. This was ascribed to non-wetting and consequent slip of the water against the lyophobized capillary bore.

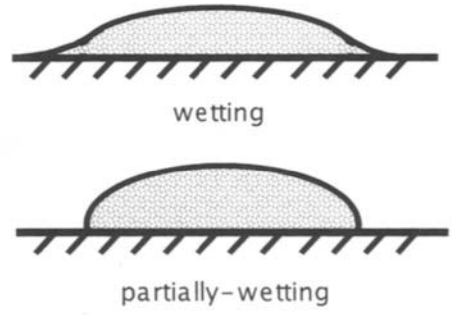


Fig. 1. Wetting and partially/non-wetting liquids

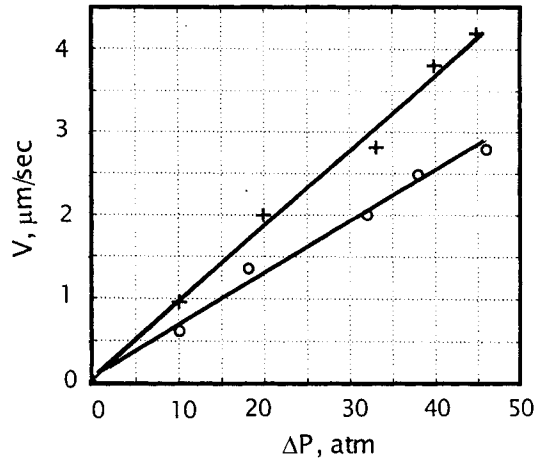


Fig. 2. Flow rate of two liquids through a horizontal capillary at 23.7°C. Crosses are water and circles CCl_4 from refs. (3)(4)

Although this effect has been repeated on many systems and by a number of workers over the last sixty years, this has not been sufficient to convince sceptics of the reality of slip at the wall. One reason for this may be poor repeatability due to difficulty in reliably treating the inside wall of capillaries. Recently, however, two studies based on flow other than in capillaries have provided evidence that is less easy to ignore.

2.2 Fluorescence Recovery Work

Pit and co-workers used a fluorescence recovery after photobleaching (FRAP) technique to assess the velocity of liquid close to a solid surface under simple Couette shear [5][6]. A 90 µm thick film of

liquid containing fluorescent marker was sheared between a moving silica and a stationary sapphire (alumina) flat. Periodically, a narrow column of fluid through the liquid was bleached by an intense laser beam and the rate of replacement of the bleached column by unbleached fluid due to flow was monitored through the transparent alumina. The probe beam employed to monitor the arrival of unbleached fluid used an evanescent wave principle and thus examined only a very thin film of liquid, within 80 nm of the alumina surface. The test liquid employed was hexadecane.

The authors found that when a monolayer of octadecyltrichlorosilane was grafted on the alumina surface so as to render it lyophobic to hexadecane, the rate at which unbleached fluid flowed into the bleached column was enhanced; indicative of more rapid flow of liquid adjacent to the alumina surface. Figure 3 compares how the intensity, I, of the bleached fluid in the column recovered over time at several different shear rates.

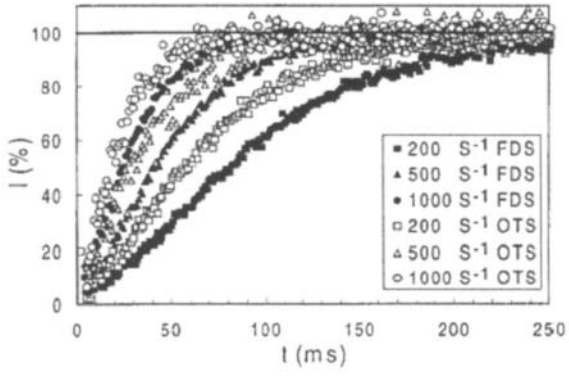


Fig. 3 Comparison of rate of recovery of fluorescence intensity between OTS-coated and FDS-coated alumina surface at four shear rates. (The OTS monolayer is believed to promote slip and FDS to inhibit it). From ref. [5]

In this figure OTS refers to octadecyl-trichlorosilane and FDS to an incomplete monolayer coating of fluorinated surfactant that appeared to inhibit slip. The results were analysed in terms a slip length model. This model is widely used in polymer flow studies and assumes that slip at the wall produces a flow enhancement equivalent to displacing the solid wall by a fixed slip length, b into the slip surface. This length is also called the "Navier length". The model is equivalent to assuming that the slip velocity at the wall, u_s is proportional to the strain rate, $\dot{\gamma}$ i.e. $u_s = b\dot{\gamma}$ and thus that the velocity of slip is proportional to the applied shear stress at the surface, τ_s, i.e. $u_s = b\tau_s / \eta$. Hence a large value of b implies greater slip.

Pit and his co-workers found evidence for slip even for hexadecane on clean sapphire, with b = 175 nm. This slip was prevented by an inhomogeneous coating of FDS. When the sapphire was coated with an OTS layer, slip length increased to b = 400 nm. When the clean sapphire was tested with stearic acid solution in hexadecane, the extent of slip increased with time of immersion of the sapphire to reach b = 350 nm.

2.3 Force Balance Work

The second strand of recent evidence for slip at the wall is due to Zhu and Granick [7][8]. These authors employed a force balance apparatus to study the hydrodynamic squeeze force between a pair of crossed mica cylinders lubricated by both tetradecane and water. As with the alumina surfaces used by Pit et al. above, the mica surfaces were chemically modified, either by grafting on them a lyophobic monolayer of long chain octadecyl-triethoxysiloxane (OTE) or by adding a surfactant (1-hexadecylamine) to the tetradecane solvent. Contact angle measurements showed that these treatments converted the normally fully-wetted mica surface to one only partially wetted by both test liquids.

For tetradecane between untreated, wetted mica, the authors found hydrodynamic squeeze forces in accord with conventional theory. However when the mica was rendered lyophobic, the hydrodynamic squeeze force was drastically reduced. This is shown in figure 4, where the ratio of the squeeze force using treated mica surfaces to that for untreated mica, f^*, is shown as a function of the minimum film thickness between the cylinders, D [7]. It can be seen that when the mica is lyophobic, the squeeze force is reduced by more than an order of magnitude.

Zhu and Granick found that their results did not obey a constant slip length model but implied slip lengths of more than 1 μm at high squeeze velocity and thin film conditions. The author has recently analysed Zhu and Granick's results and shown that they fit quite well a model of slip where the fluid slip only begins when a shear stress of about 0.1 Pa

is reached and then occurs with a very large slip length of more than 1 μm [9].

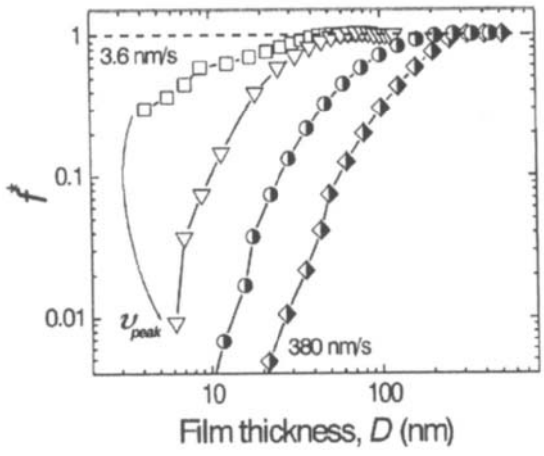


Fig. 4. Squeeze force ratios for deionised water between partially wetted mica surfaces (OTE-coated) at four squeeze velocities. Triangles are for peak velocity 40 nms-1, circles for peak squeeze velocity 100 nms-1. (Squeeze force ratio is the ratio of squeeze force for treated and untreated mica,) From ref. [7]

In a very recent paper, Zhu and Granick have also demonstrated another important requirement for slip - that the partially-wetted surface must be very smooth [8]. By depositing tiny lumps of polymer on the mica surfaces they showed that the tendency of the liquid to slip was progressively reduced as surface roughness was increased and, in the system studied, was effectively prevented when a root mean square (rms) roughness of 6 nm was reached. It should be noted that the alumina surface studied by Pit et al. was also very smooth, with rms 0.4 nm. This sensitivity of slip to surface roughness may explain some of the variability of early capillary work, since it is difficult to precisely control the roughness inside a glass or quartz capillary.

2.4 Molecular Modelling

As well as the experimental evidence outlined above, it should also be noted that several molecular dynamics simulation studies have also demonstrated enhanced mobility or slip of liquid molecules against solid surfaces with which they have only weak interaction [10][11][12].

All of the above work strongly suggests that Newtonian liquids can slip against solid surfaces in both Couette and Poiseuille flow so long as the liquid is only "weakly bonded" to the solid (i.e. only partially wetting and thus with a high contact angle) and the solid is very smooth (i.e. less than 6 nm rms).

3. MODELS OF SLIP

One contributory reason for the lack of general acceptance of the phenomenon of slip at the wall may be the lack of a proven model for how such slip might occur. Although a number of models have been suggested, none have so far been fully tested against experiment. There are, essentially, three main proposed models.

3.1 Enhanced liquid mobility at the interface

The most highly developed model is due to Tolstoi and has been very clearly reviewed by Blake [13]. It is based on an extension of the activated flow model for bulk liquid viscosity. In the latter, viscosity is related to the mobility of individual molecules and thus to the energy required to form a "hole" in the liquid into which a neighbouring liquid molecule can move. This energy depends on the "surface energy" of the hole. In Tolstoi's extension it is suggested that the energy required to form a hole in the liquid monolayer adjacent to a non-wetted solid is much less than in the bulk liquid, since the surface of the hole next to the solid will have low energy of formation due to the liquid having no affinity for the solid. In effect a liquid molecule next to a non-wetted solid surface will behave more as though it were at a vacuum/liquid interface than a solid/liquid one. This model predicts slip with a constant slip length, b that depends on the surface energy of the liquid and extraordinarily critically on the contact angle of the liquid against the solid, as illustrated in figure 5.

3.2 Gas or vapour gap

A number of researchers have suggested that slip may occur because of the formation of a very thin layer of gas or vapour between the solid and the liquid, promoted by the low energy of the liquid/solid interface [14]. Slip then takes place within this film, which is generally presumed to be of nano-scale thickness. This air/vapour gap model has recently been tentatively promoted by deGennes who shows, assuming reasonable values for its main components, that the model predicts a constant slip

length for a given air gap, which was around 7 μm [15]. This model appears to be far less sensitive to contact angle than Tolstoi's one.

Very recently, Zhang et al. investigated the effect of dissolved air on the behaviour of water in a very thin gap between a hydrophobic and a hydrophilic surface [16]. They found that very thorough degassing had no influence on film behaviour, which suggests that if a thin gap is present between liquid and solid this is more likely to be vapour or vacuum than air.

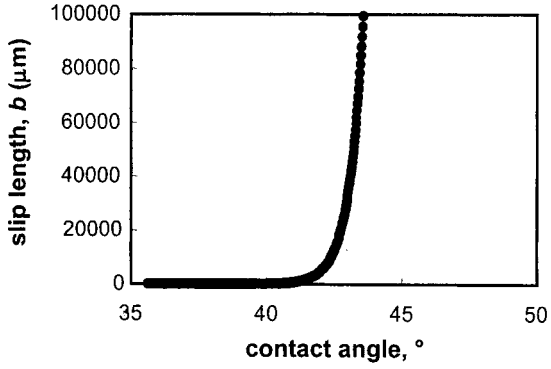


Fig. 5. Illustration of the very strong dependence of slip length on contact angle for Tolstoi's model [13]. Values assumed in this calculation are liquid surface tension 30 mN/m, T=293K, δ = 6.5 nm (δ is related to the distance between molecules in the liquid).

3.3 Low Viscosity Liquid Surface Layer

The third model put forward to explain slip of liquid at a lyophobic solid proposes that the boundary layer of liquid close to the surface of a non-wetted solid has a lower viscosity than the bulk liquid (14). This might result from an enhanced concentration of dissolved gas close to the surface or from some form of molecular alignment induced by proximity of the solid surface. However no clear or quantitative mechanism for this model has been propounded and details remain vague.

As can be seen from the above, there is currently no clear consensus as to how liquids may slip against lyophobic solids. All of the proposed models are however essentially based on the same "fluid-like" principle – of enhanced mobility or reduced viscosity of an interfacial layer, be it a monolayer, a thicker layer of liquid or a gas layer. Probably because of this, they all predict a broadly similar form of slip equation, with a constant slip length. This implies that in all cases some degree of slip will always occur - at a velocity proportional to the shear stress and the slip length and inversely proportional to the fluid viscosity. As described in section 2.3, this is not in accord with Zhu and Granick's results, which suggests either that the slip length is strongly dependent on the squeeze rate or that slip occurs only when a critical shear stress value is reached at the interface. The latter is more in accord with a pinning or "solid-like" response and, if verified by further experimental work, may lead to much fertile modelling.

One current limitation of all existing slip models is that none incorporates the observed strong dependence of slip on surface roughness. This could, however, easily be added to Tolstoi's model, since liquid/solid contact angle is known to be very dependent on roughness [17][18].

4. IMPLICATIONS FOR TRIBOLOGY

At first sight, both non-wetting and any consequent slip of liquid lubricant against solid surfaces in fluid film lubrication would appear to be unequivocally bad in practical terms. Non-wetting will tend to prevent lubricant from adhering to the rubbing solids and also inhibit replenishment of the contact inlet by capillary action. Both of these are likely to result in starvation and thus reduced film generation.

Any tendency for the lubricant to slip against the solid surfaces should also reduce the entrainment into a contact, thereby leading to a thinner, or even no hydrodynamic film formation. This has, indeed, been proposed in the analysis of viscoplastic elastohydrodynamic (EHD) lubrication [19][20]. In this regime of lubrication, the highly pressurised and thus very viscous lubricant is assumed to drastically shear thin when it reaches a critical shear stress under EHD inlet conditions, and then to yield as a plastic solid. In EHD and also hydrodynamic lubrication, so long as thermal effects are neglected, the maximum shear stress should occur at the bounding surfaces. Consequently, in viscoplastic EHD lubrication, a process equivalent to slip should take place at the surfaces at a critical shear stress. This has been shown to lead to possible film collapse.

Notwithstanding the above, there are two types of fluid film lubrication where it has been suggested that non-wetting and slip may be utilised to

advantage, as described in Sections 4.1 and 4.2 below.

4.1 The Non-wetted Bearing

In this type of bearing, both solid surfaces of a contact are non-wetted, or only very slightly wetted, by the liquid and the lubricant is also localised to the vicinity the contact itself. This offers the potential advantage of load-support due to Laplace pressure.

The Laplace pressure results from the curvature of the air/liquid surface and is given by;

$$\Delta p = \gamma_{LV}\left(\frac{1}{R_1}+\frac{1}{R_2}\right) \qquad (1)$$

where γ_{LV} is the interfacial tension of the liquid-air interface and R_1 and R_2 are the radii of curvature of the surface of the liquid in two orthogonal directions. The pressure is greatest on the concave side of the interface. For very thin films, only the radius of curvature across the film thickness need be considered and then the pressure difference approximates to;

$$\Delta p = \frac{\gamma_{LV}}{h}\left(-\cos\theta_1 - \cos\theta_2\right) \qquad (2)$$

where h is the film thickness and θ_1 and θ_2 are the contact angles of the liquid at the two solid surfaces. Figure 6 shows the two cases of a thin film of liquid in a contact where (i) the liquid wets the solids and has low contact angle, (ii) it does not wet or only partially wets the solids and has high contact angle. In the first case, the Laplace pressure is negative and thus "sucks" the surfaces together, thereby enhancing the effective load. In the non-wetted case, however, the Laplace pressure is positive and will help support any externally-applied load. The levels of pressure generated are quite small, ranging typically from –1 MPa to +0.5 MPa but the latter, positive value may be valuable in low load contacts such as in micromachines.

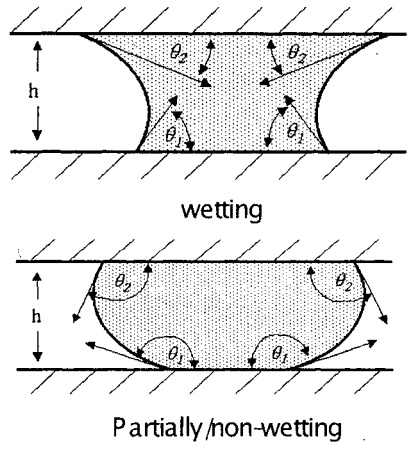


Fig. 6. Contact angles and meniscus curvature for a droplet of liquid between two solid surfaces

For such a non-wetted bearing to be possible, some means of retaining the lubricant at the non-wetted surfaces is necessary. Chappius has proposed a mercury-lubricated journal bearing (mercury does not wet most metals) in which small pockets retain the liquid [21]. Much more recently, Hiratsuka et al. have suggested a bearing based on droplets of water held between two non-wetted PTFE surfaces, with localised areas of one of the surface treated to make them hydrophilic so that tiny droplets of water cling to these regions [22].

In neither of the above non-wetted bearings has the phenomenon of slip been considered, although, as shown below, it should result in a significant reduction in friction.

4.2 The Half Wetted Bearing

The author of this paper has recently proposed a new type of hydrodynamic bearing termed the "half-wetted bearing". In this, the lubricant and solid surfaces of a bearing are selected so that the lubricant wets, and thus forms a no-slip boundary with one bearing surface, but does not wet, and thus can slip, against the other (23)(9). Based on Section 2 above, one way that this can be achieved is by making the latter surface very smooth and coating it with a lyophobic monolayer. Analysis suggests that this type of bearing may provide a combination of good load support with extremely low friction in conditions where conventional hydrodynamic bearings give unacceptably high friction – in very low load and in very thin film conditions.

A linear, half-wetted bearing is shown schematically in figure 7. The fluid is assumed to

wet the lower, moving surface, and thus to have velocity u_1 at the interface. At the upper, stationary surface the fluid is assumed to slip when the shear stress at the surface reaches the value τ_c.

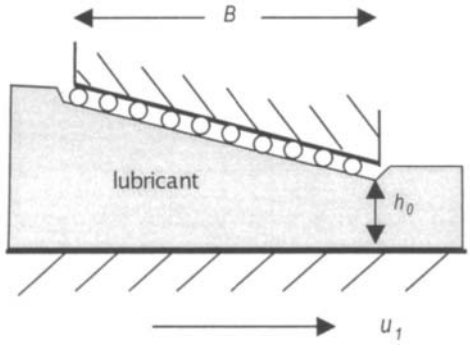


Fig. 7. Schematic diagram of a fully-flooded, linear, half-wetted bearing. The lubricant wets the lower, moving surface but can slip against the upper, stationary surface; (potential for slip is shown by small rollers). From ref. [23]

The conventional Reynolds equation cannot be used to describe such a bearing since it assumes no-slip at both surfaces. However an extended version of Reynolds equation can be developed based on one surface having a conventional no-slip boundary condition and the other a fixed shear stress boundary condition, $\tau = \tau_c$ [23]. This shows that, when slip occurs, the pressure, p, generated in a 1-D bearing is given by;

$$\frac{d}{dx}\left\{h^3\frac{dp}{dx}\right\}=3u_1\eta\frac{dh}{dx}-\frac{3}{2}\mathrm{sgn}(u_s)\frac{d}{dx}\left\{\tau_c|h^2\right\} \quad (3)$$

where h is the corresponding film thickness and u_s is the velocity of the fluid at the slipping surface. sgn(u_s) denotes the sign of u_s and is thus positive when the fluid slips in the sliding direction.

In no slip conditions, Reynolds equation applies, i.e.;

$$\frac{d}{dx}\left\{h^3\frac{dp}{dx}\right\}=6u_1\eta\frac{dh}{dx} \quad (4)$$

Comparison of equations 1 and 2 suggests that, for the "ideal slip" case, with $\tau_c = 0$, the pressure generated in a half-wetted bearing is just half of that produced in a conventional, fully-wetted bearing, Thus the load supported is half that of classical no-slip hydrodynamic bearings.

The friction in a hydrodynamic bearing is determined by integrating the shear stress in the liquid film adjacent to the lower, flat bearing surfaces, τ_1. This shear stress is simply the lubricant viscosity multiplied by the velocity gradient at the surface and, for conventional hydrodynamic lubrication, it is given by;

$$\tau_1=\eta\left(\frac{du}{dz}\right)_1=-\frac{u_1\eta}{h}-\frac{h}{2}\frac{dp}{dx} \quad (5)$$

This expression has two terms. The first is the Couette shear and represents the linear velocity gradient component across the fluid film. The second term is the superimposed Poiseuille flow resulting from the retarding or accelerating effect of the pressure gradient on the fluid within the convergent gap. In conventional hydrodynamic bearings the Couette term is generally much larger than the Poiseuille term, especially at low loads and in very thin film conditions.

For the half-wetted bearing, when slip occurs the corresponding shear stress equation is;

$$\tau_1=\eta\left(\frac{du}{dz}\right)_1=-\tau_c-h\frac{dp}{dx} \quad (6)$$

The Couette term is replaced by the value of the critical shear stress at the slipping surface while the Poiseuille term is twice as large as in the no-slip case. In the "ideal slip" case where $\tau_c = 0$, this means that there is no Couette friction, only Poiseuille friction. The latter is typically only a few percent of the former and also decreases rather than increases as lubricant film thickness is reduced.

To determine the actual performance of a half wetted bearing in terms of load support and friction, equation 3 or its 2-D equivalent must be integrated over a specific bearing geometry to determine the fluid pressure field. This can be done analytically for simple cases or using finite difference method for more complex ones such as 2-dimensional bearings or when slip occurs over part of the bearing surface but not elsewhere. It is also necessary to specify a value of the critical shear stress, τ_c at which slip occurs and this value may be allowed to vary over the bearing and may thus, if desired, be assumed to be a function of slip speed or pressure (23).

Figure 8 indicates how the predicted friction coefficient of a half-wetted 1-D, fixed, linear bearing varies with supplied lubricant thickness, h_s, at a constant bearing load. The insert shows the wedge geometry. Several τ_c conditions are shown, ranging from a very high value, which ensures no-slip under any condition, to the ideal slip case, $\tau_c = 0$. (In determining this figure, τ_c was assumed constant over the bearing, but quite similar results were obtained when τ_c was assumed to vary in proportion to the local slip speed, equivalent to the constant slip length model).

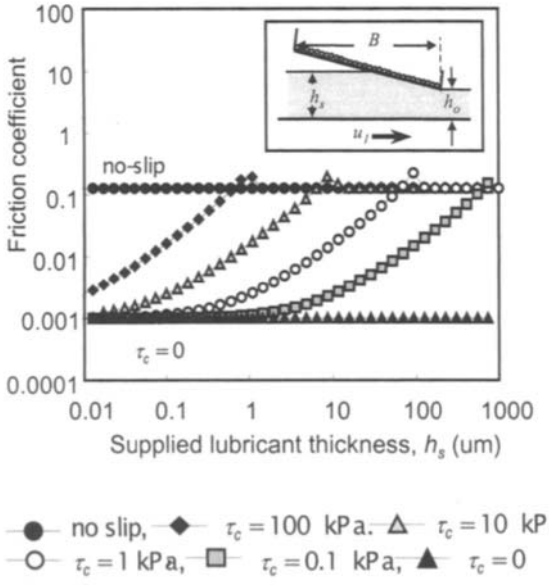


Fig. 8. Influence of supplied lubricant thickness on friction coefficient for various slip criteria. Linear bearing, width B = 1 mm, bearing slope = 0.06°, applied load per unit length, W_L = 100 N/m, sliding speed, u_1 = 1 m/s, lubricant viscosity, η = 0.1 Pas. Insert shows bearing geometry.

From figure 8 it can be seen that the no-slip condition gives a friction coefficient of about 0.1 while ideal-slip gives about one hundredth of this, since friction originates solely from Poiseuille flow. (The friction coefficient does not vary with supply thickness since, as more oil is made available, the convergence ratio remains constant; only the effective bearing width increases). For finite values of τ_c, the bearing changes from no-slip to slip conditions as the supplied lubricant thickness h_s decreases. This reflects an increase in contact stress as the film thickness is reduced. Figure 9 shows how friction coefficient varies with applied load for the same linear bearing at a constant supplied lubricant thickness. This illustrates the very large friction coefficients produced in conventional no-slip bearings at very low bearing loads; values that are curtailed when slip can occur.

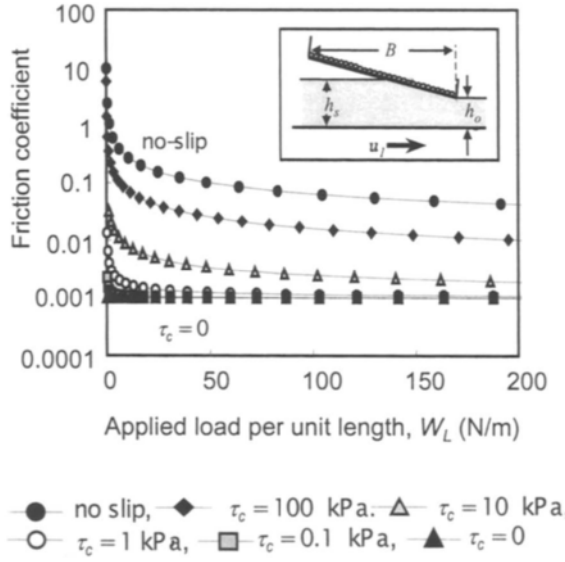


Fig. 9. Influence of bearing load on friction coefficient for various slip criteria. Linear bearing, width B = 1 mm, bearing slope = 0.06°, sliding speed, u_1 = 1 m/s, lubricant viscosity, η = 0.1 Pas, supplied lubricant thickness h_s = 100 nm. Insert shows bearing geometry.

Figures 10 and 11 show similar plots for a 1-D cylindrical bearing. In this there are both converging and diverging regions and the Reynolds boundary condition is assumed in the diverging zone. With this geometry, the half-wetted case gives a very large decrease in friction coefficient in thin film conditions. This is because, as the film thickness is decreased in this shape of bearing, the convergence ratio also decreases. For a conventional no-slip bearing this produces a greater proportion of Couette shear (which is inversely proportional to h), and thus high friction. However in a bearing with slip it simply results in less Poiseuille shear (which is proportional to h).

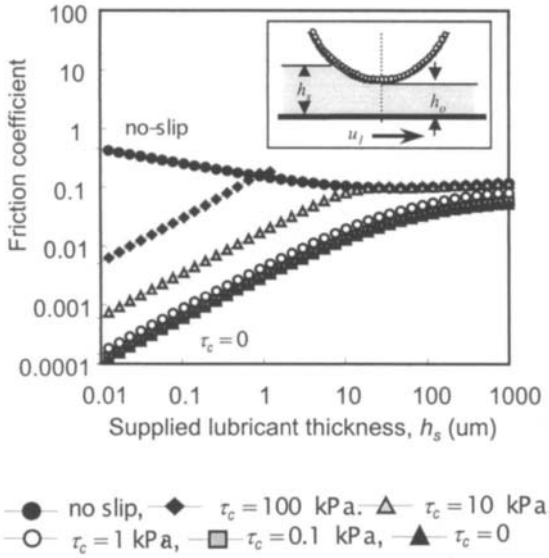


Fig. 10. Influence of supplied lubricant thickness on friction coefficient for various slip criteria. Cylindrical bearing, radius = 10 mm, applied load per unit length, W_L = 100 N/m, sliding speed, u_1 = 1 m/s, lubricant viscosity, η = 0.1 Pas. Insert shows bearing geometry.

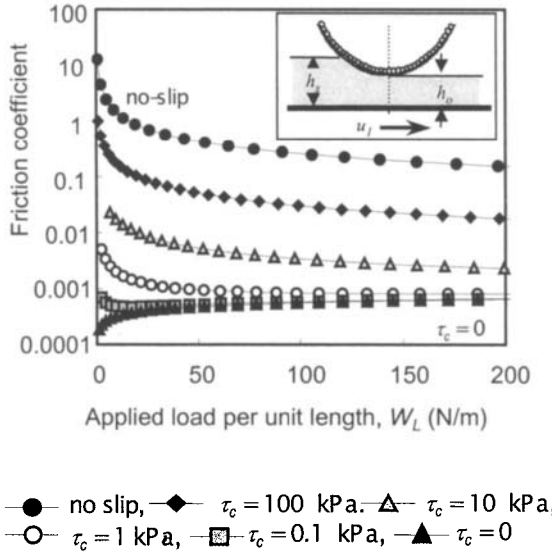


Fig. 11. Influence of bearing load on friction coefficient for various slip criteria. Cylindrical bearing, radius = 10 mm, sliding speed, u_1 = 1 m/s, lubricant viscosity, η = 0.1 Pas, supplied lubricant thickness, h_s = 100 nm. Insert shows bearing geometry.

It should be noted that a half-wetted bearing is only possible if the wetted surface moves with respect to the bearing (i.e. the convergent zone). If, in figure 7, the lower surface were the wetted one and the upper surface the non-wetted one, no entrainment and thus no load support would be possible. If, however, the lower, wetted surface were stationary and the upper, non-wetted, tilted surface were to translate from right to left, then entrainment and film build-up would occur since the bearing would now be moving with respect to the wetted surface.

The above illustrates the potential of the half-wetted bearing for providing very low friction coefficient in low load and thin film conditions. These are precisely the conditions under which conventional fluid film lubrication gives unacceptably high friction, and occur in micromachines and also in hard disc drives. Such machines also have very smooth surfaces, so it seems possible that the half-wetted bearing approach may provide an effective way of lubricating such systems.

Clearly a crucial practical issue is the value of τ_c actually attainable for a liquid against a partially-wetted surface. Calculation suggests that the non-dimensional group ($\tau_c h_o / u_1 \eta$) needs to be less that about 0.01 for full slip to occur in a half-wetted bearing, where h_o is the minimum film thickness [23]. For the sliding speed and viscosity conditions shown in figures 8 to 11, this implies that, for a half wetted bearing to be fully realized, τ_c must be less than 10 kPa when h_o = 100 nm and less than 100 kPa when h_o = 10 nm.

4.3 Boundary Lubrication

The above relates to the possible impact of slip at the wall on hydrodynamic lubrication by using surfaces tailored to encourage such slip. However it has also been shown experimentally by both Pit et al. [5] and Zhu and Granick [7] that long chain polar compounds such as stearic acid, when allowed to adsorb from hydrocarbon solution on to smooth solid surfaces, can promote slip almost as effectively as chemically-grafted, lyophobic monolayers. It is likely that related commercial organic friction modifiers will behave in a similar fashion. The immersion and in particular the rubbing of metals in surfactant solutions has long been known to promote the formation of highly lyophobic surfaces by friction modifiers, so-called "oleophobic monolayers" [24][25]

The intriguing question that then arises is whether fluid slip, and consequent lower local friction produced by equation 6 above, may be a

significant mechanism of boundary lubrication. This possibility was suggested in the 1930s by Hersey [26]

Perfect adhesion, or freedom from slip at the boundary, is assumed in the modern hydrodynamic theory. Without such tangential adhesion the necessary fluid pressures could not be maintained. Imagine two lubricants of equal viscosity, differing with respect to the maximum tangential stress which can be supported at the boundary. Then if the operating conditions are such as to exceed the critical value for either or both lubricants there will be a difference in the frictional resistance, and we have another explanation of oiliness.

Thus adsorption of additives on rubbing surfaces may produce local, very smooth regions which are highly lyophobic. In this case, under thin film, high shear stress conditions, the fluid should slip against these regions, thereby reducing Couette shear and consequent friction. Clearly such an effect must remain localised, since were most of the two rubbing surfaces to become lyophobic, dewetting and starvation would result.

This concept, that friction modifiers may reduce friction in the full fluid film as well as in the solid-solid contact regime, may help explain the puzzling observation that not only the high load-high temperature but also the low load-low temperature stages of fuel economy engine tests, in which hydrodynamic lubrication should predominate, tend to be responsive to the presence of friction modifier additives

The above are, of course, only a few possible implications of slip at the wall, which extend far beyond tribology into all aspects of micro- and macro-fluid mechanics.

5. CONCLUSIONS

This paper has examined current experimental evidence that simple fluids can slip against solid surfaces. It is clear that such slip is possible, so long as the solid surfaces concerned are extremely smooth (< 6 nm rms) and highly lyophobic to the liquid concerned. One of the simplest ways that otherwise polar solid surfaces can be made lyophobic is by depositing or grafting upon them a long alkyl chain surfactant monolayer.

Some practical implications on liquid lubrication of slip at the wall have been considered. If a lubricant does not wet either rubbing surface, lubricant cannot be supplied or retained in a contact, so fluid film lubrication is probably impossible, regardless of whether slip occurs.

One novel possibility is a "half-wetted bearing." In this, one of the bearing surfaces is deliberately made wetting and the other non-wetting. Such a bearing may provide a very promising combination of hydrodynamic pressure generation and thus load support with extremely low friction due to slip.

Liquid slip can be promoted by the adsorption of long chain surfactants from solution on to smooth solid surfaces. This suggests a possible new mechanism of boundary lubrication by organic friction modifiers – encouragement of local slip of the liquid lubricant against the bounding solids to produce a local reduction in shear stress. Such a mechanism would imply that friction modifiers might be able reduce friction in full film lubrication conditions.

REFERENCES

1. B S Massey, *Mechanics of Fluids, Section 5.6, 6th Edition,* publ Chapman and Hall, London 1989.
2. deGennes, P-G., "Viscometric flows of tangled polymers." *C.R. Acad. Sci. Ser. B*, 288, (1979) pp. 219-220.
3. Vinogradova, O.I., "Slippage of water over hydrophobic surfaces." *Int. J. Miner Process.* 56, (1999), 31-60.
4. Churaev, N.V., Sobolev, V.D. and Somov, A.N., "Slippage of liquids over lyophobic solid surfaces." *J. Coll. & Interf. Sci.* 97, (1884), 574-581.
5. Pit, R., Hervet, H. and Leger, L., "Friction and slip of a simple liquid at a solid surface." *Tribology Letters* 7, (1999), 147-152.
6. Pit, R., Hervet, H. and Leger, L., "Direct evidence of slip in hexadecane:solid interfaces." *Phys. Rev. Lett.* 85, (2000), 980-983.
7. Zhu, Y and Granick, S., "Rate dependent slip of Newtonian liquid at smooth surfaces." *Phys. Rev. Lett.* 87, (2001), paper 096105.
8. Zhu, Y and Granick, S., "Limits of hydrodynamic no-slip boundary condition." *Phys. Rev. Lett.* 88, (2002), paper 106102.
9. Spikes, H.A., "The half-wetted bearing. Part 2: potential application to low load contacts."

Accepted for publication in *Proc. I.Mech.E Part J*, 2002.

10. Thompson, P.A. and Robbins, M.O., "Shear-flow near solids - epitaxial order and flow boundary-conditions." *Phys Rev.* A41 (1990), 6830-6837.
11. Barrat, J-L. and Bocquet, L., "Large slip effect at a nonwetting fluid-solid interface". *Phys. Rev. Lett.* 82, (1999), 4671-4674.
12. Wang, H., Hu, Y.Z. and Guo, Y., "Molecular dynamics study on interfacial slip phenomenon in ultra-thin lubricating films." Proc. Intern. Trib. Conf., Nagasaki, Oct. 2000, 751-762. publ. JST, Tokyo.
13. Blake, T.D., "Slip between a liquid and a solid: D M Tolstoi's (1952) theory reconsidered." *Colloids and Surfaces* 47, (1990), 135-145.
14. Vinogradova O.I., "Drainage of a thin film confined between hydrophobic surfaces." *Langmuir* 11, (1995), pp. 2213-2220.
15. deGennes, P-G., "On fluid/wall slippage". *Langmuir* 18, (2002), 3213-3414.
16. Zhang, X., Zhu, Y. and Granick, S., "Hydrophobicity at a Janus interface." *Science* 295, (2002), 663-666.
17. deGennes, P-G., "Wetting: statics and dynamics." *Reviews of Modern Phys.* 57, (1985), 827-863.
18. Chow, T.S., "Wetting of rough surfaces." *J. Phys. Condens. Matter* 10, (1998), L445-L451.
19. Wilson, W.R.D. and Aggarwal, B.B., "A plastohydrodynamic inlet zone analysis for a viscoplastic lubricant." *Wear* 47, (1978), 119-132, (1978).
20 Wen, S. and Zhang, Y., "EHL performance of the lubricant with shear strength. Part I – boundary slippage and film failure." *Trib. Trans.* 43, (2000), 700-710.
21. Chappuis, J., "Lubrication by a new principle: the use of non-wetting liquids." *Wear* 77, (1982), 303-313.
22. Hiratsuka, K., Bohno, A. and Kurosawa, M., (2000), "Ultra-low friction between water droplet and hydrophobic surface." Proc. NATO Course, Keszthely, Hungary, August 2000, *Fundamentals of Tribology and Bridging the Gap between Macro- and Micro/Nanoscales.*, pp. 345-348, ed. B Bhushan, publ. Kluwer Acad. Publ. Netherlands.
23. Spikes, H.A., "The half-wetted bearing. Part I: extended Reynolds equation." Accepted for publication in *Proc. I.Mech.E Part J*, 2002.
24. Bigelow, W.C., Pickett, D.L. and Zisman, W.A., "Oleophobic monolayers. I Films adsorbed from solution in non-polar liquids." *J. Coll. Sci.* 1, (1946), 513-538.
25. Masuko, M., Toda, Y. and Suzuki, A., "Grazing-angle reflection FT-IR spectroscopy of adsorbed molecular films on the rubbed steel surface." Proc. Intern. Trib. Conf., Nagasaki, Oct. 2000, pp. 493-497. publ. JST, Tokyo.
26. M D Hersey, *Theory of Lubrication*, J.Wiley & Sons Inc. New York, 1936.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Calculation of EHL Contacts in Mixed Lubrication Regime

A. C. Redlich[a], D. Bartel [b] and L. Deters [b]

[a] Robert Bosch GmbH, FV/FLM, P.O. Box 10 60 50,
D-70049 Stuttgart, Germany, alexander.redlich@de.bosch.com

[b] Machine Elements and Tribology, Otto-von-Guericke-University Magdeburg
Universitaetsplatz 2, D-39106 Magdeburg, Germany

A calculation model is presented for machine elements operating under mixed lubrication regime. The following conditions are assumed: elastohydrodynamic point contact; stationary rough surface to study pure sliding conditions; non-Newtonian Ree-Eyring fluid model; linear-elastic, ideal-plastic material behaviour and isothermal conditions. Under these constraints some detailed investigations within the mixed lubrication regime are made.

1. INTRODUCTION

For solving tribological problems, simulation tools gain an increasing importance. Particularly the calculation methods for elasto-hydrodynamic contacts were developed further during the last years, so that film thickness-, pressure- and temperature distribution as well as elastic deformations of the surfaces can be calculated successfully. The aim of this study is to broaden this knowledge regarding surface topographies in mixed lubrication regime. It can be seen as a continuation of [1].

An overview of the theoretical models in the mixed lubrication regime can be found at Nakahara [2] or Chang [3]. Starting from stochastic approaches like Makino *et al.* [4] more and more deterministic studies were developed, e.g. Jiang *et al.* [5] or Hu and Zhu [6]. Unfortunately the deterministic approach has to cope with some numerical problems, especially with the treatment of the boundary condition between fluid film and solid contact regions. Jiang *et al.* use a continuous pressure boundary condition whereas Hu and Zhu switch off single terms in the Reynolds equation depending on the local film thickness.

In this work a new approach is used to cope with the fictitious solid contact spots. This is demonstrated with the help of a single bump in sliding contact. This simple surface feature enables insight to some important effects within mixed lubrication regime. After these results three different real surface topographies are compared with each other. Finally the friction calculation will be explained and two Stribeck curves are presented.

2. RESULTS WITH SINGLE BUMP

The basic concept of the EHL model is described in [1], i. e. the non-Newtonian Ree-Eyring fluid model according to Ai *et al.* [7] and the implementation of the 2D Reynolds-, elasticity- and load equation according to Venner [8]. The lubricant properties are the viscosity-pressure relationship as proposed by Mihailidis *et al.* [9] and the density-pressure relationship as proposed by Dowson and Higginson. The non-dimensional parameters in this work can be seen in table 1.

Table 1
Nondimensional parameters

W	$3.188 \cdot 10^{-5}$
U	$2.16 \cdot 10^{-11}$
G	4918
L	12.61
M	1892

2.1. Single asperity (bump) in sliding contact

A single bump is studied under pure sliding conditions. The sliding direction is always along the x-axis. Equation (1) describes the geometry of the bump with an amplitude of $h_{bump} = 500$ nm and a wavelength of $\lambda = 10\ \mu$m.

$$d(x,y) = h_{bump} \cdot 10^{-10\left(\frac{x^2+y^2}{\lambda^2}\right)} \cdot \cos\left(2\pi \frac{\sqrt{x^2+y^2}}{\lambda}\right) \tag{1}$$

Fig. 1a and 1b show the bump in the center of an unloaded ball surface.

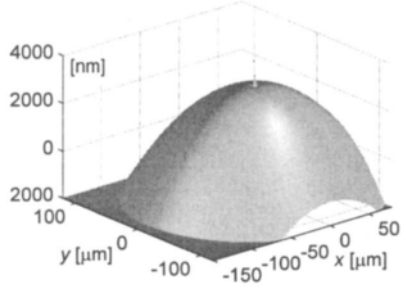


Fig. 1a. Smooth ball surface with single bump.

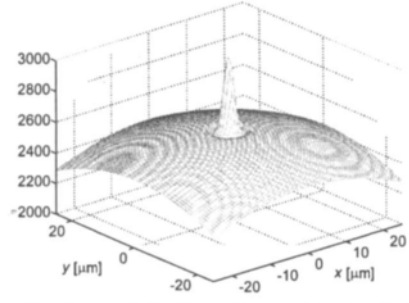


Fig. 1b. Smooth ball surface with single bump.

The elastic deformations as well as the pressure distribution under the given conditions of table 1 are shown in the following graphs of fig. 2 and 3. In the center of the Hertzian contact one can see that high pressure gradients upstream of an asperity lead to unrealistic high pressure values up to 20 GPa. This leads to strong local elastic deformations ahead of the bump. The reason for this is the Reynolds equation which opens the gap between the two surfaces by a local pressure increase. But in this case there is still a negative film thickness at the tip of the asperity (fig. 3, left side). This is just a matter of grid refinement. As soon as the asperity is too sharp for a certain grid (in this work a distance of 638 nm could be achieved between the grid points) the pressure gradients cannot be solved properly and a negative film thickness occurs. These are so called "fictitious solid contact spots". Unfortunately, this results has large consequences since it is not possible to make any statements about the number or area of solid contacts on the basis of a film thickness criterion. Also solid friction cannot be calculated with the help of the film thickness distribution. Due to this fact there is a need for a different approach.

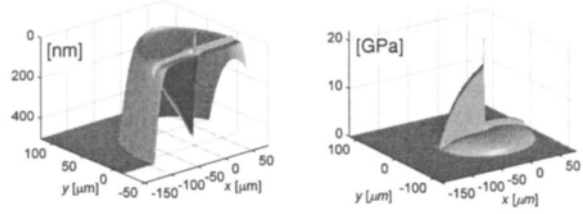


Fig. 2. Film thickness and pressure distribution with single bump in the center of the contact.

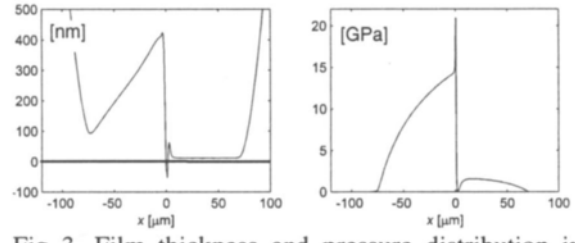


Fig. 3. Film thickness and pressure distribution in the line of symmetry.

2.2. Dry contact solution

As it will be explained in section 2.3 the mixed lubrication regime needs the solution of the dry contact problem. (Meaning of symbols see [8])

$$\begin{pmatrix} -\frac{2}{\pi^2}\Delta K^{hhhh}_{ii-1\,jj} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \delta_{i-1,j} \\ \overline{H}_{i-1,j} \end{pmatrix} + \begin{pmatrix} -\frac{2}{\pi^2}\Delta K^{hhhh}_{iijj} & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \delta_{i,j} \\ \overline{H}_{i,j} \end{pmatrix} + \begin{pmatrix} -\frac{2}{\pi^2}\Delta K^{hhhh}_{ii+1\,jj} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \delta_{i+1,j} \\ \overline{H}_{i+1,j} \end{pmatrix} = \begin{pmatrix} \tilde{H}_{i,j} \\ f\left(\tilde{H}_{i,j}\right) \end{pmatrix} \tag{2}$$

$$f(\tilde{H}_{i,j}) = \begin{cases} 0 & \tilde{H}_{i,j} \le 0 \\ -\tilde{H}_{i,j} & \tilde{H}_{i,j} > 0 \end{cases} \quad (3)$$

The system of equations (2) is derived from the EHL problem without terms of Poiseuille and Couette flow. It can be solved in the same way as the EHL solution. The constraints in equation (3) lead to a pressure build up at spots where film thickness is lower than zero whereas all other grid points do not experience any pressure changes. Therefore the pressure of grid points with both solid contact pressure and film thickness above zero must be reduced by a suitable non-linear function.

2.3. Additional pressure p_{num}

The basis of the new approach is, that the geometrical boundary condition must be valid, i.e. two surfaces cannot penetrate each other. Since the numerical solution of the Reynolds equation on "coarse" grids is not able to fulfil this condition in case of real surface topographies, a new pressure p_{num} is introduced at fictitious solid contact spots to maintain this requirement.

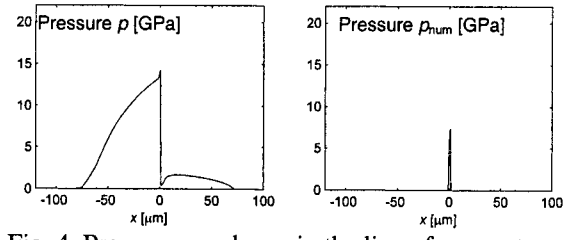


Fig. 4. Pressure p and p_{num} in the line of symmetry.

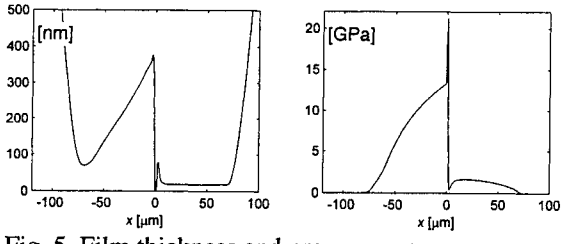


Fig. 5. Film thickness and pressure $p + p_{num}$.

Two remarks must be mentioned: The pressure p_{num} may not be coupled directly with the hydrodynamic pressure p, but must be calculated by the dry contact solution (section 2.2) and secondly the pressure p_{num} is not a solid contact pressure, since the Reynolds equation is still valid and therefore no real solid contact spots can be identified. With this approach, the results of fig. 3 change as follows. At the fictitious solid contact spot in the center of the Hertzian contact a sharp pressure spike p_{num} is calculated (fig. 4) which leads to a film thickness equal to zero (fig. 5, left). Additionally, the pressure drop downstream of the bump does not reach zero as before (fig. 5, right). From the numerical point of view, these local pressure gradients cause instabilities during solution process and therefore preventing convergence.

2.4. Linear-elastic, ideal-plastic material behaviour

The next step is to cope with the unrealistic high pressure values above material strength. Due to these high pressure values, the material will deform plastically. For this reason a linear-elastic, ideal-plastic material behaviour was introduced. This is done by the implementation of a new variable h_{plast} into the elasticity equation (4). The value of h_{plast} has to be calculated iteratively:

$$h(x,y) = h_{00} + r(x,y) + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + \frac{2}{\pi E'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(x',y')}{\sqrt{(x-x')^2 + (y-y')^2}} dx'dy' + h_{plast}(x,y) \quad . \quad (4)$$

In this work the plastic pressure limit p_{plast} is set to 4 Gpa and above this value plastic deformations occur. With the new variable h_{plast} the simulation results of the example change as follows: In figures 6 and 7 one can see that the pressure upstream of the bump is limited by the plastic pressure limit. Therefore the elastic deformations upstream of the bump are less than before. The plastic deformation of the bump with an amplitude of 500 nm is h_{plast} = 350 nm. The remaining 150 nm of the bump are deformed elastically and form a micro EHL contact as one can see on the left side of figure 7.

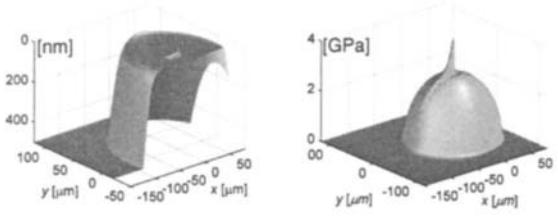


Fig. 6. Film thickness and pressure distribution with linear-elastic, ideal-plastic material behaviour (p_{plast} = 4 GPa).

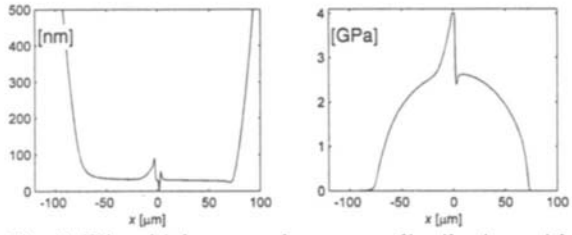


Fig. 7. Film thickness and pressure distribution with linear-elastic, ideal-plastic material behaviour (p_{plast} = 4 GPa) in the line of symmetry.

2.5. Some remarks

The question if the plastic material behaviour would have been sufficient to solve the problem with the fictitious solid contact spots should be clarified:

1) Since the pressure drops down at the fictitious solid contact spots, there will be no plastic deformation at these spots themselves. High pressure values can only be found directly upstream of the bump. Hence, the implementation of the linear-elastic, ideal-plastic material behaviour without p_{num} leads to an increasing plastic deformation upstream of the fictitious solid contact spot during the iterations and not at the contact itself. So no useful solution can be obtained.

2) Concerning the FAS-algorithm it was not possible to get a converged solution with micro asperity or real surface topographies. The local events in the vicinity of surface asperity seem to limit the multigrid algorithm which needs a smooth residual distribution during solution process. (Of course, this does not effect the multi-integration algorithm to solve the elastic deformation equation.) For that reason, solution time increases dramatically, especially in combination with the additional iterations to solve for h_{plast} and p_{num}.

3. RESULTS WITH REAL SURFACE TOPOGRAPHIES

In this section three different real surface topographies will be discussed with the help of the new mixed lubrication approach. All surfaces are measured with an atomic force microscope (AFM). The measured values have a distance of 784.3 nm and has to be transferred to the finest simulation grid by linear interpolation.

3.1. Lapped surface topography

The roughness parameters of the lapped surface topography (see fig. 8) are CLA = 0.12 μm and RMS = 0.15 μm, respectively.

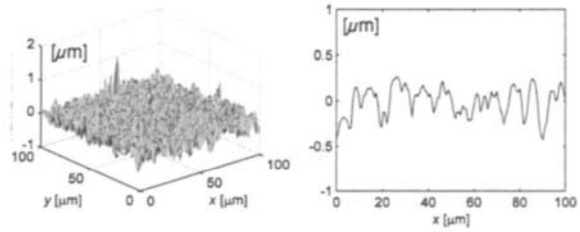


Fig. 8. Lapped surface topography.

The first simulation was done for the dry contact situation as a kind of "worst case" regarding the mixed lubrication regime.

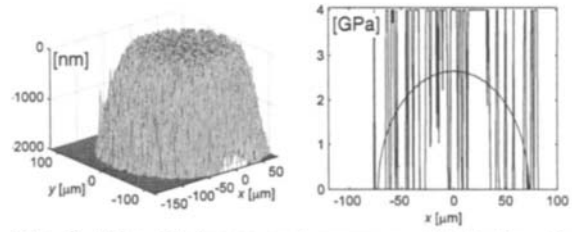


Fig. 9. Film thickness and pressure distribution for dry contact situation.

The smooth line in the pressure distribution (right side of fig. 9) represents the ideal-smooth surface

with the same conditions. In case of the real surface topography, one can see that almost every asperity in contact reaches the plastic pressure limit of 4 GPa. Based on the Hertzian contact area, the real contact area is 45% and the plastically deformed contact area is 37%, i.e. 82% of the real contact area are deformed plastically and 18% elastically.

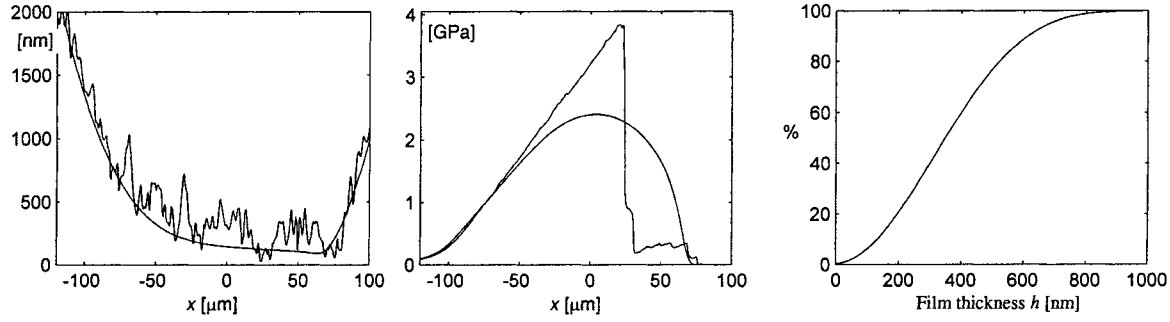


Fig. 11. Film thickness, pressure distribution and slice plane graph for $u_2 = 10.0$ m/s.

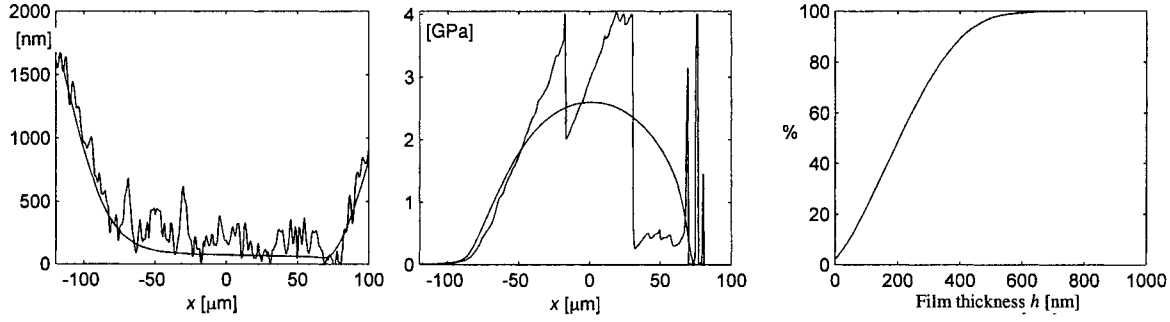


Fig. 12. Film thickness, pressure distribution and slice plane graph for $u_2 = 1.0$ m/s.

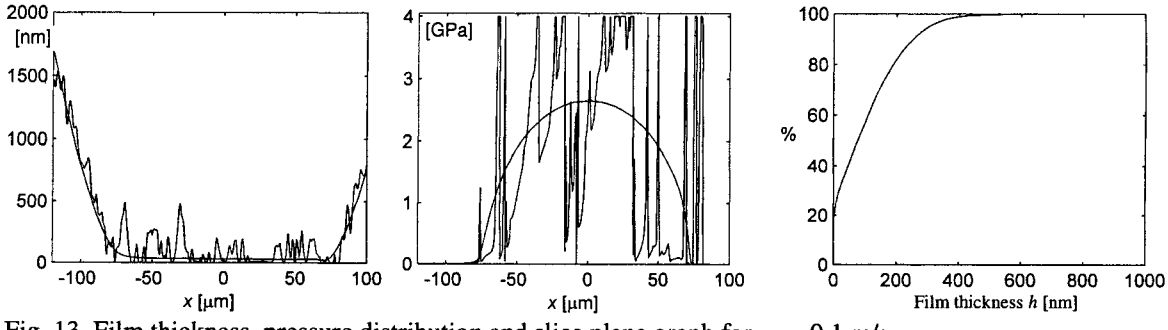


Fig. 13. Film thickness, pressure distribution and slice plane graph for $u_2 = 0.1$ m/s.

A new graph describing the material and fluid film or solid contact portion within the Hertzian contact area will be presented as follows: The slice plane graph is similar to the well known Abbott curve. Here, slice planes parallel to the x-axis through the film thickness distribution on the left side of fig. 10 result in material and fluid film areas (or air in case of the dry contact situation). These areas are then divided by the Hertzian contact area and depicted as fluid- and material portions on the right side of fig. 10. The aim of this graph is to

enable an easy view into the contact and to decide whether it is good lubricated or not.

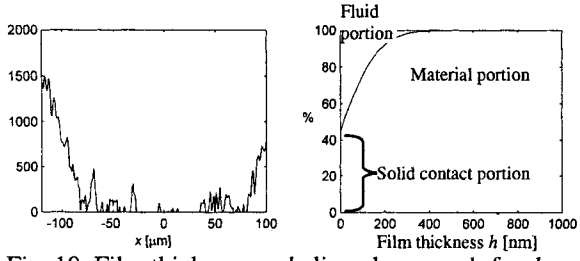


Fig. 10. Film thickness and slice plane graph for dry contact situation.

The next step is to investigate pure sliding conditions for three different velocities: $u_2 = 10.0$, 1.0 and 0.1 m/s. Again the smooth lines in the film thickness and pressure distribution correspond to the smooth surface case. Fig. 11 shows the simulation results for $u_2 = 10.0$ m/s. At this sliding speed, both surfaces are separated completely by a fluid film as one can see in the slice plane graph because there is no material portion at film thickness zero. The pressure distribution in the line of symmetry shows a pressure wedge due to the asperity at about $x = 25\ \mu$m. Downstream of this point pressure does not recover any more.

In case of sliding velocity $u_2 = 1.0$ m/s (fig. 12), the pressure wedge moves upstream due to the micro-asperity at about $x = -20\ \mu$m. But in this case pressure builds up again until the next smallest gap. The pressure spikes at the end of the Hertzian contact are due to the additional pressure p_{num} to satisfy boundary condition that surfaces may not penetrate with each other.

In case of the lowest sliding velocity of $u_2 = 0.1$ m/s (fig. 13), the overall fluid film between the two surfaces is the least. The slice plane graph shows that the material portion in the range of small film thickness has increased significantly and of course the fluid film portion has decreased. The pressure wedges still move upstream and many local pressure spikes occur in the center of the Hertzian contact.

3.2. Polished surface topography

The roughness parameter of the polished surface topography (see fig. 14) are CLA = 0.06 μm and RMS = 0.08 μm, respectively.

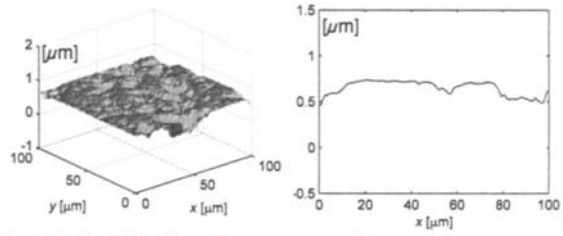


Fig. 14. Polished surface topography.

Again the first simulation studies the dry contact situation.

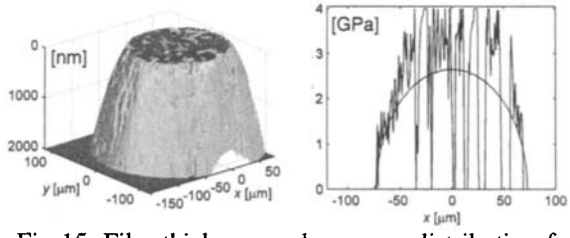


Fig 15. Film thickness and pressure distribution for dry contact situation with polished surface.

Because the polished surface is much smoother than the lapped one, the pressure distribution is more similar to the solution of the smooth surface. Only in the center of the Hertzian contact the pressure reaches the plastic pressure limit of 4 GPa.

The right side of fig. 16 shows that in this case the real contact area is 80% of the nominal contact area and only 6% of the nominal contact area is plastically deformed.

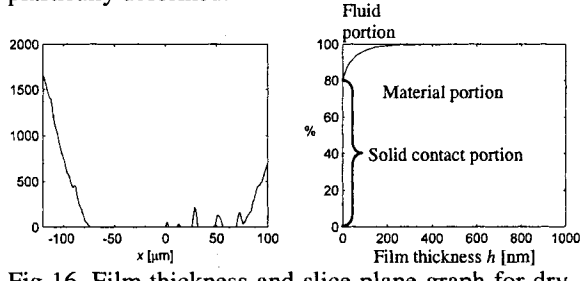


Fig 16. Film thickness and slice plane graph for dry contact situation with polished surface.

Again, the next step is to investigate pure sliding conditions for three sliding velocities (fig. 17-19). It is interesting that in all cases the slice plane graphs show no material portion at zero film thickness, i.e. also for the lowest sliding velocity it still exists an almost complete fluid film between both surfaces. Compared to the lapped, surface the run of the slice plane curves are quite different because the fluid film portions or fluid volumes, respectively (i.e. area above the curves) are considerably smaller for the polished surface.

The pressure distribution for the highest sliding velocity of $u_2 = 10.0$ m/s (fig. 17) is similar to the solution of the smooth surface. With decreasing sliding velocity, some smooth pressure wedges form but do not reach the plastic pressure limit, so no micro asperity will be deformed plastically.

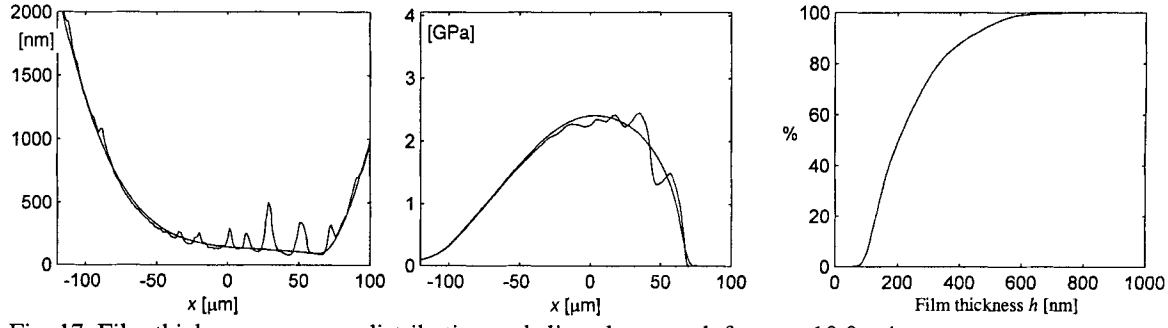


Fig. 17. Film thickness, pressure distribution and slice plane graph for $u_2 = 10.0$ m/s.

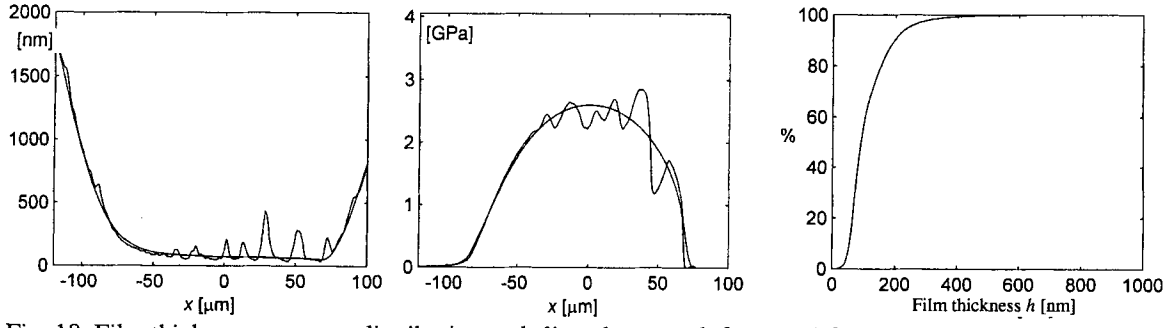


Fig. 18. Film thickness, pressure distribution and slice plane graph for $u_2 = 1.0$ m/s.

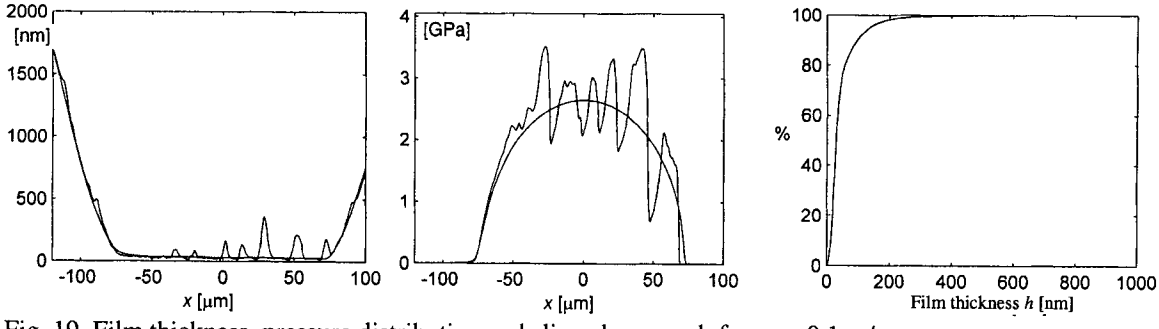


Fig. 19. Film thickness, pressure distribution and slice plane graph for $u_2 = 0.1$ m/s.

3.3. Transversal and longitudinal ground surface topographies

In this section a comparison between a transversal and a longitudinal ground surface is made, i.e. the same surface topography is studied with different sliding directions (fig. 20 and 21). The comparison indicates two main differences. First the slice plane graph of the longitudinal ground surface runs more steep than the transversal one. This means that the fluid volume inside the Hertzian contact region of the longitudinal ground surface is smaller than the fluid volume of the transversal ground surface. The third diagram shows the

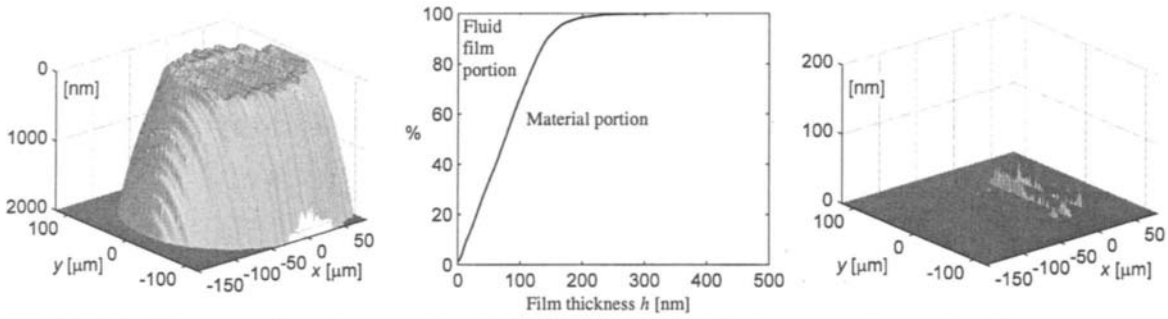


Fig. 20. Film thickness, slice plane graph and plastic deformations for transversal ground surface.

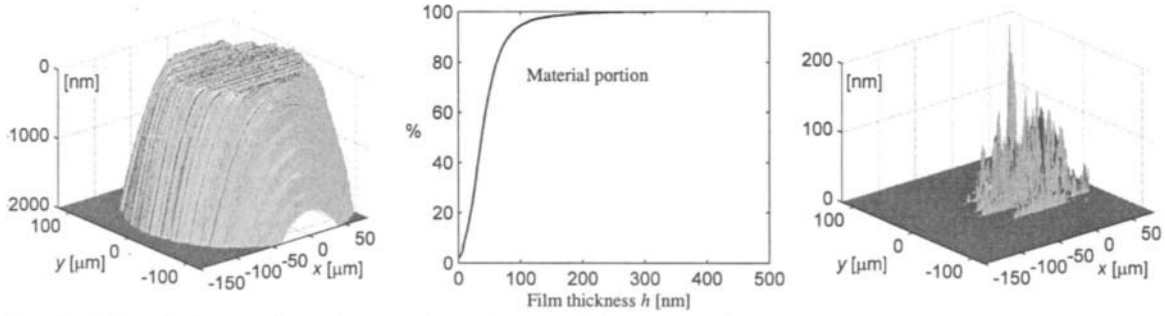


Fig. 21. Film thickness, slice plane graph and plastic deformations for longitudinal ground surface.

plastically deformed micro asperity of both surfaces. In case of the transversal ground surface, only a few asperity at the outlet of the EHL contact experience a small amount of plastic deformations. On the other hand, one can find severe plastic deformations in case of the longitudinal ground surface due to the ridges along the sliding direction.

4. FRICTION

This section describes the calculation of the friction force within the mixed lubrication regime. This is fluid friction due to shear stresses within the film and solid contact friction due to interaction between micro asperity.

4.1. Fluid friction

Fluid friction is calculated from the midplane shear stress τ_m of a Ree-Eyring fluid. This is already described in [1]. Shear stress τ_m for the line of symmetry of the Hertzian contact is depicted in fig. 22. The pressure p has the most influence on shear stress due to the exponential effect on viscosity. As a result of the introduction of the plastic pressure limit p_{plast} the effect on viscosity is

in fact limited. Therefore the maximum shear stress is also limited to a certain value which is more or less linearly dependent on the Eyring shear stress τ_0, if one neglects the influence of film thickness and pressure gradient at these spots.

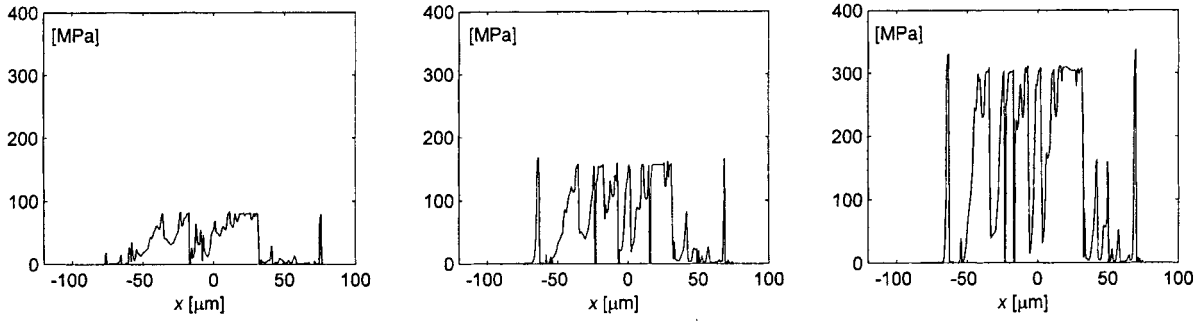


Fig. 22. Shear stress τ_m with $\tau_0 = 1.9$, 3.8 and 7.6 MPa, lapped surface, $u_2 = 0.1$ m/s.

4.2. Solid friction

The solid friction is calculated by an energetic approach which is described in [1] and in more depth by Bartel [10]. The components of the solid friction are deformation and adhesion. The first one can be calculated from the elastic-, plastic deformations of the macro and micro geometry. To calculate the adhesion force, some assumptions have to be taken into account at the moment: Adhesive bonding can only occur if plastic deformation of the material takes place and secondly if local temperatures exceed a critical value which lead to a desorption of the fluid film. Since the simulation program is not able to determine local temperatures yet, the second assumption cannot be tested. But it seems reasonable that the plastically deformed micro asperity also experience local temperature peaks.
Since it is quite sophisticated to get reliable values for the shear strength in adhesive bonding, in this work one sixth of the universal hardness or plastic pressure limit of 100Cr6 (SAE 52100), respectively, is used:

$$\tau_{ad} = \frac{HU}{6} = \frac{P_{plast}}{6} = 667 \frac{N}{mm^2} \qquad (5)$$

In this work the reference oil FVA Nr. 3 was studied. Again the value for the Eyring shear stress τ_0 is not known. Hence, the following estimation with the help of two experimental results was done. The friction curves of a ball on disc sliding wear test rig were evaluated (fig. 23).

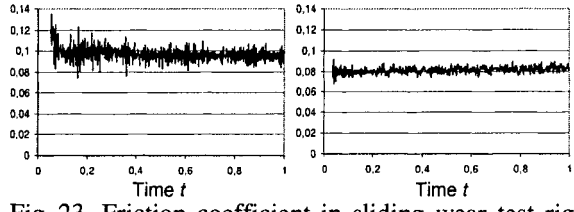


Fig. 23. Friction coefficient in sliding wear test rig for lapped and polished discs. Conditions of table 1. Ball and discs of 100Cr6 (SAE 52100).

One can see that the lapped disc starts with a friction coefficient of $\mu = 0.12$ whereas the polished surface starts at $\mu = 0.08$. These measurements were compared with simulation results for various Eyring shear stresses. It was found that an Eyring shear stress of about 6.8 MPa yields reasonable agreement between experiment and simulation.

4.3. Calculation of a Stribeck curve

The next step is to calculate Stribeck curves from boundary to hydrodynamic lubrication. With the help of the aforementioned assumptions as well as material- and fluid parameters, it is possible to calculate friction coefficients over sliding velocity. Fig. 24 shows the results for the lapped and fig. 25 for the polished surface topography.

Unfortunately, the available test rig is not capable to measure friction coefficient over the complete velocity range. Therefore it was not possible to verify all calculated points of the curve at this time. Nevertheless, one can see that in case of the lapped surface, the friction coefficient starts to increase after the hydrodynamic and mixed friction region is passed. However, in case of the polished surface, the friction coefficient is further decreasing. The reason is that no additional plastically deformed asperity will be added with decreasing velocity because of the smoothness of the polished surface.

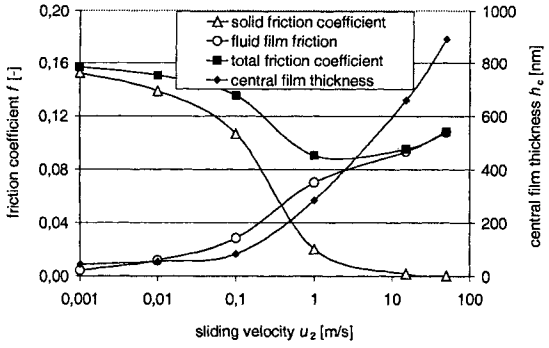


Fig. 24. Calculated Stribeck curve, ball against lapped disc (100Cr6/100Cr6).

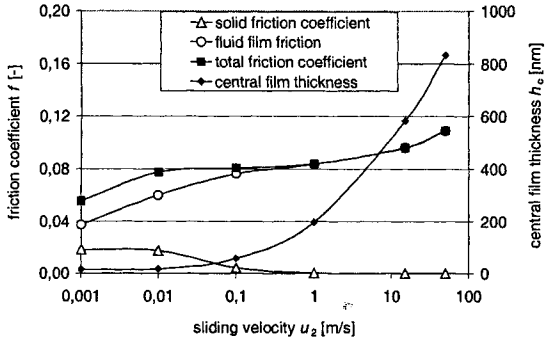


Fig. 25. Calculated Stribeck curve, ball against polished disc (100Cr6/100Cr6).

Since it is expected that the friction coefficient in the boundary regime is also above the hydrodynamic friction coefficient, it shows that the assumptions which are made concerning the friction calculation are not sufficient yet. It is supposed to be the temperature calculation which must be introduced to explain the increasing solid friction coefficient in that region. Probably the molecular thin film will be destroyed under these conditions and adhesive forces will increase.

The central film thickness is calculated according to Hu and Zhu [6]. This is 2/3 of the Hertzian diameter is taken into account for averaging.

5. CONCLUSION

The aim of the study is to investigate real rough surface topographies under pure sliding and mixed lubrication conditions. The new approach was introduced by means of an EHL problem with a single bump as surface asperity in the center of the Hertzian contact. This example offers detailed insight to some important mechanisms: On the one hand "fictitious solid contact spots" at asperity tips and on the other hand huge local pressure gradients upstream of the bumps which cause plastic deformations. The introduction of the variables p_{num} and h_{plast} enable the solution of problems with micro-asperity within the EHL contact.

With this algorithm three different real surface topographies were studied. Results for film thickness and pressure distribution were obtained. Further on new slice plane graphs were introduced to offer a quick and easy view into the contact. For example, one can see that the supply with lubricant on a lapped surface is quite different to a polished surface. The lapped surface shows a higher material portion than the polished surface but on the other hand in total the lapped surface has a bigger fluid volume within the contact.

Secondly, a comparison between a transversal and a longitudinal ground surface was made. It could be shown that the fluid volume of the transversal ground surface is bigger and the plastic deformations are less than in case of the longitudinal ground surface.

The last section describes the friction calculation with fluid and solid friction. The new model enables the calculation of Stribeck curves over several orders of magnitude. But it has proven that the assumptions concerning solid friction calculation are not sufficient yet. Since it could be shown that solid contact spots cannot be calculated by Reynolds equation but rather must be identified by a

temperature criterion. This must be done in future work.

REFERENCES

1. A. C. Redlich, D. Bartel, H. Schorr and L. Deters, Proceedings 26th Leeds-Lyon Symposium on Tribology, pp. 85-93, 2000
2. T. Nakahara, Japanese Journal of Tribology, Vol. 39, Nr. 3, pp. 335-349, 1994
3. L. Chang, Wear, Vol. 184, pp. 155-160, 1995
4. T. Makino, S. Morohoshi and K. Saki, Proc. of the 25th Leeds-Lyon Symposium, 1998
5. X. Jiang, D. Y. Hua, H. S. Cheng, X. Ai and S. C. Lee, ASME Journal of Tribology, Vol. 121, pp. 481-491, 1999
6. Y.-Z. Hu and D. Zhu, ASME Journal of Tribology, Vol. 122, pp. 1-9, 2000
7. X. Ai, H. S. Cheng and L. Zheng, ASME Journal of Tribology, Vol. 115, pp. 102-110, 1993
8. C. H. Venner, PhD Thesis, Twente University, 1991
9. A. Mihailidis, J. Retzepis, C. Salpistis and K. Panajiotidis, Wear, Vol. 232, pp. 213-220, 1999
10. D. Bartel, PhD Thesis, Aachen; Shaker-Verlag, 2001

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Some insights into lubricant film thickness in mixed EHD point contacts

I. Křupka, M. Hartl and M. Liška

Faculty of Mechanical Engineering, Brno University of Technology,
Technická 2, 616 69 Brno, The Czech Republic

This paper presents an experimental study of the effect of rolling speed and surface roughness on the mixed elastohydrodynamic (EHD) lubrication characteristics for point contact formed between a real, random, rough surface, steel ball and smooth glass disc. The Thin Film Colorimetic Interferometry measurement technique has been extended to give detailed information about in-contact deformation of the microgeometry. It has enabled to derive the amplitude reduction curve that shows progressive recovering of ball roughness features with increasing speed.

1 INTRODUCTION

An understanding of mixed lubrication, and in particular mixed lubrication of concentrated contacts sometimes referred as partial elastohydrodynamic (EHD) lubrication, is not only of fundamental interest but also practical importance to design of many machine elements. Fundamental to this phenomenon is the coexistence of a couple of different lubrication regimes within one contact region. The applied load is shared by both fluid and boundary lubricant films. In this lubrication regime there are the direct asperity contacts accompanied by altering local lubricant film thickness and pressure distribution (micro-EHD), asperity deformation and lubricant-surface tribochemical reactions. These processes may result in surface failures, such as excessive wear, scuffing and pitting. To avoid it, it is essential to study the effect of surface roughness on the mixed lubrication contact characteristics.

The analysis of partial EHD lubrication has attracted attention of many researchers during the last thirty years. The initial studies carried out e.g. by Johnson et al. [1] and Tallian [2] were analytical. With the advances in computer power and numerical techniques, full numerical solution of lubricated contact between rough non conformal surfaces became available. The first models were stochastic and gave only statistically global effect of surface roughness on EHD film thickness and pressure [3]. However, the real engineering surface topography is usually three-dimensional and complex. That is why deterministic solutions with optically measured three-dimensional rough surface data as an input have been recently introduced [4]-[7]. They ranged from a model of the full film EHD lubrication only [4], [5] to those covering entire range of lambda ratio from infinity down to nearly zero [6], [7].

As it can be seen from the above, considerable progress has been made in numerical simulation of mixed EHD lubrication of real rubbing surfaces. Conversely, only a few experimental studies on this subject can be found in the literature. The first attempt to obtain EHD film thickness map for random rough surface was made by Guangteng et al. [8] using advanced optical interference technique. Another form of this technique was recently used by Luo et al. [9] for the study of influence of operating conditions on the real contact area. Fluorescence microscopy [10] seems to be also promising tool for the study of mixed lubrication but it is still in nascent.

In this paper the Thin Film Colorimetric Interferometry (TFCI) technique [11] is used for the film thickness mapping in mixed EHD point contact formed between a real, random, rough surface, steel ball and a smooth glass disc. The main aims of the work are:

i.) to perform contact analysis of static dry contact to study the influence of load on surface roughness and contact area;

ii.) to explore the transition from the full film EHD lubrication to mixed lubrication;
iii.) to study the effect of rolling speed on the mixed EHD lubrication characteristics such as the average film thickness, lambda ratio and contact area ratio;
iv.) to measure in-situ surface roughness and determine its dependence on rolling speed.

2 EXPERIMENTAL METHOD

The experimental apparatus and measurement technique have been described in detail before [11], [12]. However, a brief account of the experimental method will be given here since the mixed lubrication study required some modification of the measurement procedure.

2.1 Experimental apparatus

Figure 1 shows a schematic diagram of the experimental apparatus. This is a high-pressure ball on disk tribometer equipped with a microscope imagining system and a control unit.

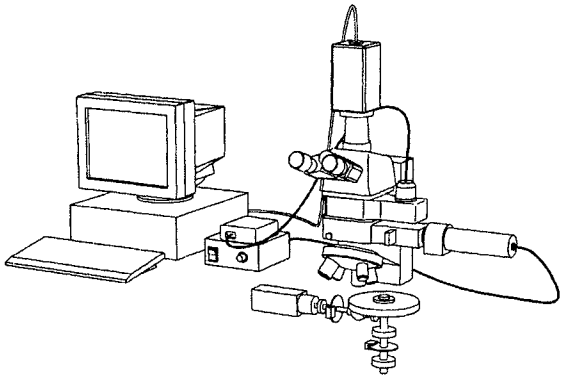


Figure 1. Experimental apparatus.

The tribometer consists of a steel ball loaded and rotated against a transparent glass disc. The contacting side of the disc is coated with a thin semi-reflective chromium layer that is overlaid by a silicon dioxide "spacer layer". The glass disc is driven, in nominally pure rolling, by the ball that is driven by a servomotor via a planet gearbox.

Chromatic interference patterns produced by the contact area are collected by microscope that incorporates a xenon flash lamp and a high-resolution digital camera. Using the xenon flash lamp with 2.9 μs FWHM pulse width, it is possible to freeze surface irregularities movement. During flash duration, observed asperities travel distance not longer than 0.3 μm. Three 1360 × 1024 pixel CCD camera together with 10× objective gives spatial resolution of 0.47 μm.

Flash discharge is triggered by a signal from the rotary encoder attached to the ball shaft, so that all measurements are carried out at the same ball position. Another encoder is fastened to the shaft connected to the glass disc. It ensures that all images are captured at the position where the spacer layer thickness was measured beforehand.

2.2 Measurement technique

The experiments reported here make use of the Thin Film Colorimetric Interferometry technique for film thickness measurement. In this technique, 3 × 8 bit SXGA still pictures are after transformation from RGB to CIELAB colour space converted to the film thickness map using appropriate colour/film thickness calibration curves. They are obtained from Newton rings for static contact formed between the steel ball and the glass disc coated with the chromium layer only. During several last years the accuracy and resolution of TFCI has been improved and it is now able to measure film thickness to within 1 nm [13].

Film thickness evaluation was performed by the program AChILES. For the mixed lubrication study, the program capabilities were extended with few plug-ins. They enable to calculate from the selected part of the contact area Ω the following three-dimensional parameters:

i.) Average film thickness defined as

$$h_{av} = \frac{1}{N} \sum_{x,y \in \Omega} \sum h(x,y) \qquad (1)$$

where N is the total number of points belonging to the area Ω and $h(x,y)$ is measured film thickness at location x, y.

ii.) Root-mean-square deviation of the rubbing surface

$$S_q = \left[\frac{1}{N}\sum_{x,y\in\Omega}\sum\left(h(x,y)-h_{av}\right)^2\right]^{1/2} \quad (2)$$

iii.) Real contact area Ω_c that is determined by summing up all elemental areas where measured film thickness is equal or lower than the spacer layer thickness.

In this paper, the area Ω is a circle having its centre identical with the centre of Hertzian contact zone and radius equals of 2/3 of the Hertzian contact radius.

2.3 Test materials and conditions

Commercially provided 25.4 mm diameter steel balls of two different roughness values were employed in this study. The balls have no preferential surface roughness pattern. The glass disc has 150 mm in diameter and its surface is optically smooth. The silicon dioxide spacer layer about 195 nm thick covers almost entire underside of the glass disc with the exception of a narrow radial strip where only chromium layer is deposited. Scanning probe microscopy examinations of silicon dioxide films shows that they are extremely smooth [14]. Material properties of the contact bodies are summarised in Table 1. The surface roughness parameters of steel balls measured by stylus technique are given in Table 2.

Table 1
Material properties of contact bodies

	Elastic modulus (GPa)	Poisson's ratio
Ball	212	0.3
Disc	81	0.208
Spacer layer	70 [15]	0.17 [15]

Table 2
Ball surface roughness parameters

Ball name	Centerline average (CLA) roughness (nm)	Root mean square (rms) roughness (nm)	Maximum peak-to-valley height (nm)
Smooth	4	6	35
Rough	12	15	93

All experiments were carried out with naphtenic additive-free base oil SUN 500 N. Its properties at a temperature of 25°C are listed in Table 3.

Table 3
Properties of the test oil at a temperature of 25°C

Atmospheric viscosity (Pas)	Pressure-viscosity coefficient (GPa^{-1})	Refractive index
0.321	31	1.5024

The test conditions used in this study are summarised in Table 4.

Table 4
Test conditions

Contact load (N)	29
Hertz pressure (GPa)	0.52
Contact radius (μm)	165
Rolling speeds (ms^{-1})	0.0001 – 0.05
Temperature (°C)	30

3 RESULTS AND DISCUSSION

3.1 Static contact

An asperity contact is the major feature of mixed lubrication. That is why it is essential to start experimental investigation with the contact analysis of a static contact.

The static contact can be considered as a contact between an isotropically randomly rough ball and elastic half-space with a smooth surface layer. Because the elastic moduli of silicon dioxide layer and glass disc are close, the presence of a surface layer has very weak effect on the contact behaviour [16]. When the ball is subjected

to a normal load, its surface peaks make contact with the layer at discrete contact spots. As the load is increased, the peaks are crushed into the layer and their shapes are changed. The sum of the areas of all the contact spots constitutes the real area of contact that is a small fraction of the nominal contact area. The nature of contact deformation can be determined using Mihic's plasticity index [17]

$$\psi = \frac{E'}{H}\Delta_q \tag{3}$$

Taking the effective modulus $E' = 173$ GPa, the steel ball hardness $H = 7.7$ GPa and the root mean square slope of a profile $\Delta_q = 0.006$, it follows that $\psi = 0.14$ and the contact is expected to be predominantly elastic.

The capability of TFCI technique to measure nanometric separations has been demonstrated earlier [13]. In order to make sure that this technique can provide credible estimation of the real contact area, the contact topography was also observed by darkfield microscopy. Figure 2 shows comparison of the results given by both techniques for 30 × 70 μm part of lightly loaded contact between the rough ball and the glass disc. The dark areas represent

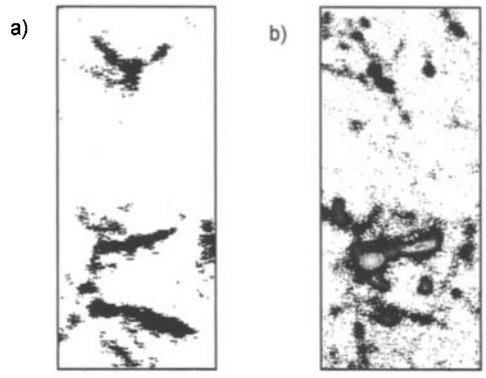


Figure 2. Comparison of the contact topography visualisation obtained using (a) TFCI and (b) darkfield techniques.

the valleys on the ball surface i.e. the regions that reflect oblique light back into the objective under darkfield illumination. As it can be seen there is satisfy agreement between both techniques.

Figures 3 (a) – 3 (c) show the TFCI measured results of the real contact area under various applied loads. Figures 3 (d) – 3 (f) depict the corresponding three-dimensional surface roughness taken from the central part of contact. It can be seen that with

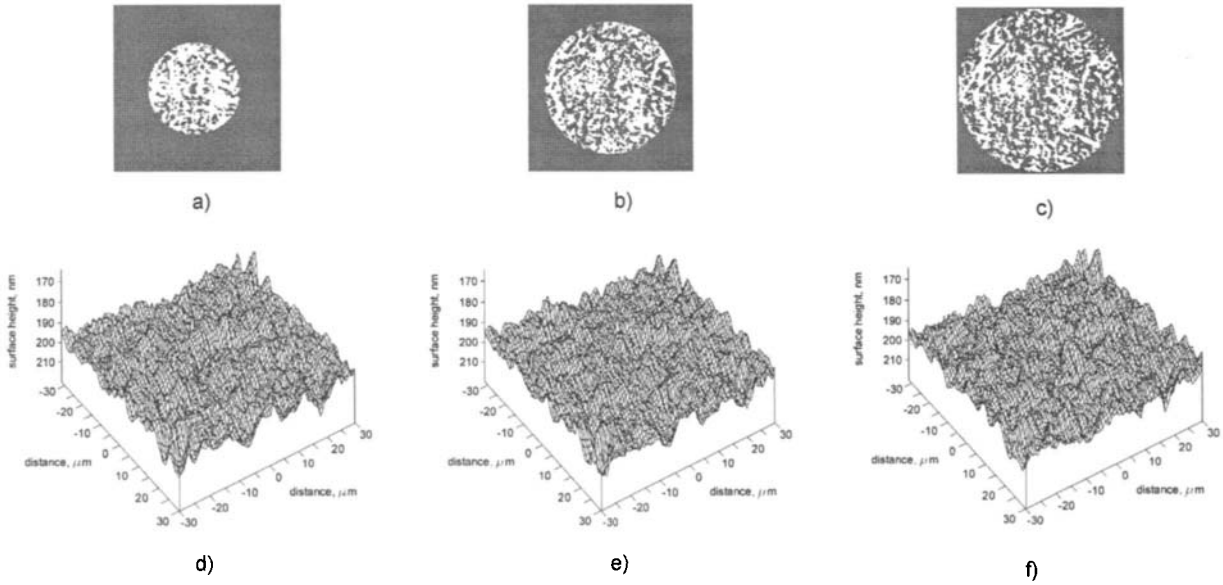


Figure 3 Contact between the rough ball and SiO_2 layered glass disc. Real contact area for the load of 6.5 N (a), 19N (b) and 36.5 N (c) and corresponding 3D surface roughness in the central part of contact (d - f).

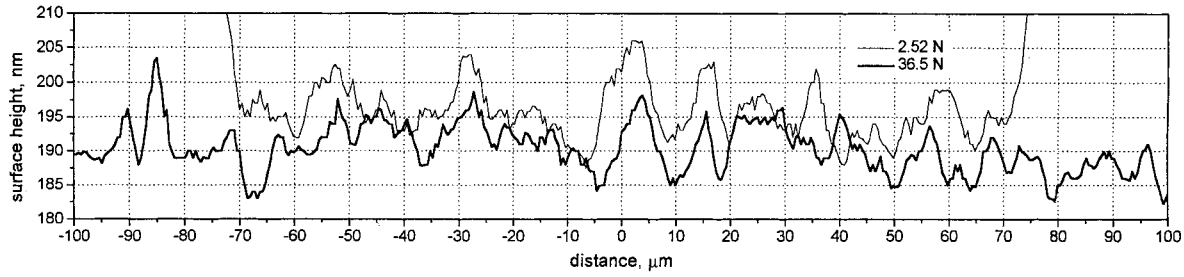


Figure 4 Surface profiles at a load of 2.52 and 36.5 N.

the increase of the applied load, the deformation occurs and asperities are partially flattened. Deformations were found to be recoverable. Between contacting surfaces there is some intercontact space volume filled with air that can potentially influence the measurement accuracy. To prevent it, a low viscosity oil with the same refractive index as the silicon dioxide layer was introduced between surfaces before measurement. Further evidence of surface roughness conformity effect can be seen in Fig. 4, which shows surface profiles at two different loads. The displacement of reference line that is about 2 nm is caused by the compression of spacer layer. If the surface roughness measurements are performed for several different load values, relationship between root-mean-square deviation of the ball surface and applied load can be generated. This is on Fig. 5 for 60 μm diameter circle area located at the centre of the contact.

The relationship between the real contact area and applied load is shown in Fig. 6. For a comparison, the ideal smooth case is also plotted. It can be seen that the rough surface contact area is linearly proportional to the applied load and represents only a small percentage of the smooth contact result. It is consistent with Greenwood and Trip [18] extension of the Hertzian theory for the case of the elastic contact of rough spheres. Figure 7 shows the variation of real contact area with applied load for four different spatial resolutions. For each spatial resolution, a nearly linear behaviour is observed. The slope of the lines decreases with increasing spatial resolution. Therefore at the constant load the contact pressure increases and real contact area decreases with spatial resolution. This is in accordance with previously published works [19], [20].

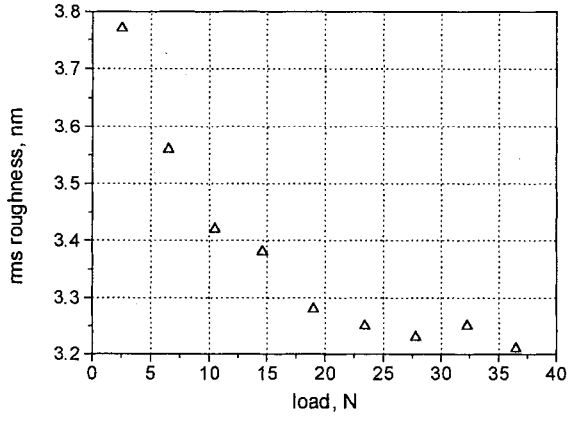


Figure 5 Variation of rms deviation of the ball surface with applied load.

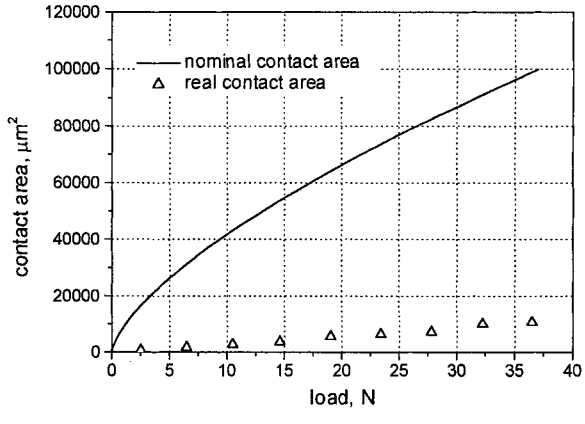


Figure 6 Variation of contact area with load.

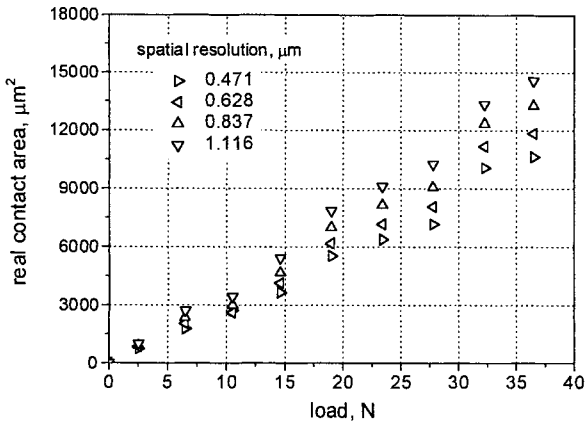


Figure 7 Variation of real contact area with load for different spatial resolutions.

3.2 Running contact

Besides the average film thickness and rms deviation of the rubbing surface that were introduced in Section 2.2, three another common terms are used in this study to describe mixed EHD lubrication. They were originally defined in the papers devoted to the numerical analysis e.g. [21] and include nominal lambda ratio, lambda ratio and contact area ratio. Nominal lambda ratio is defined as the ratio of average film thickness for the smooth surface to the composite rms roughness for the rough ball. Lambda ratio is defined as the ratio of average film thickness for rough surface to the composite rms roughness for rough ball. Contact area ratio is a fraction real contact area over the total area of the Hertzian contact zone.

Figure 8 shows the evolution of a lubricant film shape for the rough ball at the transition from the full EHD to mixed lubrication. The lambda ratio ranges from 1.03 to 0.49. The contour line intervals are 15 nm. It is observed that the ball topography and particularly asperity-scale geometry is significantly changed as the rolling speed decreases.

Figure 9 shows the comparison between film thickness profiles taken along the contact for the smooth and rough ball. Smooth ball measurements demonstrate typical full film EHD characteristics, as it can be seen in Fig. 9(a). Figure 9(b) illustrates the process of the surface roughness attenuation with decreasing rolling speed. Comparing the film thickness profiles shows that for a smaller speed the elastic deformation of the microgeometry is noticeably higher. Similar effect has been described in numerical studies e.g. [22] and experimental studies of the artificial surface features e.g. [23].

In Fig. 10 the ratio of the average film thickness for the rough ball to that for the smooth ball is plotted against rolling speed. The results show that the effect of surface roughness on the average film thickness do not become significant until rolling speed becomes below approximately 0.001 - 0.002 m/s. It corresponds to the nominal lambda ratio value around 0.8.

Figure 11 shows the contact area ratio and both lambda ratio and nominal lambda ratio as a function of rolling speed. Because the value of contact area ratio is quite sensitive to the spacer layer thickness that is measured with accuracy better than 1 nm, error bars are also plotted to indicate the estimated error in a measurement. It is important to see that the curves in Fig. 11 have the same trends as those

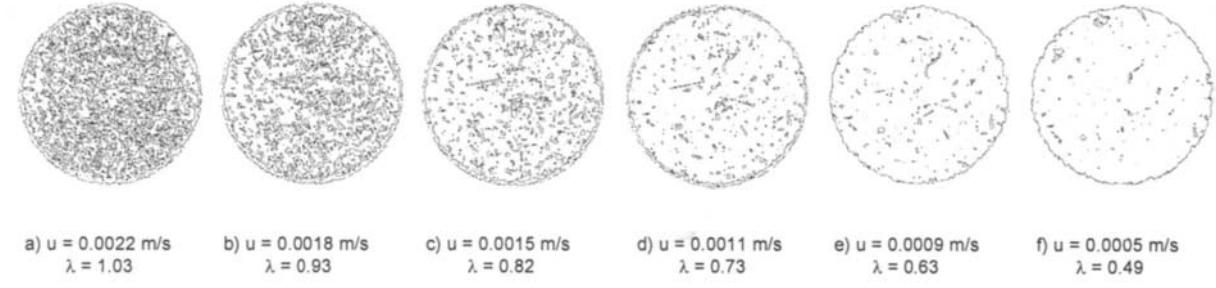


Figure 8 Film thickness contour maps for rough ball.

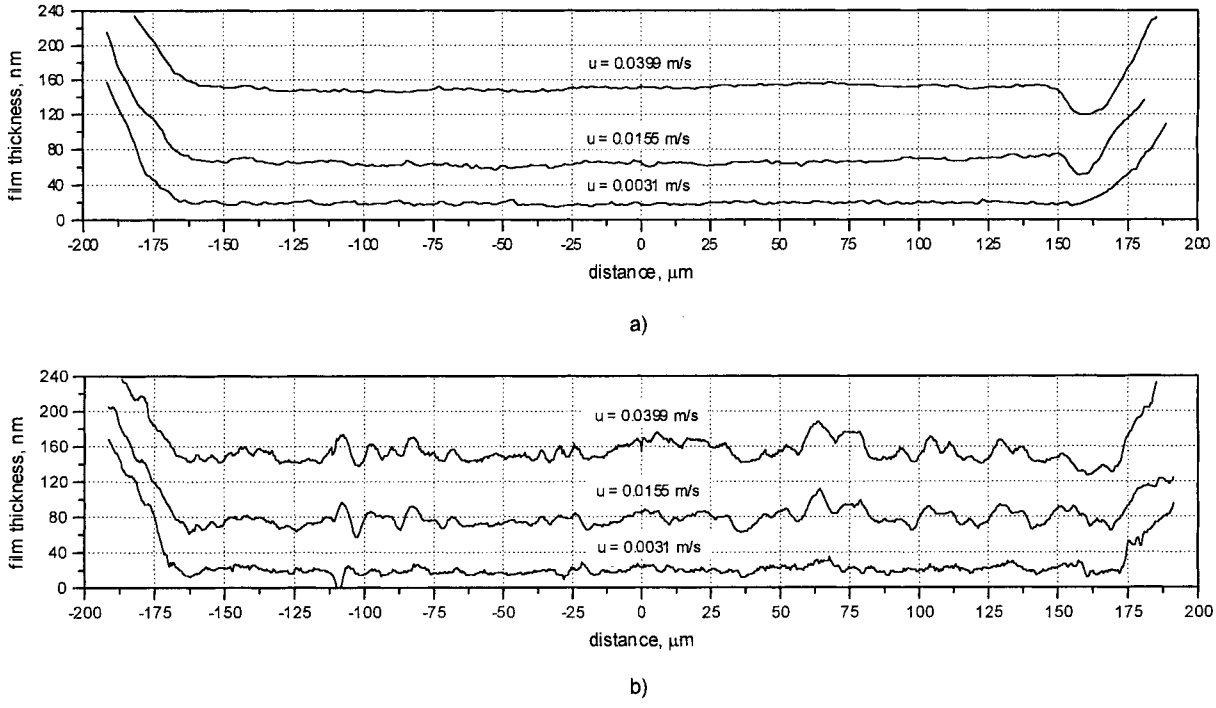


Figure 9 Film thickness profiles along the contact for (a) smooth ball and (b) rough ball.

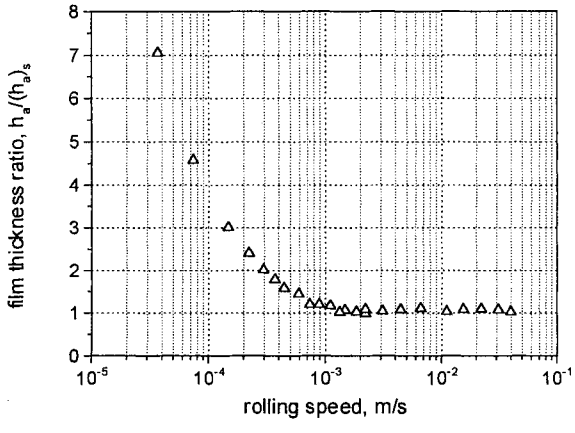


Figure 10 Ratio of the average film thickness for rough ball to that for smooth ball as a function of the rolling speed.

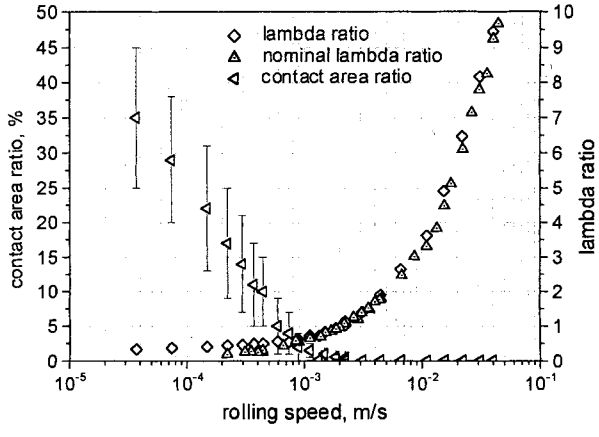


Figure 11 Variation of contact area ratio and lambda ratio with rolling speed for rough ball.

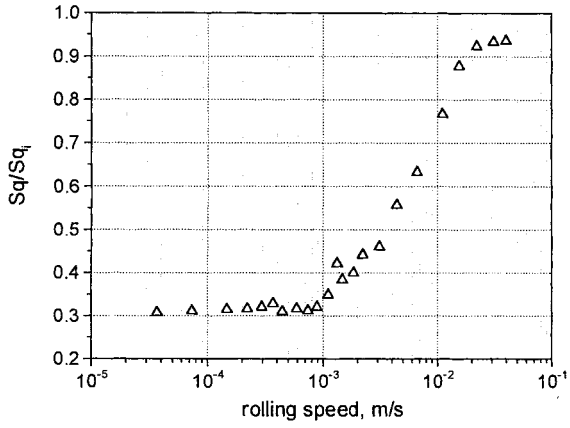


Figure 12 Variation of amplitude reduction with rolling speed for rough ball.

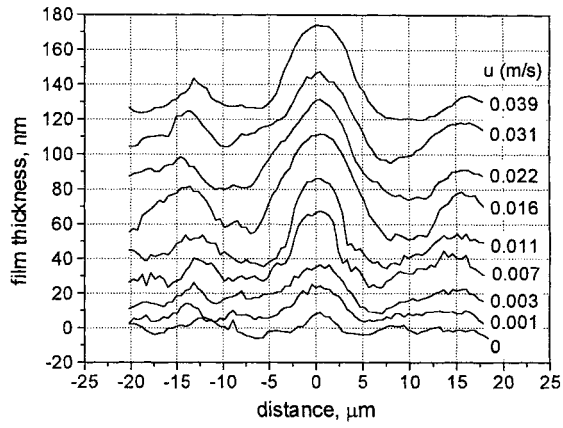


Figure 13 Rough ball film thickness profiles along the contact showing in detail the deformation of asperities.

obtained from deterministic numerical solutions with measured surface roughness [6], [24]. The contact area ratio decreases as the rolling speed increases, from 35 % at very low speed down to almost zero at higher speeds. Because of the low value of load parameter used in this test, the real contact area is lower than hydrodynamic area for the whole speed range. The reduction of the rolling speed decreased both lambda ratios that are identical down to 0.001 m/s. Below this speed, the lambda ratio is greater than nominal lambda ratio because the roughness makes the film thicker.

To determine the extent of surface roughness attenuation in EHD contact, Venner et al. [25] introduced a parameter called amplitude reduction. It is defined as the ratio between deformed and undeformed roughness amplitudes. Figure 12 shows the ratio of the rms roughness measured in contact S_q as a function of rolling speed to that for undeformed surface S_{qi}. For the low speeds, the roughness features on the ball surface are deformed to approximately 30 % of their original height and S_q is close to the value measured in the static contact. As the rolling speed increases the roughness features recover their original form and S_q approaches S_{qi}.

Figure 13 gives detail insight into lubricant film thickness behaviour in the vicinity of real asperity. As it has been already described, with decreasing rolling speed the asperity is flattened. It is illustrated in Fig. 14 that shows the variation of asperity height with rolling speed. Comparing the film thickness profiles in Fig. 13 foreshadows the existence of a wedge-like film shape at the asperity tip. This effect should be attributed to the micro-EHD lubrication. Similar behaviour has been previously observed by Kaneta and Cameron [26] that studied three dimensional model asperities. They suggested to explain it by a micro squeeze effect.

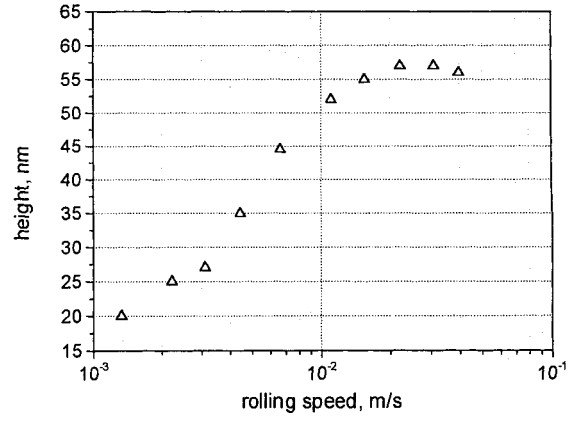


Figure 14 Variation of rough ball asperity height with rolling speed.

4 CONCLUSIONS

The following main conclusions can be drawn from this work:

i.) The TFCI measurement technique has been extended to enable the study of real surface roughness behaviour in mixed EHD point contact.

ii.) When the technique has been applied to static dry contact, the linear relationship between the real contact area and applied load has been observed. The real contact area has been also found to decrease with spatial resolution.

iii.) In the case of mixed EHD contact, the effect of rolling speed on the mixed EHD characteristics such as the average film thickness, lambda ratio and contact area ratio has been investigated.

iv.) The real surface amplitude reduction curve has been derived from measured in-contact deformation of the microgeometry. It shows progressive recovering of ball roughness features with increasing speed.

ACKNOWLEDGEMENTS

This research was supported by the Grant Agency of the Czech Republic and the Ministry of Education, Youth and Sports of the Czech Republic under grants Nos.101/00/0155 and J22/98:262100002. The test oil was supplied by Sun Oil Company Belgium. CeramTec CZ should be acknowledged for stylus surface roughness measurements.

REFERENCES

1. K. L. Johnson, J. A. Greenwood and S. Y. Poon, Wear, 19 (1972) 91.
2. T. E. Tallian, Wear, 21 (1972) 49.
3. N. Patir and H. S. Cheng, ASME Jour. of Lubr. Tech., 100 (1978) 12.
4. G. Xu and F. Sadeghi, ASME Jour. of Trib., 118 (1996) 473.
5. D. Zhu and X. Ai, ASME Jour. of Trib., 119 (1997) 375.
6. D. Zhu and Y. Z. Hu, Trib. Trans., 44 (2001) 383.
7. D. Zhu and Y. Z. Hu, Trib. Trans., 44 (2001) 391.
8. G. Guangteng, P. M. Cann, A. V. Olver and H. A. Spikes, ASME Jour. of Trib., 122 (2000) 65.
9. J. Luo, S. Liu and S. Wen, Sci. China Ser. A, 44 (2001) 78.
10. J. Sugimura, M. Hashimoto and Y. Yamamoto, Proc. of the 26th Leeds-Lyon Symp. on Tribology, (2001) 609.
11. M. Hartl, I. Křupka, R. Poliščuk, M. Liška, J. Molimard, P. Vergne and M. Querry, Trib. Trans., 44 (2001) 270.
12. M. Hartl, I. Křupka, R. Poliščuk and M. Liška, Trib. Trans., 42 (1999) 303.
13. M. Hartl, I. Křupka and M. Liška, Proc. of the 28th Leeds-Lyon Symp. on Tribology, (2001) in press.
14. IEEE Micro Electro Mechanical Systems Workshop, San Diego 1996, 97.
15. M. T. Kim, Thin Solid Films, 283 (1996) 12.
16. S. J. Cole and R. S. Sayles, ASME Jour. of Trib., 114 (1992) 334.
17. B. B. Mihic, Int. J. Heat Mass Transfer, 17 (1974) 205.
18. J. A. Greenwood and J. H. Tripp, ASME Jour. Apple. Mech., 89 (1967) 153.
19. C. Y. Poon and B. Bushan, Wear 202 (1996) 83.
20. B. N. J. Persson, F. Bucher and B. Chiaia, Phys. Rev. B 65 (2002) 65.
21. Y. Z. Hu and D. Zhu, ASME Jour. of Trib., 122 (2000) 1.
22. C. H. Venner and W. E. ten Napel, ASME Jour. of Trib., 114 (1992) 616.
23. M. Kaneta, T. Sakai and H. Nishikawa, ASME Jour. of Trib., 114 (1992) 779.
24. X. Jiang, D. Y. Hua, H. S. Cheng, X. Ai and S. C. Lee, ASME Jour. of Trib., 121 (1999) 481.
25. C. H. Venner, F. Couhier, A. A. Lubrecht and J. A. Greenwood, Proc. of the 23rd Leeds-Lyon Symp. on Tribology, (1996) 103.
26. M. Kaneta and A. Cameron, ASME Jour. of Trib., 102 (1980) 374.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

ON THE EHL FILM BREAKDOWN

Dong Zhu
Eaton Corporation, 26201 Northwestern Highway, Southfield, Michigan 48037, USA, e-mail: DongZhu@Eaton.com

SUMMARY

This paper, based on a recently developed deterministic approach for mixed lubrication, discusses the EHL film breakdown caused by three main reasons: speed (or viscosity) reduction, load increase and surface roughening. It has been observed that as the speed/viscosity is continuously decreased, the reduction of film thickness can be gradually accelerated, and lubrication breakdown may be significantly faster than that predicted by the conventional EHL theory. At a constant load there is apparently a critical speed, below which the EHL film cannot be established and severe surface contact may take place. This critical speed appears to be a function of load and contact ellipticity. The heavier the load is, for example, the higher the critical speed would be. When the speed is kept constant and the load is continuously increased, on the other hand, a similar phenomenon is observed. The EHL film may quickly drop with increasing load. After the load goes beyond a certain limit, the EHL film crashes due to insufficient hydrodynamic action that can no longer support the increased load. This critical load is evidently a function of speed. Commonly used EHL formulae may overestimate the film thickness, especially under severe conditions. The EHL film breakdown can also be observed when the roughness increases, but this appears to be a gradual process due to the fact that the overall average film thickness in the contact may not change significantly with increased roughness. A typical curve for the load sharing between the hydrodynamic film and the surface asperity contact vs. roughness is presented in this paper.

Keywords: Mixed Lubrication, Elastohydrodynamic Lubrication (EHL), Lubrication Breakdown, Lubrication Failure

1 INTRODUCTION

Elastohydrodynamic Lubrication (EHL) has been given great attention since 40 some years ago, because it is important to many engineering applications, such as bearings, gears, cams and followers, and traction drives. It is always preferred to maintain an EHL film in the contact, where the power and motion are transmitted and various types of surface failure are originated. The EHL film breakdown, on the other hand, often becomes the major cause of low operating efficiency, high frictional loss/energy consumption and drastic temperature rise, as well as early failures, such as pitting due to contact fatigue, scuffing and excessive wear. Therefore, a good understanding of the transition from the full-film EHL to the lubrication breakdown is essential for design optimization, failure analysis, problems solving and efficiency/performance improvement.

Due to the strong non-linearity and instability of the EHL problems, numerical simulation under severe operating conditions has been difficult until recently. Early studies by Dowson and Higginson [1], and Hamrock and Dowson [2] and others (e.g. [3-5]) have laid a solid foundation for the conventional EHL theory that has been widely used in tribology research and industrial applications. Since the computational power was very limited and the numerical algorithms were not well developed in the past, these early studies were unable to cover wide parameter ranges, especially cases under severe conditions. In the last 10-15 years computer technologies have been advancing greatly, so that sophisticated algorithms and large-scale computational analyses can now be executed on personal computers. This enables researchers to overcome the difficulties associated with the non-linearity and instability. Full-scale deterministic simulations under very severe conditions towards the lubrication breakdown now become possible. Early deterministic models include those by Lubrecht et al [6], Kweh et al [7], Ai and Cheng et al [8], Venner et al [9], Xu and Sadeghi [10] and Zhu and Ai [11]. It seems that, however, all the cases published in these studies were still in the full-film EHL regime and no significant surface contact/lubrication breakdown was presented.

Most recently Jiang et al [12] published a deterministic EHL solution that handled hydrodynamic lubrication and asperity contacts (up to 12% of Hertzian contact area in presented cases) simultaneously. Shortly after that Zhu and Hu [13,14] presented a unified approach that is capable of simulating the entire transition from the full-film, mixed, down to boundary lubrication. The contact load and area ratios could vary from zero up to 100%, while the film thickness ratio, λ, could be reduced from infinity down to zero. The numerical model has been proven to be robust under severe operating conditions (see [13-16]). Based on this model, a systematic study has been conducted by the present author for the effects of load and speed on the EHL film thickness in extended parameter ranges [17].

This paper focuses on the discussion of EHL film breakdown, caused by severe conditions such as low speed or viscosity, heavy load, and large surface roughness. The objective is to better understand the transition towards lubrication failure, and hopefully this will benefit the development of prediction tools.

2 CASES ANALYZED AND PARAMETER RANGES COVERED

The present study is based on the deterministic model and the full numerical solution method described in detail in [13-15], which will not be repeated here.

It is important to note that commonly used EHL film thickness formulae were originally from curve fitting based on limited numbers of analyzed cases in relatively narrow parameter ranges due to the lack of computing power in the past. Hamrock and Dowson's formulae [2], for example, were based on 34 cases in the following speed and load parameter ranges:

Speed Parameter U^*: $0.08416\text{x}10^{-11} \sim 5.050\text{x}10^{-11}$
Load Parameter W^*: $0.1106\text{x}10^{-6} \sim 1.290\text{x}10^{-6}$
(Maximum Hertzian Pressure P_H: 0.183~0.415 GPa)

However, actual operating conditions in typical engineering applications, such as gears, bearings, cams and traction drives, may vary in much wider parameter ranges. The maximum Hertzian contact pressure, P_H, for example, is often around 1.5~3 GPa in gears and rolling element bearings, and it could possibly exceed 5~6 GPa in traction drives. In the above-mentioned recent study by the present author [17], the speed and load parameter ranges have been extended to:

Speed Parameter U^*: $2.298\text{x}10^{-17} \sim 3.448\text{x}10^{-8}$
Load Parameter W^*: $6.299\text{x}10^{-9} \sim 7.880\text{x}10^{-3}$
(Maximum Hertzian Pressure P_H: 0.146~6.851 GPa)

These ranges cover more than nine orders of magnitude for the speed and more than six orders for the load. Note that the lower speed limit has been greatly reduced down to practically zero, while the upper load limit increased so that the maximum Hertzian pressure can be as high as 6.851 GPa. This parameter range extension appears to be necessary for the study of lubrication breakdown.

In order to focus on the speed and load effects, some other parameters are fixed for all the cases analyzed. These include the material properties, E'=219.78 GPa, η_o=0.096 Pa.s, and α =18.2 GPa $^{-1}$, and the geometric parameter R_x=19.005 mm. The other geometric parameter, R_y, will be either 19.005 mm for a circular contact (k=1.0), or 504.80 mm for a long elliptical contact of k=8.0 with its minor axis coincident with the rolling direction. This elliptical contact can be considered as a better representation of so-called "line contact" (that can be commonly found in gears, roller bearings and many other components). This is because the ideal 2-D line contact does not allow any direct surface contact in the mixed lubrication study with deterministic models, and it does not exist in reality either. In addition, the presentation and discussion of the results in this paper will be focused only on the central (or average) film thickness, as those of the minimum film thickness have similar trends. Also, the minimum film thickness does no longer seem to be a representative/critical parameter whenever surface contact may possibly take place.

3 LUBRICATION BREAKDOWN DUE TO LOW SPEED/VISCOSITY

A large number of cases have been analyzed and full numerical solutions obtained. The results for the central film thickness vs. rolling speed are summarized in Figure 1 for smooth contact cases at the contact ellipticity of k=*1* (circular contacts) and in Figure 2 for those at k=*8*. All the results are compared with those from Hamrock and Dowson's formula illustrated by straight lines labelled with "H&D" in the figures. Note that in an isothermal lubrication analysis the effect of viscosity is very similar to that of speed. The analysis and discussion in this section, therefore, will be focused on the speed effect, but the related results and conclusions would also be useful for understanding the viscosity effect, which is important in engineering applications where temperature increase may cause a significant viscosity reduction.

It is observed that, at a constant load, when the speed is moderate and in a relatively narrow range, the present numerical solutions show good agreement with the conventional theory. The relationship between the film thickness and the speed can be approximated by the straight lines on a log-log scale, as shown in Figures 1 and 2. However, after the speed range is greatly extended, the relationship is found to be quite different from the simple power-law function described by the conventional theory, especially when the load is heavy. The slope of the film thickness curve is no longer constant. It is gradually reduced as the speed rises, but, more importantly, rather quickly increased if the speed decreases, showing an accelerated reduction of the EHL film thickness.

If the speed is further reduced down below a certain limit, the central film thickness drops down to zero, demonstrating the lubrication breakdown with a severe surface contact due to insufficient hydrodynamic action. Since zero cannot be illustrated on the log scale, a very small value corresponding to 0.47 nm has been used in Figure 1, and that of 0.58 nm in Figure 2, as a practical representation of zero film. Since no hydrodynamic film can be so thin in reality, it is reasonable to use such small values as "practically zero" when presenting the results.

It can be seen from Figures 1 and 2 that the present results better agree with the conventional theory if the load is light. For example, the results for P_H=0.397 GPa agree quite well with Hamrock and Dowson's formula, which was developed through curve-fitting based on 15 cases for the speed effect at P_H=0.3444 GPa (see [2]). When the load is heavy, however, there may be a significant discrepancy, and the conventional theory seems to often overestimate the film thickness, especially when the speed is low. Note that the critical speed below which the lubrication breaks down is considerably higher when the load is heavier. Also, both the circular and the long elliptical contact cases show the same trends, although there are some noticeable quantitative differences. The long elliptical contact appears to have a significantly higher critical speed on the same level of Hertzian pressure, demonstrating a better load-bearing capacity.

Figure 3 presents a set of sample cases at k=8, showing some detailed characteristics for the speed effect, especially the transition towards lubrication breakdown

due to a continuous speed reduction. Note that in the figure the solution domain has been normalized.

4 LUBRICATION BREAKDOWN DUE TO HEAVY LOAD

After the load parameter range is greatly extended, the relationship between the film thickness and the load is found to be quite different from the simple power-law function. Typical curves of film thickness vs. load based on a set of numerically analyzed cases for k=8.0 are given in Figure 4, in comparison with the results from Hamrock and Dowson's formula illustrated by straight lines. It can be seen that, at a constant speed, when the load is light and within a relatively narrow range, the results from the present work agree quite well with those from the conventional EHL theory. The film thickness decreases slowly as the load increases. However, if the speed is not high and the load becomes heavy, the film thickness may quickly fall with the increasing load, considerably faster than expected by the conventional theory. When the load is further increased beyond a certain limit, the EHL film could eventually crash down to zero and severe surface contact may take place. This is because that at the constant speed the hydrodynamic influence is limited. If the load continuously increases, beyond a certain point the hydrodynamic action will no longer be able to support the increased load.

It is evident that the critical load beyond which lubrication breaks down is a function of speed. When the speed is very high, e.g. U=15 m/s, it is difficult to observe the EHL failure because the hydrodynamic action is strong. However, the film would eventually break if the load continued to go infinitely heavy. When the speed is low, on the other hand, the critical load becomes light as a result of weak hydrodynamic influence, and the lubricant film can easily drop down to zero. The higher the speed, therefore, the heavier the critical load would be. Note that a very small value of h_c/R_x =0.00000015 has been used as the "practically zero" in Figure 4.

Figure 5 gives some sample results that may be helpful for a better understanding of the transition towards lubrication breakdown due to load increase.

5 LUBRICATION BREAKDOWN DUE TO ROUGHNESS

In a recent paper by the present author [16] the effect of roughness on the EHL film thickness is studied in a wide RMS roughness range, from 0 and 0.1 μm up to 1.5 μm, with different types of machined surface. The results of the study indicate that, as roughness increases, the lubrication mode could change from the full-film to the mixed and even to the boundary lubrication with the contact load ratio greater than 60-80%, while the average film thickness may be kept relatively constant. Figure 4 summarizes the results from a set of sample cases with a shaved surface running through the EHL interaction zone, and Figure 5 shows some details for four typical cases. It can be seen that when the roughness is small, 0.1 μm, the film thickness ratio, λ, can be as large as 2.5 and there is almost no contact. This is in the full-film lubrication. When the roughness is increased to 0.2-0.4 μm, significant asperity contacts can be observed but there is still a large percentage of the load supported by the hydrodynamic film, showing typical characteristics of mixed lubrication. If the roughness is further increased to 1.0-1.5 μm, the asperity contacts become severe and the λ ratio can be reduced down to about 0.2. This can be considered as lubrication breakdown, because the majority of the load is now supported by the surface contact. Note that the transition from the full-film to the boundary lubrication due to increasing roughness appears to be a gradual process, because the overall average film thickness in the contact may not change greatly.

6 CONCLUSIONS

1) It has been observed that as the speed/viscosity is continuously decreased the film thickness reduction could be gradually accelerated, and the lubrication breakdown can be significantly faster than predicted by the conventional EHL theory.
2) At a constant load there is a critical speed, below which the EHL film cannot be established and severe surface contact may take place. This critical speed appears to be a function of load and contact ellipticity. The heavier the load, the higher the critical speed.
3) When the speed is kept constant and the load is continuously increased, the EHL film may quickly drop, and the commonly used EHL formulae may overestimate the film thickness especially for the cases under severe conditions. After the load goes beyond a certain limit, the EHL film crashes due to insufficient hydrodynamic action. This critical load is evidently a function of speed.
4) The EHL film breakdown can also be observed when the roughness increases, but this appears to be a gradual process due to the fact that the overall average film thickness in the contact may not change greatly.

ACKNOWLEDGMENTS

The author would like to express his sincere gratitude to the continuous support by Mr. Bo Lisowsky and Mr. John Bair of Eaton Corporation. He would also like to thank Professor Yuanzhong Hu of Tsinghua University, China, for his valuable discussion during the present study.

NOMENCLATURE

a, b = semi-axes of Hertzian contact ellipse in the x- and the y-directions, respectively
E' = effective Young's modulus
G^* = $\alpha E'$, dimensionless material parameter
g_E = $W^{*8/3}/U^{*2}$, dimensionless elasticity parameter
g_V = $G^* W^{*3}/U^{*2}$, dimensionless viscosity parameter
h = local film thickness (or gap)
h_a = average film thickness in the contact zone
h_c = central film thickness for smooth contact

H_c = h_c/R_x, dimensionless central film thickness
k = b/a, contact ellipticity
p = pressure
P_H = maximum Hertzian pressure
R_q = RMS roughness in contact zone after deformation
R_x = effective radius of curvature in the x-direction
R_y = effective radius of curvature in the y-direction
U = rolling velocity
U^* = $\eta_o U/(E'R_x)$, dimensionless speed parameter
W = load
W_c = contact load ratio (load supported by surface contacts divided by total load)
W^* = $W/(E'R_x^2)$, dimensionless load parameter
x = x-coordinate (coincident with rolling direction)
X = x/a
y = y-coordinate
Y = y/b
α = pressure-viscosity exponent
η_o = viscosity under ambient condition
λ = h_a/σ, film thickness ratio
σ = composite RMS roughness before deformation

REFERENCES

[1] Dowson, D., and Higginson, G. R., Elastohydrodynamic Lubrication, Pergamon Press (1966).

[2] Hamrock, B.J., and Dowson, D., Isothermal Elastohydrodynamic Lubrication of Point Contacts, Part I — Theoretical Formulation, Part II — Ellipticity Parameter Results, Part III — Fully Flooded Results, Part IV — Starvation Results, *ASME J. of Lubrication Technology*, **98** and **99** (1976-77).

[3] Patir, N., and Cheng, H.S., 1978, "An Average Flow Model for Determining Effects of Three-Dimensional Roughness on Partial Hydrodynamic Lubrication," *ASME Journal of Lubrication Technology*, Vol.100, pp.12-17.

[4] Chittenden, R. J., Dowson, D., Dunn, J. F., and Taylor, C. M., "A Theoretical Analysis of the Isothermal Elastohydrodynamic Lubrication of Concentrated Contacts (Part 1 and Part 2)," *Proceedings of the Royal Society*, London, **A397**, pp.245-294 (1985).

[5] Zhu, D., and Cheng, H.S., 1988, "Effect of Surface Roughness on the Point Contact EHL," *ASME Journal of Tribology*, Vol.110, pp.32-37.

[6] Lubrecht, A. A., ten Napel, W. E., and Bosma, R., 1988, "The Influence of Longitudinal and Transverse Roughness on the Elastohydrodynamic Lubrication of Circular Contacts," *ASME Journal of Tribology*, Vol. 110, pp. 421-426.

[7] Kweh, C. C., Evans, H. P., and Sindle, R. W., 1989, "Micro-Elastohydrodynamic Lubrication of An Elliptical Contact with Transverse and 3-Dimensional Roughness," *ASME Journal of Tribology*, Vol. 111, pp.577-584.

[8] Ai, X., Cheng, H.S., and Zheng, L., 1993, "A Transient Model for Micro-Elastohydrodynamic Lubrication with Three Dimensional Irregularities," *ASME Journal of Tribology*, Vol.115, pp.102-110.

[9] Venner, C. H., and Lubrecht, A.A., 1996, "Numerical Analysis of the Influence of Waviness on the Film Thickness of a Circular EHL Contact," *ASME Journal of Tribology*, Vol.118, pp.153-161.

[10] Xu, G., and Sadeghi, F., 1996, "Thermal EHL Analysis of Circular Contacts With Measured Surface Roughness," *ASME Journal of Tribology*, Vol.118, pp.473-483.

[11] Zhu, D., and Ai, X., 1997, "Point Contact EHL Based on Optically Measured Three-Dimensional Rough Surfaces," *ASME Journal of Tribology*, Vol.119, pp.375-384.

[12] Jiang, X., Hua, D.Y., Cheng, H.S., Ai, X., and Lee, S.C., 1998, "A Mixed Elastohydrodynamic Lubrication Model With Asperity Contact," *ASME Journal of Tribology*, Vol.121, No.3, July.

[13] Zhu, D., and Hu, Y.Z., "The Study of Transition from Full Film Elastohydrodynamic to Mixed and Boundary Lubrication," *Proceedings of 1999 STLE/ASME H.S. Cheng Tribology Surveillance*, pp.150-156 (1999).

[14] Hu, D., and Zhu, D., "A Full Numerical Solution to the Mixed Lubrication in Point Contacts," *ASME Journal of Tribology*, **122**, pp.1-9 (2000).

[15] Zhu, D., and Hu, Y.Z., "Effects of Rough Surface Topography and Orientation on the Characteristics of EHD and Mixed Lubrication in Both Circular and Elliptical Contacts," *Tribology Transactions*, **44**, pp.391-398 (2001).

[16] Zhu, D., "Effect of Surface Roughness on Mixed EHD Lubrication Characteristics," presented at the 2nd World Tribology Congress, Vienna, Austria, September 3-7, 2001, and to appear in *Tribology Transactions*.

[17] Zhu, D., "Elastohydrodynamic Lubrication in Extended Parameter Ranges, Part I: Speed Effect, Part II: Load Effect," to appear in *Tribology Transactions*.

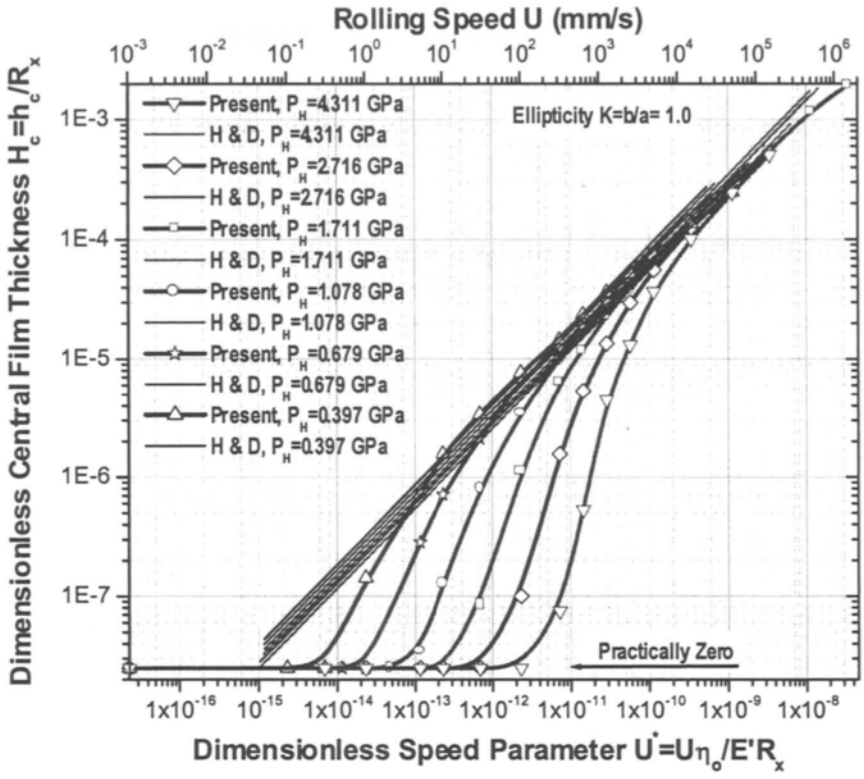


Figure 1. Effect of Speed on the Film Thickness (k=1.0)

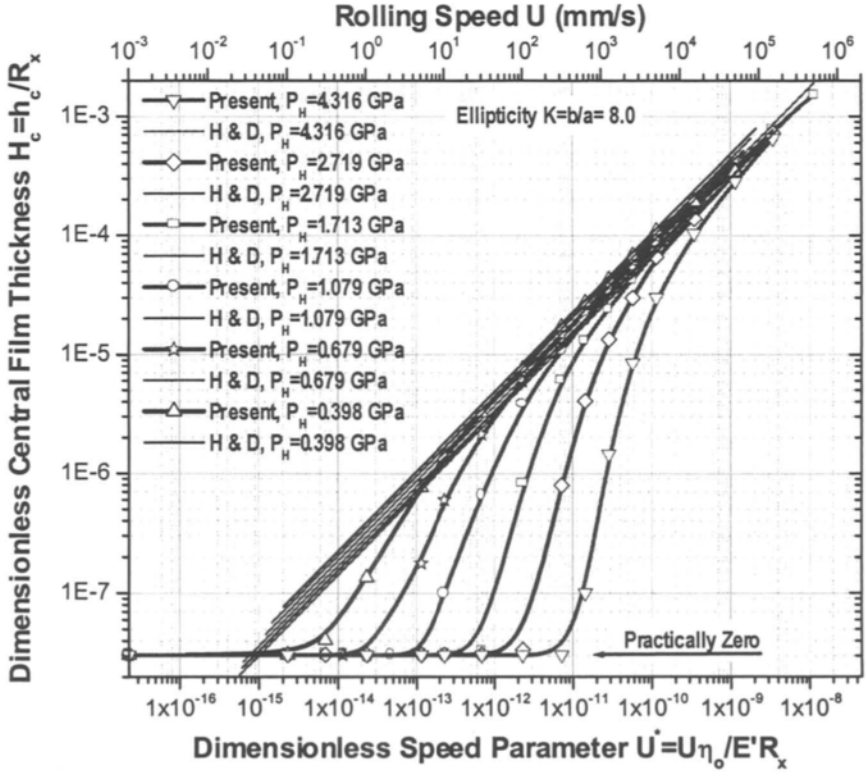


Figure 2. Effect of Speed on the Film Thickness (k=8.0)

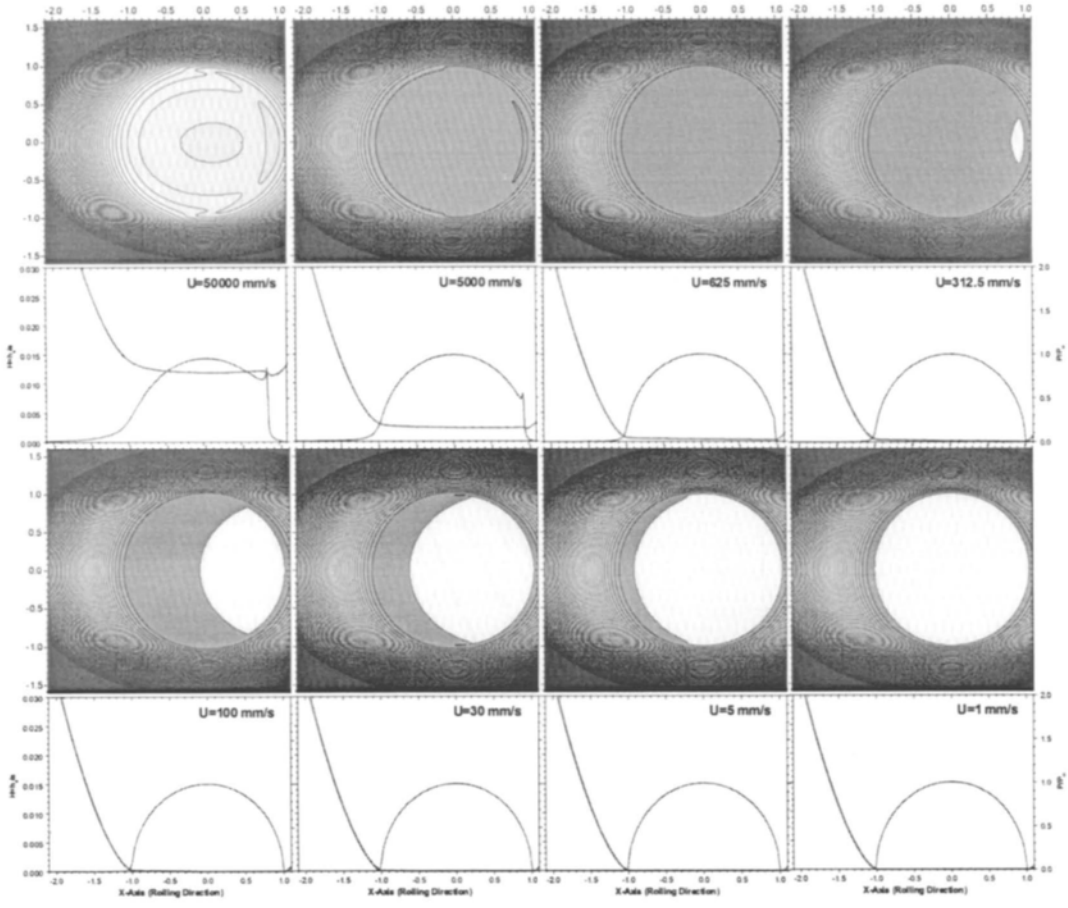


Figure 3. Sample Cases for Speed Effect (k=8, P_H=1.712 GPa)

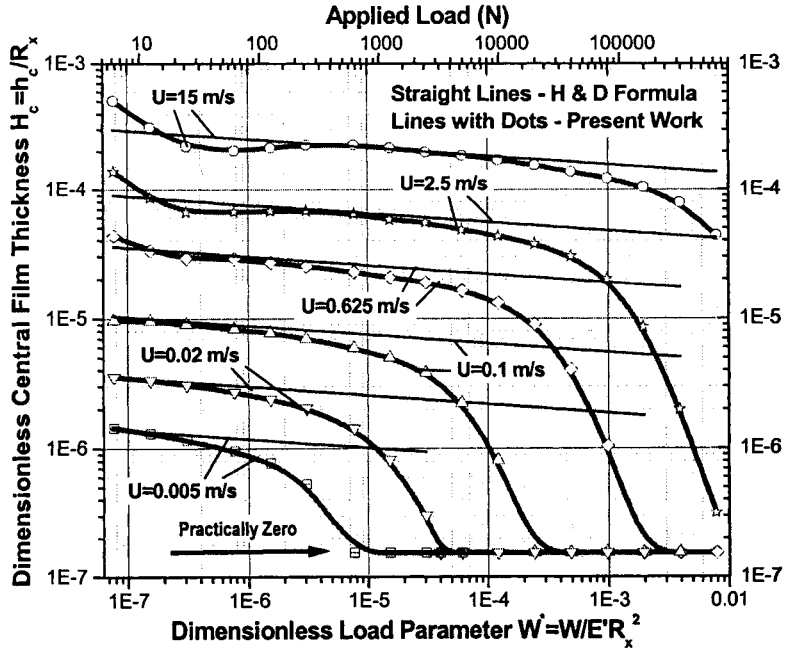


Figure 4. Effect of Load on the Film Thickness (k=8.0)

Paper XV (iv) "On the EHL Film Breakdown" by Dong Zhu

Supplementary Discussion / Presentation.

"This paper promoted considerable debate following its presentation. The Editors felt that it would be of interest and useful for delegates and readers alike to have access to the substantive points raised and the responses of the author as a supplement to the full paper. Further contributions can be found in the Written Discussion Section at the end of the Volume of Proceedings"

D. Dowson and M Priest

Review

This paper is the latest of a number of challenging and extremely controversial offerings from the author on the subject of the transition from fluid film (elastohydrodynamic) to solid-solid conditions.

The author predicts that, under critical conditions of load and speed, the film thickness declines abruptly to zero, in strong contrast to earlier findings. If true, the prediction could be extremely important. However, the present reviewer is unconvinced by this - or the author's earlier papers - for the following reasons:

1. The author does not say what feature of his model gives rise to this behaviour nor does he justify it in physical terms. This suggests that he does not understand how the collapse behaviour shown by his model comes about. In turn, this fails to give the reader any confidence that the predicted behaviour is realistic. The author should remedy this if he can.

2. The presence of a boundary between fluid film and solid-solid contact under steady state conditions, is incompatible with either mass conservation or with no-slip at the solid surfaces. At least some fluid must be entrained into the dry region unless there is slip at the wall. The author should address this point making clear which condition has been relaxed and why.

3. The predictions of film collapse are well within the range explored by earlier experiments and no such behaviour has been observed. Examples are given in the two papers referenced below which, for a circular contact, cover speed parameters (U^*) of down to around 10^{-14} with load parameters (W^*) of about 2×10^{-6}. It will be seen that most real lubricants are observed to behave in precisely the converse fashion, with boundary films, often of many tens of nm thickness, persisting at vanishingly low speeds. The author should at least comment on the discrepancy with observation.

In conclusion, it appears that the author's predictions are physically implausible, poorly justified and contrary to experimental observation. The paper could be very misleading and it is important that the author convincingly addresses the above points.

Additional References

Smeeth, M, Spikes, H , and Gunsel, S, Boundary Film Formation by Viscosity Index Improvers, Trib Trans, 39(1996) 3, 726-734.

Guangteng, G, and Spikes, H A, Boundary Film Formation by Lubricant Base Fluids, Trib Trans, 39(1996) 2, 448-454.

Author's Reply The author sincerely appreciates the comments from this reviewer, because he believes that deep understanding is often a result of thorough discussions.

It should be mentioned first that the results presented in this paper well agree with the Hamrock-Dowson formulae in the parameter ranges where the formulae were originally developed (e.g. at relatively light loads and moderate speeds). Such comparison actually confirms the validation of the model used in the present study. The result is also in good agreement with the experimental data such as those shown in the papers the reviewer referenced (see reply to question No. 3).

It is also important to mention that the present paper discusses the EHL film break down merely due to insufficient hydrodynamic action. It does not intend to cover the thin film collapse in a general sense. The author agrees that, as the lubricant film thickness continuously decreases, there may be a transition from the EHL film to a

boundary film and then to the boundary film collapse and possible surface failure. This transition may sometimes take place in a quite wide parameter range, say speed range, so that it may be difficult to separate the EHL film from the boundary film in an experiment. From the modeling point of view, however, the EHL film can be easily separated. The boundary film formation and breakdown are more difficult for computer simulation, but they are beyond the scope of the current paper.

1 By the term "EHL film breakdown", the author meant to explore the mechanical parameters that may reduce the hydrodynamic influence and, as a result, reduce the EHL filmthickness to practically zero, at which other films may become dominant in supporting the load and protecting the surfaces. The current topic is different from "film collapse", which is a more complicated phenomenon in general. Typically, the EHL film is at least one or two orders of magnitude thicker than the boundary film. Therefore, the term "practically zero" for the EHL film can mean a sufficiently small value at which the hydrodynamic action becomes negligible and the boundary film (or other films or even dry contact) becomes dominant. In most cases it may not necessarily mean an absolute zero gap between the two surfaces. The revised paper better reflects this definition.

2 The current model does not violate mass conservation as permitted by the Reynolds equation commonly solved in EHL research, nor the flow continuity condition. When the film thickness is sufficiently small (may not have to be zero), the left-hand side pressure driven flows become negligible, because they are related to the third power of the film thickness. When this happens, the majority of the fluid flows around the contact, but there may still be some lubricant passing through the very thin gap under the no-slip boundary condition. The treatment by the reduced Reynolds' equation does not conflict with this. It will yield a zero Couette flow simply because the flow through such a thin gap does not generate significant hydrodynamic pressure.

This is not a violation to the no-slip boundary condition, but a proper treatment for the terms that vanish under certain conditions. It is similar to the treatment of neglecting the mass of the fluid when solving fluid dynamic problems for low-Reynolds number flows, where mass exists, but just too small to affect anything. A numerical experiment has been conducted in order to study how important the left-hand side pressure flow terms are when the gap is very small.

The results are summarized in Figure 1. A small value, e, was used in the experiment as a criterion for turning the pressure flow terms on and off. Whenever and wherever the film thickness (or gap) is smaller than e, the reduced Reynolds equation without the pressure flow terms is used. Otherwise the full Reynolds equation is used. It can be seen that, as long as ε is sufficiently small (below 56.75 nm in this case), the pressure flow terms appear to be negligible and changing ε does not have a notable effect on the final solution.

3. The author had read the two papers referenced in the review and noticed that the results reasonably agree with the measurement data for base stocks in these papers.. Based on a Young's modulus of 73 GPa for glass (neglect the coating effect, although the chromium coating has a different Young's modulus), 220 GPa for 52100 steel and a radius for the steel ball of 0.019 m, the author found that the maximum Hertzian pressure due to the load of 20 N in both papers is about 0.33 GPa. To the speed range down to U* ~10 -14 , the film thickness down to Hc ~ 5.3 10 –8 (1 nm), and load W* up to ~ 10^{-6} the model used in the present paper does not predict EHL film breakdown. Actually, corresponding to this maximum Hertzian pressure at the given fixed load, the current model well agrees with the Hamrock-Dowson formula that was used by the two reference papers. At this Hertzian pressure, the model predicts a near straight-line relation between the film thickness and U* in a double logarithm scale, as indicated by the arrow (see Figure 2 below which is copied from the present paper), almost the same as that for the base oil given in the paper by Smith et al (also Fig. 2 in Guangteng and Spikes). In these ranges of data, the current model slightly deviates from the Hamrock-Dowson formula at Hc ~ 10^{-7} (about 1 nm or 2 nm), where the measured data exhibit notable scattering (Figure 2 in both Smith et al and Guanteng and Spikes). There, the rough surface

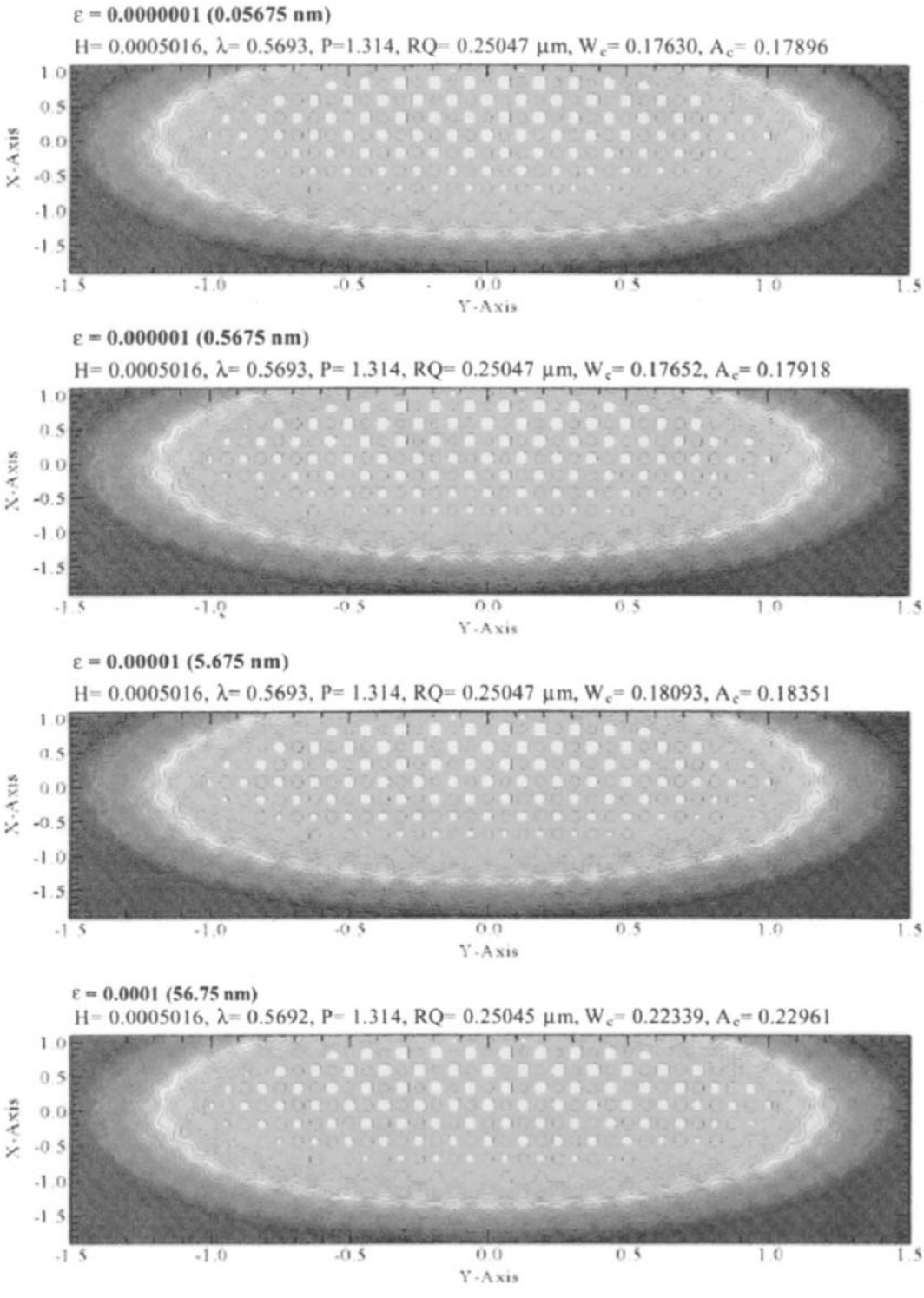

ε = 0.0000001 (0.05675 nm)
H= 0.0005016, λ= 0.5693, P=1.314, RQ= 0.25047 µm, W_c= 0.17630, A_c= 0.17896
ε = 0.000001 (0.5675 nm)
H= 0.0005016, λ= 0.5693, P= 1.314, RQ= 0.25047 µm, W_c= 0.17652, A_c= 0.17918
ε = 0.00001 (5.675 nm)
H= 0.0005016, λ= 0.5693, P= 1.314, RQ= 0.25047 µm, W_c= 0.18093, A_c= 0.18351
ε = 0.0001 (56.75 nm)
H= 0.0005016, λ= 0.5692, P= 1.314, RQ= 0.25045 µm, W_c= 0.22339, A_c= 0.22961
X-Axis
Y-Axis
1.0
0.5
0.0
-0.5
-1.0
-1.5

Figure 1

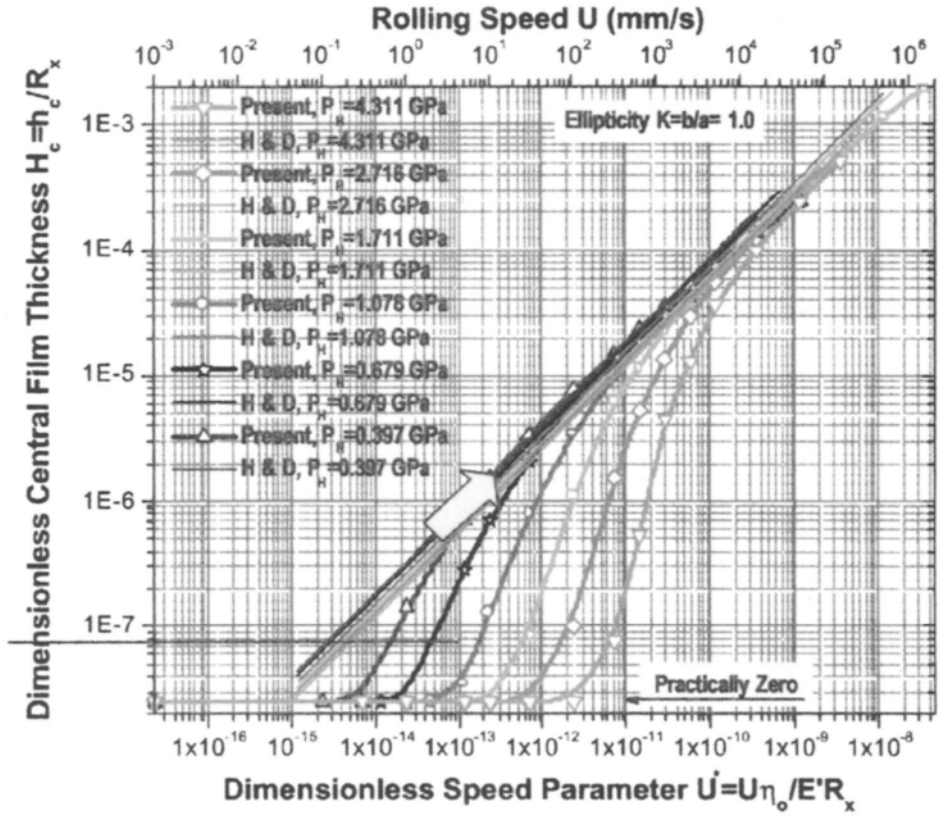


Figure 2

asperity effect was neglected, although the ball surface roughness is much larger than the measured film thickness.

The author had realized that in the two papers referenced by the reviewer no film collapse was observed as the speed continuously decreased. Instead, the film thickness reduction is in many cases slower than that predicted by the conventional EHL formulae, and the curve is often pretty much leveled when the film becomes very thin, in the range of several nano-meters or a few tens of nano-meters. This is actually the phenomenon caused by the boundary film, instead of the EHL film. The two papers were well written focusing on the topic of boundary film formation.The boundary film is more complicated, related to complex lubricant and interfacial chemistry and physics. However, what the present paper trying to discuss is still the hydrodynamic film that can be modeled and predicted numerically with continuum mechanics. The breakdown discussed in the paper is also strictly for the EHL film.

The author is aware of the fact that in the last 40 some years there have been hundreds, or even thousands, of papers related to the EHL experiments. There have also been tons of testing data by commercial companies not published. However, not all experimental results perfectly demonstrate the straight-line relation mentioned above. Figure 3 shows, for example, the work published in 1999 by Luo and Wen et al [1]. It can be seen that the slope of the curves for the film thickness vs. speed is not constant. The heavier the load is, the greater the negative slope can be

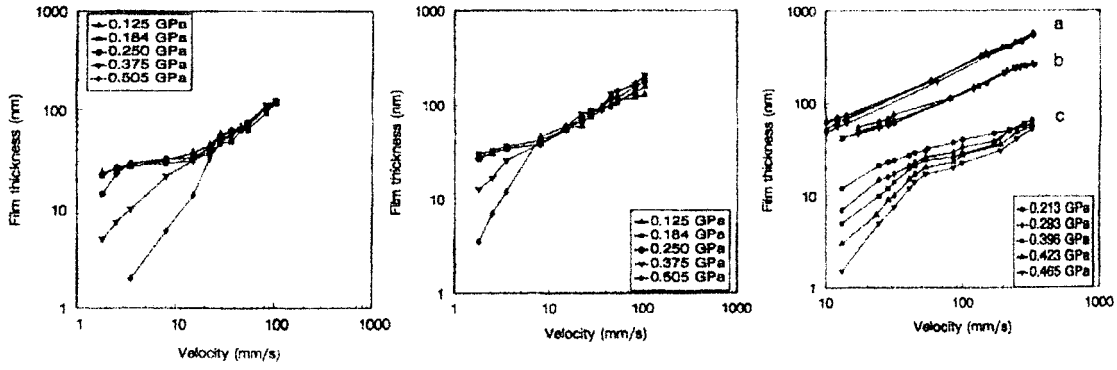


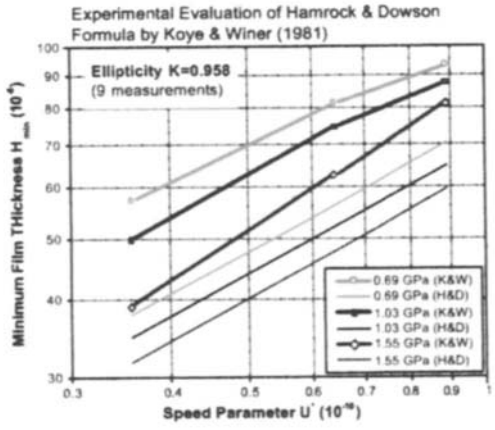
Figure 3

observed. There could be a quick decline as the speed decreases, especially when the load is heavy. This is qualitatively in good agreement with the conclusions in the present paper.

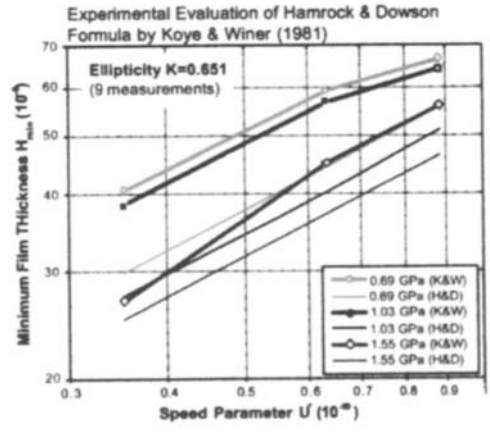
Figure 4 illustrates some results from a well-known paper by Koye and Winer published in 1983. It is observed that, even though the data are within a narrow speed range, the slopes of the film thickness curves do not seem to be constant, and, similarly, the slope appears to be a function of the speed and the load. The trends of the slope change are in agreement with the present paper.

Such disagreement with the straight-line relation can also be found from the experiments cited in Greenwood's 1970 paper[2]. The author has recently been involved in a CVT project with a major auto maker, who has conducted the EHL experiments at very high Hertzian pressure around 3-4 GPa. Their measured film thickness results are significantly lower than that predicted by the conventional formulae. Unfortunately, they have no interest in publishing their results.

It is also important to note that most experimental results obtained in the past were within relatively narrow parameter ranges under different conditions. The speed, for example, typically covered one or two orders of magnitude or less (see Figure 3 and Figure 4). However, the present paper extends the speed range to cover nearly 10 orders of magnitude, and the conclusions drawn are based on the extended ranges. The load range for most optical



(a)



(b)

interferometric measurements, similarly, is also relatively narrow and the maximum Hertzian pressure is still relatively light. Therefore, there is currently a lack of experimental data that can directly validate the present work in which the speed covers about 9-10 orders of magnitude and the maximum Hertzianpressure reaches 6-7 GPa.

The author has mentioned in the present paper that further experimental verification is certainly needed. However, this will certainly take a lot of time and efforts. He hopes to have more satisfactory answers to some questions in the near future.

Additional References:

[1] Jianbin Luo, Linmao Qian, Shan Wen, Liu Wen and Shizhu Wen, "The Failure of Fluid Film at Nano-Scale," Tribology Transactions, Vol. 42, (1999), pp.912-916.

[2] J.A. Greenwood, "A Re-examination of EHL Film Thickness Measurements," Wear, Vol.15 (1970), pp.281-289.

Further comment received from the Reviewer

1. The value of Rx in the papers by Smeeth et al and by Guangteng et al was 0.009525 mm. (3/4 inch dia. ball on flat disc) E(glass)= 75 GPa, E*=58 GPa, E'=116 GPa. Consequently the (Hertz) contact pressure was 0.527 GPa not 0.33 GPa.

2. The correct comparison with theory is achieved not by making the contact pressures equal but by making the dimensionless groups equal. The value of W* in the experiments was

20/(116E9 x 0.0095^2) = 1.91E-6

In the paper, for k=1, W* = 0.126E-6 corresponds to p0 = 0.397 GPa. (Noting that W* = (pi^3/12)(p0/E*)^3, this would be right for steel on steel) so the corresponding p0 for W* = 1.91E-6 would be(1.91/0.126)^(1/3) x 0.397 = 0.98 GPa. So the nearest curve is the one for p0 = 1.08 GPa. This deviates from the straight line below Hc = 1E-5 which would correspond to about 95 nm. In contrast, the experiments confirm Dowson & Hamrock behaviour down to a few nm for simple fluids.

A possible reason for the discrepancy is the low number of mesh points in the inlet region; at high loads and low speeds the pressure rises very rapidly and it appears that a fixed value of dX of about 0.01 has been used so there may only be a couple of mesh points in the narrow region of the inlet where the film thickness is determined.

The point that the dh/dx = 0 assumption for the thin film region makes no difference is accepted. Therefore why include it?

Author's Reply:

1 The author feels sorry that he mistakenly used 0.01905 m as the radius of the ball when evaluating the contact pressure. It was actually the diameter. However, this only caused a small quantitative difference. Even at P_H=0.527 GPa, the present results still predict a near straight-line relationship, which agrees with the Hamrock-Dowson formula quite well. The main points still hold.

2 The author doubts if the comparison by making the dimensionless load parameter equal is better. He believes that the contact pressure better represents the real physics, which is more closely correlated with lubricant rheology, lubrication mechanism, material performance and efficiency evaluation as well as possible failures. If the dimensionless load parameter is kept constant, for example, but the contact ellipticity changes from 1/8 to 8, the contact pressure would be reduced down to about 11.7%, at which the contact and lubrication mechanisms may be completely different. On the contrary, when keeping the contact pressure and R_x the same, for the same ellipticity change the dimensionless load needs to be increased more than 620 times. Even for the same ellipticity, we believe that the contact pressure is better correlated with the physics and the failures. Also, comparison by making the dimensionless load equal only yields a small quantitative difference.

Actually ,the author's main point is that the referenced experimental work was indeed focused

on the characteristics of the boundary film (clearly indicated by the titles of these two papers) at the nano/molecular scales, at which the lubricant may no longer be considered a continuum fluid and most likely the hydrodynamic influence already vanishes. This is beyond the scope of the present paper, which focuses strictly on the EHL film behaviour.

Relative to the number of mesh points, the author agrees that the mesh size is important to the numerical accuracy. In most cases a dimensionless mesh size of 0.011719 was used, but for some cases this size was further reduced down to 0.0058594 in order to double check the numerical accuracy. No significant difference in the calculated film thickness has been found between these two mesh sizes. These mesh sizes are already among the smallest, compared with other studies the author can possibly find in the published literature.

This Page Intentionally Left Blank

SESSION XVI
COMPONENT TRIBOLOGY (2) - TRANSMISSIONS

Chairman : Dr R S Dwyer-Joyce

Paper XVI (i) **Gearing Tribology**

Paper XVI (ii) **Tribology of Spline Couplings**

Paper XVI (iii) **Prediction of the Friction Coefficient Between Gear Teeth Using a Mixed Film Lubrication Model**

Paper XVI (iv) **Influence of Temperature Distribution on the Tribological Performance of Automotive Synchronisers**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Gearing Tribology

R. W. Snidle, H. P. Evans and M. P. Alanou.

School of Engineering, Cardiff University, Cardiff CF24 3TA, UK, e-mail: snidler@cf.ac.uk

Keywords: gears, scuffing, pitting, friction, elastohydrodynamic lubrication.

Abstract: *The paper presents a brief review of developments in understanding of gear tooth contact lubrication in relation to problems of surface durability and distress. Gear tooth contacts tend to operate under conditions where the lubricating oil film is thin compared to surface roughness. This feature is shown to have a significant effect on scuffing capacity, friction and micropitting. Recent developments in thin film micro elastohydrodynamic lubrication (EHL) theory are described and these should lead to a better understanding of the behaviour and modes of surface distress in gears.*

1. INTRODUCTION

Amongst the earliest references to gearing is that to the work of Archimedes who invented what is now called the worm gear in about 250 BC. The oldest surviving gears from antiquity are those in the famous Antikythera Mechanism [1] which was discovered in a sunken ship in the Greek islands in 1900 and dates from around 87 BC. It is perhaps not surprising to learn that Leonardo da Vinci designed gears and invented tooth profiles similar to those in use today. The Swiss mathematician Euler worked out the fundamental "law of gearing" (the requirement for conjugate action to give a constant speed ratio) in about 1754 [2]. During the industrial revolution the art of gearing developed rapidly and scientific principles began to be applied. In 1893 Lewis [3] published his famous formula for calculating bending stresses in gears which is still in use today, and of course Hertz [4] provided the basis for the determination of contact stresses between gear teeth in 1881. Gears have therefore been used and developed successfully for over 2000 years and they will undoubtedly survive long into the current millennium in all kinds of machinery and vehicles because of their advantages in transmitting power and matching the speeds and torques of one machine to another. In spite of their complexity and manufacturing challenges gears have survived because of their high power to weight ratio, their compact nature, and their high efficiency (typically better than ninety percent).

Gears transmit power through rolling/sliding contacts between their teeth which must operate reliably for many thousands of hours with negligible wear. Lubrication of tooth contacts, both as a means of minimising surface damage and as a way of reducing friction due to sliding, is therefore of crucial importance. Gears have been made from various materials, including polymers, but in this paper we are concerned with steel gears which are the most important type in use in vehicles, aerospace and industrial applications. It is now known that elastohydrodynamic lubrication (EHL) is the mechanism responsible for the protection of steel gear teeth. The basic kinematics of involute gears (by far the most important type) necessitate a relatively low geometrical conformity between the mating teeth (Figure 1) and this gives rise to contact pressures, which can be as high as 1.5 GPa if high strength surface-hardened steels are used. Under these conditions the viscosity of typical lubricating oils is increased by several orders of magnitude and significant "Hertzian" elastic deformation of the surfaces occurs. Sliding also takes place between the teeth given by

$$u_s = (\omega_1 + \omega_2)s$$

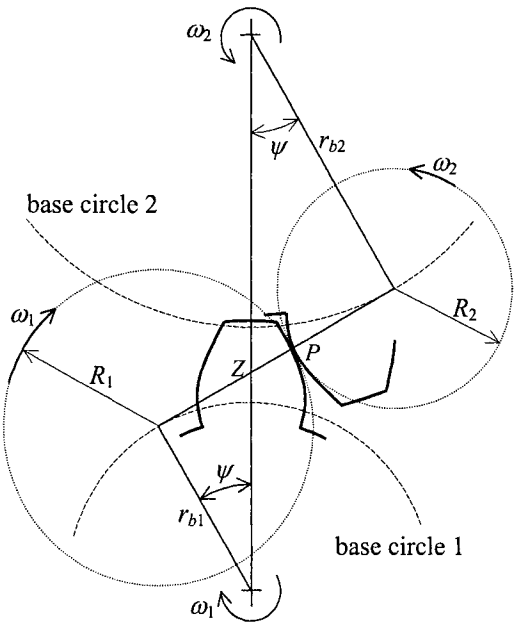


Figure 1. Illustration of gear teeth in contact showing equivalent rolling/sliding circles. Sliding is proportional to $s = ZP$, where Z is the pitch point and P is the instantaneous contact between the teeth.

where s is the distance of the instantaneous contact between the teeth from the pitch point Z shown in Figure 1 and ω_1 and ω_2 are the angular velocities of the two gears. The entraining velocity, which is responsible for the formation of any oil film by hydrodynamic action, is given by

$$\bar{u} = (r_{b1}\tan\psi + s)\omega_1/2 + (r_{b2}\tan\psi - s)\omega_2/2$$

Both sliding and entraining velocities therefore vary throughout the meshing cycle as does the relative contact geometry. A typical ratio of the sliding to entraining velocities at the extremes of tooth contact is about 0.5, but can be as high as unity in very coarse gears. Contact loads are uncertain due to load-sharing between simultaneous, overlapping tooth contacts, and dynamic effects at high speeds. A necessary compromise in high speed gear design is the need to minimise sliding by using rather small teeth whilst choosing teeth of sufficient size to resist the bending stresses caused by tooth loading. This explains the use of rather fine pitch teeth combined with relatively large face-width in high speed naval reduction units.

Lubricant film thickness in involute gear teeth is usually assessed using EHL line contact theory [5] in the form of Dowson and Higginson's formula

$$h_{\min}/R = 1.6(\alpha)^{0.6}(\eta_0\bar{u})^{0.7}(E')^{0.03}R^{-0.57}w'^{-0.13}$$

which has been experimentally validated for smooth surfaces, most notably by Dyson et al [6]. Modifications are available for situations where the underlying assumptions do not apply, e.g. to correct for temperature rise in the inlet to the contact [7]. This theory is for infinitely long "line" contacts which is the situation in the majority of both spur and helical involute gears. Gears that have "point" contacts such as those in "Wildhaber-Novikov" [8] conformal gears, and angle-drive gears such as hypoids and spiral bevels, require formulas for predicting film thickness under side-leakage conditions [9,10].

Under ideal conditions gear teeth operate with films which effectively separate the surfaces, and the above EHL film formula may be used. However, under demanding conditions of low speed or high temperature, the film predicted by classical EHL theory may be small compared with the roughness present on the teeth. The gears then operate under conditions described as "partial", "mixed" or "micro" EHL. A conventional measure of the severity of lubrication is provided by the lambda-ratio, which is defined as the mean oil film thickness divided by the combined RMS roughness of the two surfaces, i.e.

$$\Lambda = h/\sqrt{\sigma_1^2 + \sigma_1^2}$$

Parameter Λ has been used as an indicator of relative life in rolling element bearings [11]. Most gears, however, are much rougher than bearings and rarely have a finish better than about 0.4 μm R_a with nominal oil films typically 1 μm at most. Clearly, many practical gears operate with Λ significantly less than unity. In theoretical solutions of both the dry contact and micro-EHL problems significant roughness leads to a severe rippling in the contact pressure distribution with maximum values far in excess of the Hertzian values for smooth surfaces. Surface roughness

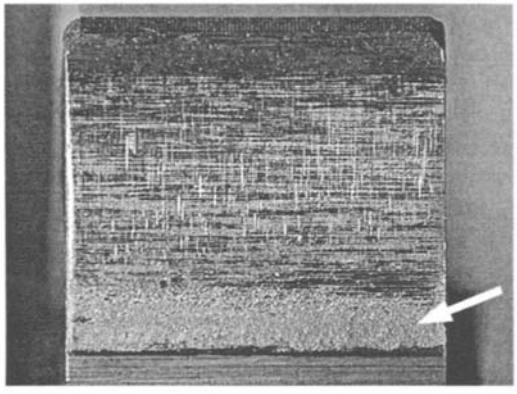


Figure 2. Gear tooth showing micropitting damage (indicated by arrow) in the root region. (*Design Unit*, Newcastle University)

effects are important in the gearing problems of "micropitting" and "scuffing" [12] which are discussed in Sections 2 and 3.

This paper briefly reviews some current research on gear tooth contact lubrication that will lead to a better understanding of the problems of micropitting, scuffing and gear and transmission lubrication in general.

2. MICROPITTING

Micropitting is pitting (rolling contact fatigue) on the scale of the roughness, as opposed to classical pitting which occurs on the scale of the nominal Hertzian, or macro, contact. With increased use of hardened gears it has become of widespread concern to the gearing industry. Although it can affect all types of gears, it has become particularly troublesome in heavily loaded gears with hardened teeth. The problem has been extensively studied by researchers [13,14,15,16] and is well known to gear designers [17]. Micropitting is characterised by the presence of fine surface pits and the occurrence of local plastic deformation and shallow surface cracks. It produces significant wear of the surfaces causing loss of profile of the teeth, leading to noise. Serious cases can precipitate scuffing and even complete fracture of the tooth. The problem is widespread and its solution has a high industrial priority. Figure 2 shows a gear tooth with characteristic micropitting damage.

The causes of micropitting are not fully understood, but surface roughness and high contact stresses at asperity or micro contacts within the overall Hertzian contact are key features, so that a thorough understanding of the stresses at asperity contacts is clearly needed. Useful insights into the behaviour of the contact of real surfaces can be obtained using dry elastic contact simulations [18,19] or elastic/plastic contact simulations [20]. However, it is necessary to consider the micro-EHL effects of the oil film since these influence the contact pressures and, through non-Newtonian and time-dependent behaviour, determine the traction forces at the micro contacts. The important sub-surface stress field is therefore affected by the presence of a micro EHL film. In general when roughness is present the maximum sub-surface shear stress field occurs much closer to the surface as seen in the micro EHL simulations in section 4.

The mechanism of plastic ratchetting brought about by asperity pummelling or cyclic stressing at the edges of micro contacts described by Kapoor and Johnson [21] may also be a factor in micropitting, and work directed towards an understanding of failure of railway wheel contacts undergoing plastic deformation may also be relevant [22]. Cavitation occurring within the nominal EHL contact due to asperity action discussed in Section 4 may also be a factor contributing to micropitting.

Recent fatigue testing of gears in the ground and superfinished state [23] shows the benefit to be obtained in terms of pitting life by superfinishing.

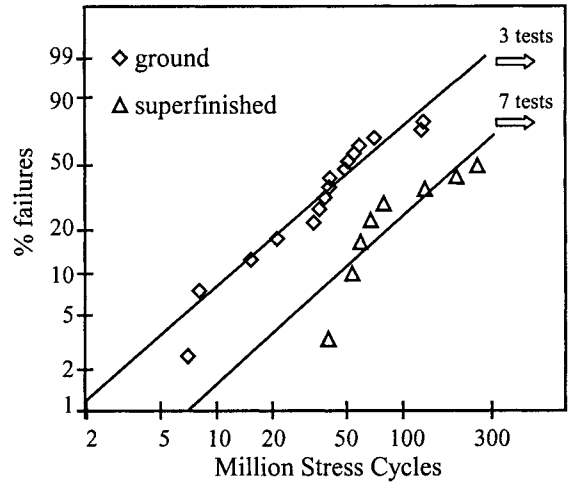


Figure 3. Results of micropitting fatigue tests on ground and superfinished gears.



Figure 4. A superfinished steel test disc showing high quality finish achieved.

Figure 3 is a Weibull plot comparing performance of hardened, ground steel gears (mean R_a = 0.38 μm) with superfinished gears (mean R_a = 0.07 μm) of the same material. The lives of the gears with superfinished teeth in this programme were found to be about four times greater than those of the ground gears.

3. SCUFFING

Scuffing is a serious form of surface failure that can affect gears running at high speeds or at high temperatures. Although there is no accepted explanation for the mechanism of scuffing, in many situations it is linked to physical failure of the EHL mechanism. To avoid scuffing problems gear designers often specify chemically active extreme pressure (EP) additives in the lubricating oil, some form of surface treatment for the gears such as case-carburising/hardening, nitriding or coatings, or a combination of these. Two particularly promising methods for upgrading the scuffing resistance of steel gears are: improvement of surface finish; and use of super hard coatings, e.g. diamond-like carbon (DLC) coatings. Polishing hardened and ground steel discs can produce a significant improvement in scuffing performance [24]. Superfinished discs used in scuffing tests showed a two fold load limit increase in comparison with ground disks, with the additional benefit of significantly lower friction and bulk metal temperatures. It should be emphasised that the

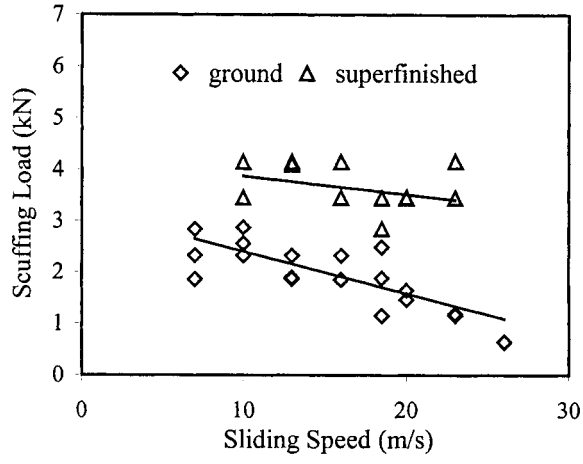


Figure 5. Results of scuffing tests on ground and superfinished discs.

ground and superfinished discs were of the same steel which was case-carburised and hardened to the same specification for both types of finish. The quality of the polishing process used to achieve these results is illustrated in Figure 4, and Figure 5 shows collected results from the series of tests. The lubricant in all tests reported in this paper was Mobiljet 2 at a feed temperature of 100°C

Thin hard coatings have been found to offer benefits in many non-gearing applications in terms of wear resistance, chemical inertness and low coefficient of sliding friction. They are deposited using different variations of the physical vapour deposition technique using magnetron sputtering [25]. The DLC types have been selected for a number of industrial applications both for tooling and machine elements. Part of this same family are the metal doped DLC coatings (Me-DLC), which appear to offer a better adherence and greater toughness than bare DLC coatings and show particular promise as coatings for gears. A second type of coating that has been considered is boron carbide (B_4C) which is an attractive alternative to Me-DLC, as outlined by Eckardt *et al.* [26], giving enhanced chemical and thermal stability combined

Table 1 Details of Scuffing Test Coatings

Coating	Thickness	Elastic Modulus	Hardness
Me-DLC	3.2 μm	121 GPa	11.4 GPa
B_4C	2.7 μm	180 GPa	17 GPa

with extreme hardness. Moreover, the coating process requires a less reactive gas, which is advantageous. The chemical structure of the coating is amorphous with a majority of B_4C carbides. Some boron oxides as well as boron oxy-carbides are also normally present with about 15-18 % oxygen located mostly at the surface [27].

Me-DLC and B_4C coatings, whose details are given in Table 1, were used in scuffing tests carried out at sliding speeds ranging from 7 m/s to 20 m/s. The tests demonstrate the potential benefits of these coatings on both case-carburised and nitrided steels, although superfinishing achieved a similar performance in terms of scuffing load capacity as may be seen in Figure 6 which ranks a wide range of different material and finish combinations (see Table 2) which have been tested as part of an ongoing programme. The triple combination of a nitrided steel substrate plus superfinishing and a diamond-like coating (configuration 9) shows particularly impressive scuffing resistance. The figure also shows the relative performance of the combinations in terms of average mean bulk temperature of the discs just prior to scuffing and the average traction coefficient under the same conditions. Again, superfinishing appears to confer similar benefits in terms of reduced temperature and traction as that obtained from super hard coatings.

Analysis of EHL between typical gear tooth surfaces in contact strongly suggests that eventual scuffing of ground surfaces can be explained by the physical failure of the lubrication mechanism due to sideways leakage of lubricant near the edges of nominal point contacts, and within nominal line contacts due to waviness in the direction transverse to that of rolling/sliding. Figure 7 shows a scuffing scar on an elliptical contact in a disc machine test. It is possible to envisage the leakage of oil from the loaded area in a transverse direction along the relatively deep valleys between asperities produced by the grinding process as a lubrication failure mechanism. It is interesting to note that for rough ground disc tests, scuffing was initiated at the edges of the contact without exception, and not near the contact centreline where pressures and "flash temperature" are highest. Roughness valleys are therefore seen as vulnerable features of the surface when considering lubrication effects. This concept has been pursued theoretically [28] in a simple model of EHL film failure in which a dry contact

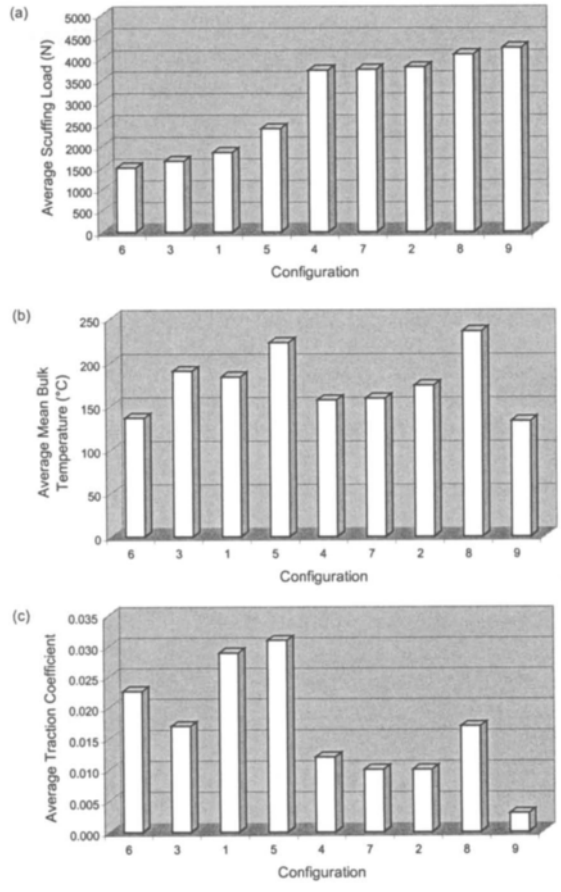


Figure 6 Comparison of performance between the different configurations (defined in Table 2) for tests performed at 16 m/s sliding speed. Uncoated configurations added for comparison. (a) scuffing load; (b) average mean bulk temperature at scuffing load; (c) average traction coefficient at scuffing load.

analysis was used to provide the geometry of gaps (due to valley features of roughness) between lubricated surfaces under thin film conditions. The magnitude of the transverse pressure gradients needed to cause significant sideways leakage in these gaps, and hence collapse of the film, is consistent with that present at the edges of real EHL contacts under similar operating conditions. These ideas are being explored more fully with the aid of detailed micro EHL solutions that include these side leakage effects, as discussed in section 4.

Table 2 Different substrate/finish/coating configurations tested

Configuration	Substrate	Finish	Coating	No of tests at each sliding speed		
				7 m/s	16 m/s	20 m/s
1*	case carburised	ground	none (reference)	3	1	1
2**	case carburised	superfinished	none	-	2	2
3	case carburised	ground	Me-DLC	3	3	3
4	case-carburised	ground	B_4C	2	4	3
5	nitrided	ground, compound layer off	none	4	3	3
6	nitrided	ground, compound layer on	none	4	3	3
7	nitrided	superfinished	none	-	3	3
8	nitrided	ground	Me-DLC	-	1	-
9	nitrided	superfinished	Me-DLC	-	3	-

* one other test was performed at each of the sliding speeds of 10, 13, 18.5 m/s, plus two other tests at 26 m/s.
** two other tests were performed at each of the sliding speeds of 10, 13, 23 m/s, plus three other tests at 18.5 m/s.

4. MICRO EHL THEORY AND APPLICATION TO GEARING

Considerable progress has been made with modelling of roughness in EHL contacts starting with idealised single features or sinusoidal waves and progressing to measured (real) roughness on both line and point contacts. Most analyses are based on a two-dimensional or nominal line contact model. This is a reasonable assumption in the case of conventional involute, parallel axis gears where rolling/sliding takes place across the finishing marks which are close to being two-dimensional, as may be judged from Figure 2.

The early attempts to model micro EHL were based upon a sinusoidal representation of roughness under steady state conditions (Goglia et al [29]; Lubrecht et al [30]; Kweh et al [31]; Elsharkawy and Hamrock [32]) More recently real surfaces as represented by profilometer traces of actual surfaces have been used in steady-state solutions (Sadeghi [33]; Kweh et al [34]; Venner, et al [35]). The problem of non-steady (i.e. moving roughness) conditions has been addressed by Chang and Webster [36], Chang [37], and Chang et al [38] using sinusoidal roughness, and by Ai and Cheng [39] with random roughness. The important influence of non-Newtonian (shear thinning) effects in micro EHL solutions has been reviewed and studied by Chang and Zhao [40]. The inclusion of

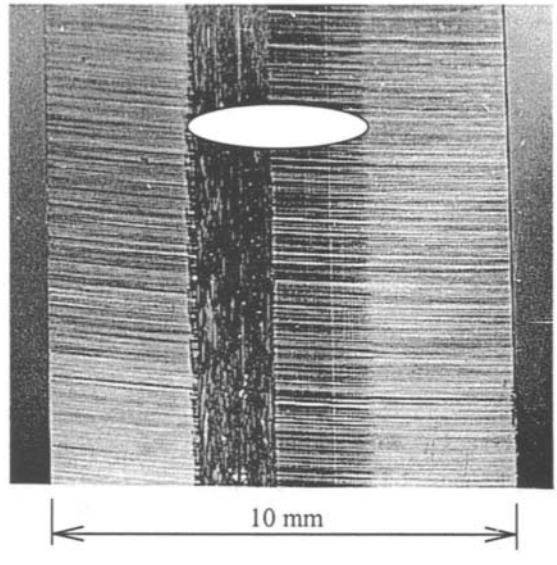


Figure 7. Scuffing scar on an axially-finished test disc. Superimposed ellipse indicates approximate equivalent Hertzian contact.

non-Newtonian lubricant behaviour and consideration of the roughness moving relative to the contact tend to reduce the pressure ripples but also tend to give thinner oil films.

Zhu and co-workers [41,42] were among the first to develop a "mixed" lubrication EHL model for elliptical contacts in which films below an arbitrary low value are regarded as creating contacts with a corresponding simplifying modification of the Reynolds equation presumed at the contacting areas. Their analyses have predicted a considerable degree of direct surface contact [43], which is stated to be insensitive to the (small) film level at which contact is taken to occur. The treatment of contact in this model is somewhat arbitrary however, and the assumptions made in dealing with contact remain to be verified. In contrast our contribution to rough surface EHL under severe thin film conditions is based on a theoretical model for line contacts (the appropriate configuration for spur gears) that solves the elastic and hydrodynamic film thickness equations simultaneously using a consistent, mass-conserving, fully coupled method [44, 45]. This new robust numerical technique has enabled results to be obtained for extremely low Λ values with contact, if it occurs, established directly from solution of the time-dependent equations. A second important feature that emerges using this approach, dependent upon the oil rheology model, is that of lubricant cavitation under severe conditions of thin films/high roughness [46].

In some circumstances contact occurs between the surfaces as part of the solution process. If, in a particular timestep a converged result is obtained with a negative calculated film thickness, then the most negative film thickness is set to zero. The hydrodynamic equation at this nodal point is deleted from the problem matrix, but the elastic deflection equation is retained. (In this way the film thickness value of zero is a boundary condition for the elastic deflection equation which ensures that the pressure distribution obtained remains entirely consistent with the film shape. The pressure developed in the coupled solution at the contacting node is then an automatic boundary condition for the hydrodynamic equation on both sides of the contacting node.) The timestep is then re-calculated. This procedure is repeated for the timestep, adding no more than one contact point per timestep re-calculation, until a converged result with no negative calculated film thicknesses is obtained. Contact, as calculated in this way, is found to be a relatively infrequent event in the line contact results obtained as discussed below.

A further instance where a full film solution is not obtained at each node point is when what may be termed "in-contact cavitation" occurs. Cavitation at the exit of a contact, where the calculated pressure falls to sub-ambient values, is a natural part of any EHL solution. Situations occur in rough EHL contacts where composite valley features are increasing in volume with time. If flow into such a feature from the micro contacts at either side is insufficient, then the pressure in the composite valley feature will fall. If this effect is sufficiently strong then the oil 'trapped' in the composite valley can become entirely de-compressed, returning progressively to ambient pressure. When the calculated pressure becomes negative the lubricant is regarded as cavitated. In these solutions the occurrence of in-contact cavitation is found to increase with the degree of sliding and this suggests, perhaps, a possible erosion wear mechanism that may be related to micropitting damage.

Both localised contact and in-contact cavitation are effects determined from the numerical analysis. The sensitivity of these calculated events to the details of the numerical scheme is the focus of a current investigation by the authors that will be reported in due course. There seems to be little doubt that micro contact does indeed take place between rough surfaces in EHL, as electrical contact resistance measurements provide strong supportive evidence of this phenomenon. The occurrence of calculated in-contact cavitation is a new feature of this kind of analysis, which may be potentially damaging to the surfaces as it takes place in isolation from the ambient atmosphere.

The line contact micro-EHL analysis technique has been applied to a set of profiles corresponding to the later (most heavily loaded) stages in an FZG gear test. The results form part of an investigation programme comparing micropitting performance of test gears using different oils. Use of experimental gear tooth profiles under the real operating conditions of the gears is judged to be an important requirement as

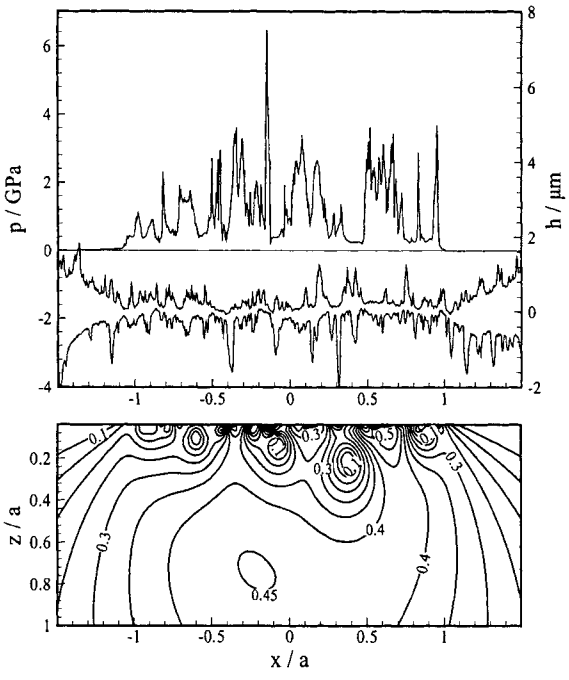


Figure 8. Pressure and film thickness at a particular timestep in a micro EHL solution for gear tooth contact in an FZG test. Corresponding maximum Hertzian pressure = 1.4 GPa. Also shown are contours of τ/GPa, the sub-surface maximum shear stress.

surface profile topography is significantly modified by running-in, and engineering interest must be primarily focussed on the stress and film thickness response actually encountered by the components.

Some results from this work are shown as follows. Figure 8 shows the pressure and film thickness variation at a particular timestep corresponding to load stage 9 in a standard FZG test in which the equivalent Hertzian (smooth surface) maximum contact pressure is 1.4 GPa. The profiles used in the simulation are taken from the gears after running at this stage, and the properties of the oil are those for the lubricant used in the tests. The FZG pinion speed was 2250 rpm and oil viscosity (0.028 Pas) was taken to be that at the feed temperature of 60°C. Non-Newtonian lubricant behaviour was modelled according to the Eyring equation. For the result shown in Figure 8 a pressure-dependent Eyring stress given by $\tau_0 = 0.03p$ was used subject to a minimum value of 10 MPa. Also shown beneath the film thickness distribution is a contour plot of the subsurface maximum shear stress corresponding to the pressure distribution shown. It is clear that pressures well in excess of the Hertzian values are generated on individual micro asperity contacts. Pressure values in the composite valley areas are much lower, but clearly contribute significantly to the load carrying of the overall EHL film. The elevated values and rapid spatial variation of pressure leads to concentrations of shear stress levels close to the surface. Maximum shear stress values as high as 0.7 GPa occur compared with the smooth surface value of 0.42 GPa. Video sequences made up of such figures enable the variation of pressure and film thickness to be assessed in a qualitative way and are invaluable for developing an understanding of the mechanisms at work as individual asperity features enter the Hertzian contact area and subsequently pass through it, as discussed in [45]. Videos for results such as those presented in this paper are available from the authors on CD by request. Contact between individual asperities and internal cavitation have been observed in such numerical solutions, as discussed above and in [46] where a thinner oil (η_0 = 0.0048 Pas, α = 11.1 GPa^{-1}) was used for the analysis.

The EHL response of these contacts requires a statistical evaluation of loading severity with parameters obtained by averaging over a sufficiently long transient analysis. Figure 9 shows two such parameters (defined in the *Notation* section) used to compare the EHL response of profiles taken from FZG gear teeth at four different load stages. Parameter $\Phi(p)$ is the fraction of the Hertzian contact area that has a pressure that exceeds p. Load stage 6 has the least amount of its contact area subject to extremes of pressure, and load stage 9 has the greatest amount. However all cases tend to the same level of extreme pressure with about 0.25% of the area subject to pressures of 4.5 GPa. The maximum pressures are thus seen to be significant in relation to the hardness of the gear teeth (700 Vickers hardness number) suggesting that surface modification is achieved by plastic deformation under the action of the EHL asperity pressures. The fraction of the contact area subject to pressures between 1 and 2 GPa is seen to rank the profiles according to the applied load, whereas true extremes of pressure may well be governed by

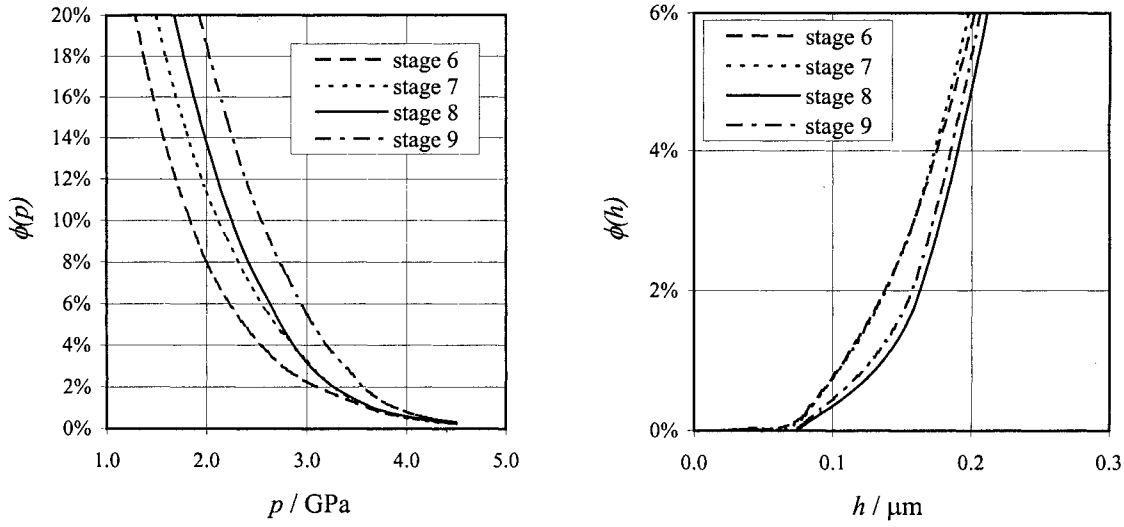


Figure 9. Maximum pressure $\Phi(p)$ parameter and low film parameter $\Phi(h)$ obtained by averaging results of micro EHL solutions corresponding to four different load stages in an FZG gear test.

surface hardness considerations. Figure 9 also gives the cumulative film thickness distribution for small film thickness values for the four profiles. Parameter $\Phi(h)$ is the fraction of the Hertzian contact area that has a film thickness that is less than h. The lowest film values observed are of the order 0.075 μm, and the earlier load stages have a greater tendency towards the lower film values, for example 2% of the contact area for load stage 6 is below 0.15 μm, a figure that reduces to 1% for load stage 9. These results thus suggest that the surface modification (running in) induced at each load

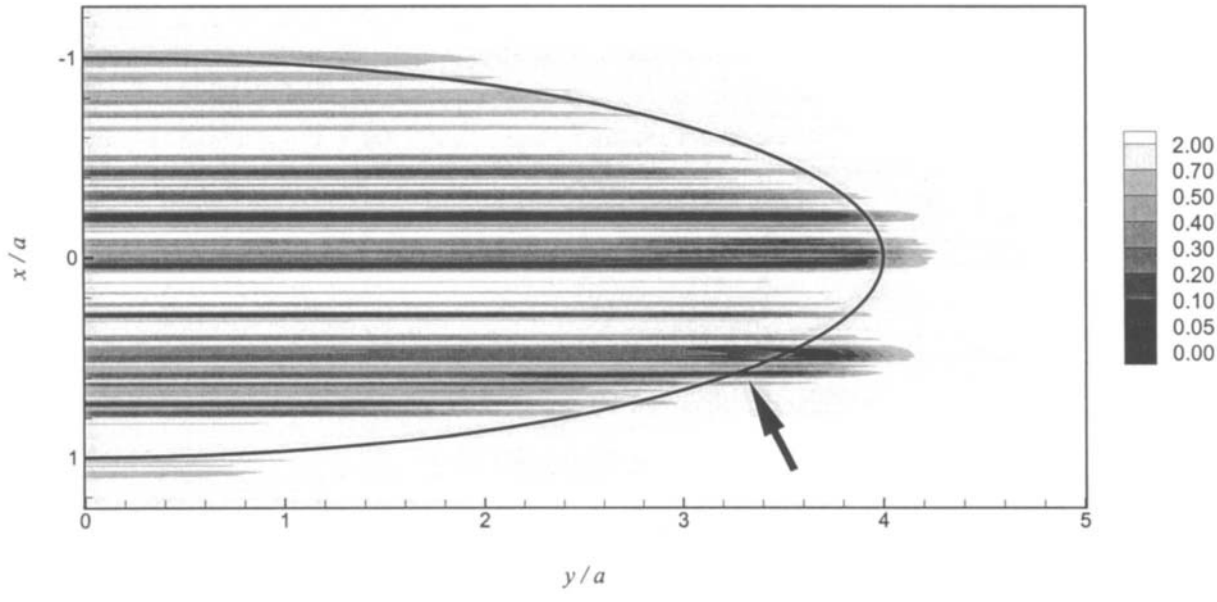


Figure 10. Contours of film thickness/μm in micro EHL solution corresponding to the contact in a crowned disc test as shown in Figure 7. Arrow indicates position of minimum film thickness (0.011 μm).

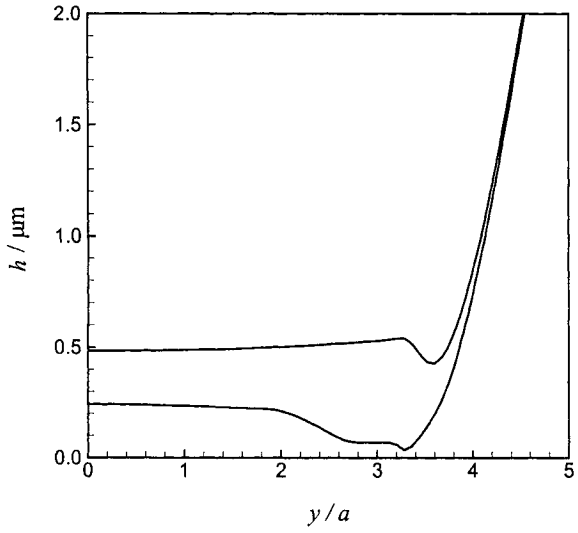


Figure 11. Transverse film thickness profiles through position of minimum film thickness corresponding to Figure 10. Upper curve shows smooth surface result and lower curve is the rough surface result.

stage causes the rough surface to adopt a shape that is more able to generate lubricant films on the micro asperities.

The above micro-EHL analyses are based on a line contact model, but in all real gear contacts side-leakage effects will occur even in nominal line contact gears due to edges and waviness of the finish in the axial direction. In gears which are deliberately arranged to have point contacts (such as crowned spur gears, crossed axis gears and special types such as Wildhaber-Novikov) a full side-leakage analysis becomes a necessity. Progress in modelling real roughness on an elliptical contact corresponding to the Cardiff crowned disc machine has been made by Holmes [47]. Figure 10 shows predicted contours of the film under rolling/sliding conditions with roughness of R_a = 0.26 μm (0.25 mm cut off) on both surfaces. Entrainment is in the direction of the minor axis of the contact. Over most of the contact the behaviour is effectively one-dimensional, but thinning of the film occurs at the transverse edges due to sideways flow of the lubricant in the deep valley features. This thinning (compared to that on the centre line, y = 0) is much more severe than that occurring in the corresponding smooth-surface case as is evident from the comparison shown in Figure 11. This predicted behaviour correlates, qualitatively, with the observation that scuffing in the disc tests initiated at the transverse edges of the contact.

5. WORM GEARS

The lubricated contacts in worm gears are similar to those occurring in high conformity Wildhaber-Novikov gears [48] giving rise to particularly severe conditions because of the elongated shape of the dry contact region. The contact geometry is non-symmetrical/non-Hertzian and although entrainment of the oil is predominantly along the major axis of the contact its direction varies throughout the area of the conjunction (i.e. a significant spin component of velocity is present). High sliding and heat generation in the contact necessitate a thermal/non-Newtonian EHL treatment. A typical solution obtained for a worm gear contact [49] is shown in Figure 12. An interesting feature of the contours shown in this figure is the appearance of a narrow region of almost zero film thickness (within the 0.47 μm contour towards the exit) which effectively separates the overall contact into two regions. This behaviour is explained as follows. A feature of EHL is that within the thin-film main load-bearing part of the contact (corresponding roughly with the equivalent Hertzian region) flow of the lubricant is dominated by the entraining motion of the surfaces

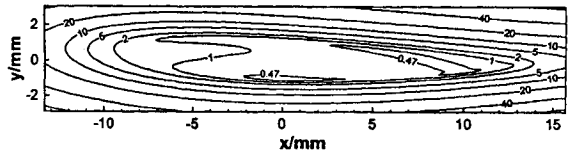


Figure 12. Contours of film thickness/μm obtained from EHL solution for a worm gear contact. (Inlet is on the left)

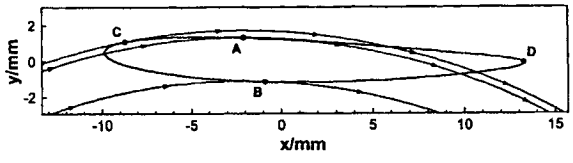


Figure 13. Dry elastic contact corresponding to Figure 12 with the key entrainment velocity flowlines superimposed.

(Couette flow). The pressure-induced (Poiseuille) flow is negligible because of the tremendous increase in viscosity due to pressure. Oil entrained into this region will therefore follow the entrainment streamlines, and these may be determined from purely kinematic considerations of the gears. It also follows that at points where entrainment lines are tangential to the edge of the high pressure region there will be zero net entrainment of oil into the contact at the tangency point, and this indicates a "line" of poor film formation. This hypothesis is reinforced by consideration of Figure 13 which shows the entrainment streamlines corresponding to Figure 12 together with the dry contact region. The striking similarity between the region of severe film thinning and the path of the tangential entrainment streamline (tangential at point A) within the contact is clear.

6. FRICTION

The efficiency of work transfer in a gear mesh is generally in excess of ninety percent, so in many industrial applications overall efficiency of a geared transmission system is not of major concern. However, in high technology applications such as aerospace the losses arising from friction can become of critical importance. For example, the need for ever greater efficiency of propulsion systems for large transport aircraft has led to serious consideration of geared propeller engines or "geared fans" where a reduction gear drive is used to match the speed of an optimum-efficiency gas turbine to that of an efficient propeller or fan. The powers involved are high, typically 30 MW or more. In these applications reduced gear tooth friction can potentially lead to significantly lower amounts of heat being generated at the tooth contacts and hence less bulky cooling arrangements.

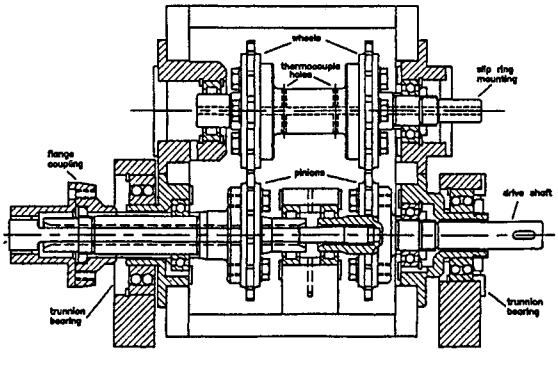


Figure 14. Gear friction rig

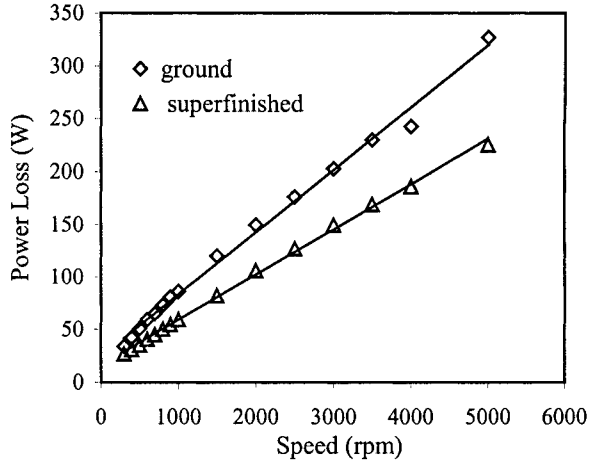


Figure 15. Measured gear mesh friction versus speed for ground and superfinished steel gears.

Measurements of losses in gearing have been made by Britton et al [50] using a special gearbox mounted on trunnion bearings as illustrated in Figure 14. Power is recirculated between the two pairs of meshing gears which are loaded by means of a torque shaft. After calibrating and deducting the churning and bearing losses, the torque required to drive the rig gives the frictional power at the tooth contacts. Figure 15 shows a comparison of tooth friction losses for both conventionally ground and superfinished gears. This work demonstrates the potential energy-saving benefits of improved surface finish in gearing. Work is now underway to develop theoretical models of gear tooth traction based upon both Eyring and limiting shear stress formulations for the lubricant.

7. CONCLUSIONS

This brief review of elastohydrodynamic lubrication effects in gears demonstrates the importance of the real effects of surface roughness including: non-steady behaviour due to the motion of roughness features relative to the nominal contact; the presence of relatively high degrees of

sliding and the consequent frictional heating; non-Hertzian and non-symmetrical geometries. Encouraging progress is being made in the development of new, robust solvers that can deal with the severe conditions of thin films/rough surfaces that characterise real gear tooth contacts. The reality of what has long been called "mixed" lubrication in an otherwise hydrodynamically-lubricated conjunction is beginning to be modelled and seen in these solutions. A second real feature is that of what we have called "in-contact cavitation" due to the presence of deep valley features within the contact, particularly under sliding conditions. The tendency for micropitting to occur in practice at the beginning and end of the meshing cycle, where sliding is greatest, leads us to suggest cavitation as a possible mechanism for erosive wear that may be part of the micropitting explanation. Edge effects in rough contacts are seen to be of crucial importance in allowing the otherwise stiff and robust EHL mechanism to weaken due to sideways leakage of lubricant. This effect is present in rough contacts to a far greater extent than occurs with perfectly smooth surfaces and reinforces the view that it is strongly related to the real phenomenon of scuffing.

8. ACKNOWLEGEMENTS

The authors acknowledge the contribution of Dr K J Sharif and Dr M J A Holmes (Cardiff University) and Mr C Aylott (Newcastle University) in providing some of the figures in the paper. Financial support for the research was provided by the British Gear Association and EPSRC.

NOTATION

b	semi-width of Hertzian line contact	m
E	elastic modulus	Pa
E'	$=0.5\{(1-\nu_1^2)/E_1+(1-\nu_1^2)/E_1\}^{-1}$	Pa
h	film thickness	m
h_{min}	minimum film thickness	Pa
p	pressure	Pa
r_b	base circle radius	m
R	radius of relative curvature of contact	m
R_1,R_2	profile radii of teeth at the contact point	m
s	distance between contact and pitch point	m
$\bar{u}$	mean entraining speed	m/s
u_s	sliding speed	m/s
w'	load per unit length of line contact	N/m
α	pressure coefficient of viscosity	Pa^{-1}
η	viscosity	Pas
η_0	viscosity at zero pressure	Pas
Λ	ratio of film thickness to composite roughness	
ν	Poisson's ratio	
σ	rms surface roughness	m
τ_0	Eyring shear stress	Pa
$\Phi(h)$	fraction of dry contact area whose film thickness *is less than h*	
$\Phi(p)$	fraction of dry contact area whose pressure *exceeds p*	
ψ	pressure angle	rad
$\omega_{1,2}$	Angular velocity of wheel and pinion	rad/s

REFERENCES

1. Price, D J D, *Gears from the Greeks*, Science History Publications, 1975.
2. Dudley, D W, *The Evolution of the Gear Art*, American Gear Manufacturers Association, 1969.
3. Lewis, W, "Investigations of the strength of gear teeth", Proc Engineers Club of Philadelphia, 1893, pp 16-23.
4. Hertz, H, "On the contact of elastic solids", J. reine und angew. Math., Vol 92, 1881, 156-171.
5. Dowson, D and Higginson, G R, 1966, *Elasto-hydrodynamic Lubrication*, Pergamon Press, Oxford.
6. Dyson, A, Naylor, H and Wilson, A R, "The measurement of oil film thickness in elastohydrodynamic contacts", Proc. Instn Mech. Engrs Vol 180(3B), 1965, pp 119-134.
7. Greenwood, J A and Kauzlarich, J J, "Inlet shear heating in elastohydrodynamic lubrication", Trans ASME Jn. Lub. Tech. Vol 95, 1973, pp 417-426.
8. Dyson, A, Evans, H P and Snidle, R W 'Wildhaber Novikov circular arc gears: some properties of relevance to their design', Proc. R. Soc., London, A425 (1989), pp. 341-363.
9. Chittenden, R J, Dowson, D, Dunn, J F and Taylor, C M, "A theoretical analysis of the isothermal elastohydrodynamic lubrication of concentrated contacts", Proc. R. Soc. Lond., Vol A397, 1985, pp 245-269, 271-294.
10. ESDU, 1989, "Film thickness in lubricated Hertzian contacts. Part 2: point contacts"

11. Harris, T A, *Rolling Bearing Analysis*, Second Edition, Wiley, New York, 1984, p 508.
12. O.E.C.D., *Glossary of Terms and Definitions in the Field of Friction, Wear and Lubrication*, 1969, pp 52-53.
13. Berthe, D, Flamand, L, Foucher, D and Godet, M, 1980, "Micro-pitting in Hertzian contacts", Trans ASME, J. Lubr. Tech., Vol 102, pp 478-489.
14. Olver, A V, Spikes, H A and MacPherson, P B, 1986, "Wear in rolling contacts", Wear Vol 112 pp 121-144.
15. Zhou, R S, Cheng, H S and Mura, T, 1989, "Micropitting in rolling and sliding contact under mixed lubrication", Trans ASME, J. of Tribology, Vol 111, pp 605-613.
16. Webster, M N and Norbart, C C J, 1995, "An experimental investigation of micropitting using a roller disk machine", Tribol. Trans, Vol 38, pp 883-895.
17. Shotter, B A, 1981, "Micropitting, its characteristics and implications, in performance and testing of gear oils and transmission fluids", Institute of Petroleum.
18. Webster, M N and Sayles, R S, 1986, "A numerical model for the elastic frictionless contact of real rough surfaces", Trans ASME, Jn Tribology, Vol 108, pp 314-320.
19. Snidle, R W and Evans, H P, 1994, "A simple method of elastic contact simulation", Proc. Instn Mech. Engrs Part J Jn of Engng Tribology Vol 208, pp 291-293.
20. Mayeur, C, Sainsot, P and Flamand, L, 1995, "A numerical elastoplastic model for rough contact", Trans ASME, J. of Tribology, Vol 117, pp 422-429.
21. Kapoor, A and Johnson, K L, 1995, "Plastic ratchetting as a mechanism of erosive wear", Wear Vol 190, pp 86-91.
22. Tyfour, W R and Beynon, J H, 1994, "The effect of rolling direction reversal on fatigue crack morphology and propagation", Tribology International, Vol 27, pp 273-282.
23. Krantz , T L, Alanou, M P, Evans, H P and Snidle, R W, "Surface fatigue lives of case-carburized gears with an improved surface finish", Trans ASME, Journal of Tribology Vol 123, pp 709-716, 2001.
24. Patching, M J, Kweh, C C, Evans, H P and Snidle, R W, 1994, "Conditions for scuffing failure of ground and superfinished steel discs at high sliding speeds using a gas turbine engine oil", Trans ASME Journal of Tribology Vol 117, pp 482-489.
25. Sproul, W. D., Physical vapor deposition tool coatings, Surface & Coatings Technology, Vol 81 (1996), pp. 1-7
26. Eckardt, T., Bewilogua, K., van der Kolk, G., Hurkmans, T., Trinh, T., Fleischer, W., Improving tribological properties of sputtered boron carbide coatings by process modifications, Surface and Coatings Technology, Vol 126 (2000), pp 69-75
27. Hu, T., Steihl, L., Rafaniello, W., Fawcett, T., Hawn, D. D., Mashall, J. G., Rozeveld, S. J., Putzig, C. L., Blackson, J. H., Cermignani, W., Robinson, M. G., Structures and properties of disordered boron carbide coatings generated by magnetron sputtering, Thin Solid Films, Vol 332 (1998), pp. 80-86
28. Evans, H P and Snidle, R W, 1996, "A model for elastohydrodynamic film failure in contacts between rough surfaces having transverse finish", Trans ASME Journal of Tribology, Vol 118, pp 847-857.
29. Goglia, P R, Cusano, C and Conry, T F, 1984, "The effects of surface irregularities on the elastohydrodynamic lubrication of sliding line contacts. Part I - single irregularities; Part II - wavy surfaces", Trans ASME, Jn. of Tribology, Vol. 106, pp 104-112, 113-119.
30. Lubrecht, A A, Ten Napel, W E and Bosma, R, 1988, "The influence of longitudinal and transverse roughness on the elastohydrodynamic lubrication of circular contacts", Trans ASME Jn of Tribology, Vol 110, pp 421-426.
31. Kweh, C C, Evans, H P, and Snidle, R W, 1989, "Microelastohydrodynamic lubrication of an elliptical contact with transverse and three-dimensional roughness", J. of Tribology, Trans ASME, Vol 111, pp 577-584.
32. Elsharkawy, A A and Hamrock, B J, 1991, "Subsurface stresses in micro-EHL line contacts", ASME Journal of Tribology, Vol 113, pp 645-655.
33. Sadeghi, F, 1991, "A comparison of the fluid model's effect on the internal stresses of rough surfaces", ASME Journal of Tribology Vol 113, pp 142-149.
34. Kweh, C C, Patching, M J, Evans, H P, and Snidle, R W, 1991, "Simulation of elastohydrodynamic contacts between rough surfaces", J. of Tribology, Trans ASME, Paper 91-Trib-36.

35. Venner, C H and ten Napel, W E, 1992, "Surface roughness effects in an EHL line contact", ASME Journal of Tribology, Vol 114, pp 616-622.
36. Chang, L and Webster, M N, 1991, "A study of elastohydrodynamic lubrication of rough surfaces", ASME Journal of Tribology, Vol 113, pp 110-115.
37. Chang, L, 1992, "Traction in thermal elastohydrodynamic lubrication of rough surfaces", ASME journal of Tribology, Vol 114, pp 186-191.
38. Chang, L, Webster, M N and Jackson, A, 1993, "On the pressure rippling and roughness deformation in elastohydrodynamic lubrication of rough surfaces", ASME Journal of Tribology, Vol 115, pp 439-444.
39. Ai, X and Cheng, H S, 1993, "A transient EHL analysis for line contacts with measured surface roughness using multigrid technique", ASME Journal of Tribology, Paper No 93-Trib-56.
40. Chang, L and Zhao, W, 1995, "Fundamental differences between Newtonian and non-Newtonian micro-EHL results", ASME Journal of Tribology, Vol 117, pp 29-35.
41. Zhu, D. and Hu, Y Z., 1999, "The study of transition from full film elastohydrodynamic to mixed and boundary lubrication" Proc STLE/ASME H.S. Cheng *Tribology Surveillance*, 150-156.
42. Hu, Y Z, and Zhu, D., 2000, "A full numerical solution to the mixed lubrication in point contacts," ASME Journal of Tribology, Vol 122, 1-9.
43. Zhu, D. and Hu, Y Z., 2000, "Effects of rough surface topography and orientation on the EHD and mixed lubrication characteristics" Proceedings International Tribology Conference, Nagasaki, 625-630.
44. Hughes, T.G., Elcoate, C.D. & Evans, H.P. Coupled solution of the Elastohydrodynamic line contact problem using a differential deflection method. Proc. Instn Mech. Engrs Part C Jn of Mechanical Engineering Science, Vol 214, pp 585 598, 2000.
45. Elcoate, C D, Hughes, T G, Evans, H P., and Snidle, R W., 2001, "Transient elastohydrodynamic analysis using a novel coupled differential deflection method" Proc. Instn. Mech. Engrs Part J, Jn of Engng Tribology , Vol 215, pp 319-337.
46. Tao, J, Hughes, T G, Evans, H P, & Snidle, R W., 2002, "Elastohydrodynamic response of transverse ground gear teeth". To appear Proc. 28th Leeds-Lyon Symp. on tribology, Elsevier, Amsterdam.
47. Holmes, M.J. *Transient analysis of the point contact elastohydrodynamic lubrication problem using coupled solution methods.* PhD thesis, University of Wales, 2002.
48. Evans, H P and Snidle, R W, 'Wildhaber-Novikov circular arc gears: Elastohydrodynamics', Trans ASME Jn of Tribology, Vol 115, pp 487-492, 1993.
49. Sharif, K, Kong, S, Evans, H P and Snidle, R W, "Contact and elastohydrodynamic analysis of worm gears: Part 2, results", Proc. Instn Mech. Engrs, Part C: Journal of Mechanical Engineering Science Vol 215 (2001), pp 831-846.
50. Britton, R D, Elcoate, C D, Alanou, M P, Evans, H P and Snidle, R W, "Effect of surface finish on gear tooth friction", Trans ASME Jn. of Tribology, Vol 122, pp 354-360, 2000.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribology of Spline Couplings

A. V. Olver and S. Medina

Tribology Section, Department of Mechanical Engineering, Imperial College of Science, Technology and Medicine, London SW7 2BX, England

Toothed spline couplings are ubiquitous in mechanical transmission systems and have been studied experimentally for many years. Recently it has become possible to construct useful three-dimensional numerical models of complete couplings and this has led to a number of surprising findings which cast new light on the mechanics and durability of the contacting spline teeth. The analysis presented includes misalignment, elasticity and friction. Among the new findings are the existence of discontinuous tooth loadings, of unsymmetrical pressure-slip behaviour and of a "toppled" regime in which the two halves of the coupling are suddenly displaced through their mutual clearance. The results of the study are used to reassess wear testing and to propose a new, more logical approach to coupling design.

NOMENCLATURE

A	Area of contact
b	Facewidth of male coupling
b^*	Dimensionless facewidth, b / m
c	Axial co-ordinate of centre of pressure
C_{max}	Maximum relative displacement of centre of pressure, $[c / b]_{max}$
d	Pitch diameter, $d = zm$
E	Young's Modulus
E^*	Dimensionless elasticity/torque, $(Em^3) / T$
m	Module
M	Bending moment
N	Number of cycles or revolutions
p	Pressure
P	Tooth load
P_{max}	Maximum tooth load
P_{av}	Average tooth load during shaft revolution
R	Pitch radius, $R = zm/2$
s	Slip displacement
T	Torque
W	Normal load in a general contact
z	Number of teeth
z_w	Local depth of wear
Z_w	Wear depth parameter (Eq. 16)
α	Pressure angle
β	Friction angle (Fig. 1)
μ	Limiting coefficient of friction
θ	Misalignment angle

1. INTRODUCTION

Toothed spline couplings are ubiquitous in mechanical transmission systems and have been studied experimentally for many years with particular attention being given to wear behaviour. Examples include the experiments of Ku and co-workers [1], Newley [2] and Baker [3].

Analysis of spline contacts has also been of occasional interest. Perhaps best-known is the paper by Dudley, [4] who concentrated on the prediction of strength. Buckingham [5] suggested that when misalignment was present contact would tend shift onto two opposite teeth and toward the edge or end of the teeth. He identified a condition "cramping" in which the teeth would exhaust their mutual backlash and in

TABLE 1 - DIMENSIONLESS GROUPS

	Symbol	*Description*	*Dimensionless Group*
Input Parameters	E^*	Dimensionless elasticity / torque	$\frac{Em^3}{T}$
	b^*	Dimensionless facewidth	b / m
	z	Number of teeth	
	α	Pressure angle	
	μ	Friction coefficient	
	θ	Misalignment angle	
Output Parameters	C_{max}	Maximum deviation of centre of pressure	
	P_{max}/P_{av}	Maximum tooth load	
	Z_w	Wear depth parameter	
	M/T	Dimensionless moment	

E = elastic modulus; m = module; T = torque; b = facewidth of male coupling; C_{max} = maximum relative axial displacement of centre of pressure from tooth centre.

TABLE 2. PRESSURE REGIMES IN SPLINE COUPLINGS (After [7]).

REGIME	DEFINITION
Uniform	Contact over the whole tooth surface. Some variation of pressure during rotation
Cyclic	Part of each tooth surface becomes unloaded during rotation
Discontinuous	Entire tooth becomes unloaded during rotation
Toppled	Contact outside normal tooth flanks (e.g. on reverse face)

↓ Increasing misalignment or bending moment; decreasing torque

which reverse face contact would occur. Neale [6] also recognised that tooth contact would shift to the opposite ends of opposed teeth and used this to estimate the maximum bending moment that a spline could exert. These earlier analyses were unable to account for the effects of elasticity of the components in distributing the applied loads in either the axial or the profile direction and so they were not able to predict tooth stresses or contact conditions accurately.

However, recently it has become possible to construct three-dimensional numerical models of complete couplings [7-9]. These have been used to optimise designs for specific applications and to make predictions of pressure and slip in the presence of applied torque and moments or misalignments.

In an earlier paper [7], we presented the results of a study of a generic involute spline coupling, as a function of a number of dimensionless quantities which control the contact behaviour. The results showed that the behaviour during rotation of the coupling could be classified according to the regime of friction (cyclic stick-slip or gross slip) and according to the manner in which the distribution of the contact pressure between male and female teeth varied during rotation of the shaft (uniform, cyclic, discontinuous or toppled). Table 1 shows the dimensionless quantities which control the spline

behaviour and Table 2 gives definitions of the regimes of pressure. Use of the non-dimensional groups allows the identification of different configurations with similar behaviour.

A significant finding from the earlier work was that applied misalignment or bending moment—even with a steady applied torque—caused the contact to oscillate up and down the teeth in an axial direction during rotation. Because of this, there were both cyclic stresses, which are related to fatigue and a complex cycle of pressure and slip, which is related to the occurrence of fretting wear damage.

In the present paper, we present a simple rigid body model of a spline coupling and use it to determine normal and tangential displacements in the contact region when misalignment is imposed. We also consider the overall equilibrium and stability of the coupling. These results are then compared to those from a boundary element model which includes elasticity and stick-slip friction in the contact. Wear depth predictions are also made. A locked spline, in which the shaft axial location is controlled by a nut and a shoulder is also examined.

2. RIGID ANALYSIS

2.1 Introduction

Simple models of spline tooth contact have been presented by Buckingham [5] and Neale [6]. These are based effectively upon the complete slip theory of contact [10] in which the elastic deformations are considered to be small compared to the slip displacements. The externally imposed displacements due to misalignment of the male spline with respect to the female are used to find local (rigid-body) displacements within the contact. Components of the relative displacement in the plane of the tooth contact provide a first approximation to magnitude and direction of the local slip. In addition it is expected that the contact will occur within the region in which there is an interference (i.e. a negative value of the relative normal displacement) between the mating components.

2.2 Misalignment

Consider the situation illustrated in Figure 1. The global axes x, y, z are chosen so that the normal to the tooth surface at P is in the y direction, while z is directed along the shaft axis. The origin is at the centre of the coupling. At time, t, the male coupling is displaced through an angle, θ, due to a remotely imposed misalignment, around an axis along the vector $\boldsymbol{u}$ which is in the xy plane forming an angle αt with the x axis. The misalignment rotates with respect to the stationary coupling at an angular speed, ω.

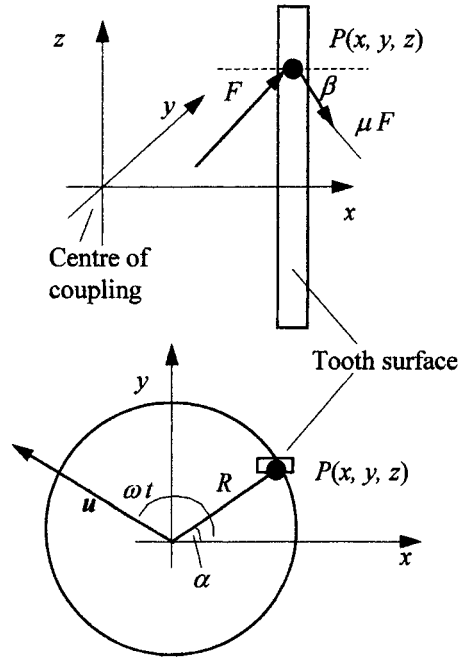


Figure 1. Global co-ordinates of a point, P on the male tooth surface. The axes are chosen so that the tooth surface is normal to the y axis and the origin is in the centre of the coupling with the z axis aligned with the shaft.

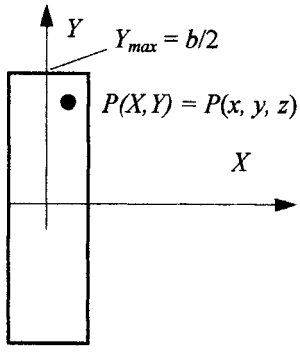


Figure 2. Local co-ordinate axes. The Y axis is parallel to the shaft and the origin is in the centre of the contacting region of the tooth which extends to $Y = \pm b/2$ at each end of the facewidth.

Local axes on the contacting tooth surface are shown in Figure 2. The point P is displaced by the rigid body misalignment of the male coupling. The X and Y components, u and v of the displacement vector, assuming the angle, θ, to be small, are [11]

$$u = Y\theta\sin\omega t$$
$$v = R\theta\sin\alpha\cos\omega t - (R\cos\alpha + X)\theta\sin\omega t \quad (1)$$

where $R = zm/2$ is the pitch radius of the coupling and α is the local pressure angle. If the maximum value of X (i.e. of the tooth contact semi-width in the profile direction) is small compared to the pitch radius, R, then Eq. (1) reduces to:

$$u = Y\theta\sin\omega t$$
$$v = R\theta\sin(\omega t + \pi - \alpha) \quad (2)$$

Thus the amplitude of the slip path is sinusoidal in t in both directions with the X and Y components being (π - α) out of phase. The X and Y amplitudes are $Y\theta$ and $R\theta$ respectively.

The relative *normal* displacement (the gap) is:

$$w = -Y\theta\cos\omega t \quad (3)$$

This is only negative when $-\pi/2 < \omega t < \pi/2$ for $Y > 1$, so contact is *intermittent*, oscillating from top to bottom of the tooth during rotation of the shaft. The maximum interference occurs at the top of the tooth ($Y = b/2$) when $\omega t = \pi$, as may easily be discerned from figure 1.

Figure 3 shows the results of the rigid body analysis. The slip path ABCD is elliptical and has maximum interference at point A. There is no contact possible along the path BCD because the surfaces are separated. It is noted that the displacements depend only on the orientation of the tooth with respect to its local pressure angle and of the instantaneous misalignment vector. Therefore, all teeth in the coupling experience the same displacement history as the shaft rotates under a constant misalignment.

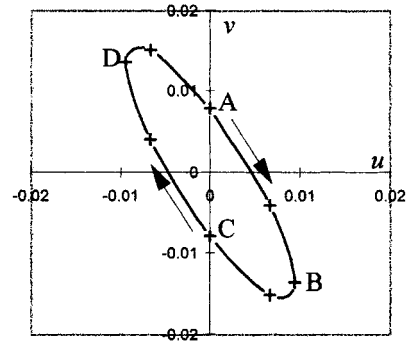


Figure 3. Locus of displacements, u and v traced out by a point on the female tooth relative to a point on the male tooth, for a location having a positive Y co-ordinate, according to the rigid model. Displacements are shown normalised by the module for $\theta = 0.1$ degrees, $b*/z = 0.6$.

2.3 Moments

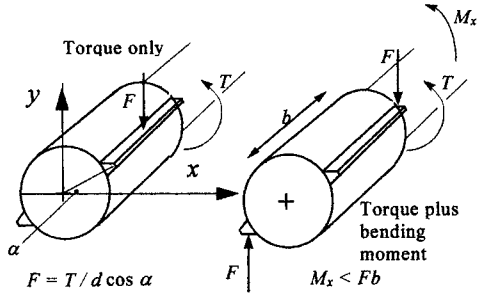


Figure 4. Loads on a two-tooth frictionless male spline shaft.

We now consider equilibrium between the contact forces and the external loads, consisting of a torque, T and a bending moment, M. For simplicity, following Neale [6] the coupling is first considered to have only two pairs of mating teeth at angles α and $\pi - \alpha$ to the x axis. If $M = 0$ and neglecting friction, then the forces on each of the two teeth are $F = T/(d\cos\alpha)$ and will act at the same axial position (the same Y co-ordinate) on each tooth pair (Fig. 4a). However, introduction of a bending moment M_x requires the tooth forces to be displaced axially along the tooth in opposite directions to maintain equilibrium (Fig. 4b). A maximum bending moment therefore occurs when each tooth pair is loaded at opposite ends.

$$\hat{M}_x = \frac{bT}{d\cos\alpha} \quad \text{or} \quad \frac{\hat{M}_x}{T} = \frac{b}{d\cos\alpha} \qquad (4)$$

where $d = 2R$ is the pitch diameter. Now considering a sliding friction force μF directed along the slip path shown in Fig. 3, it can be seen that the Y component of the friction force exerts an additional moment, $\mu Fd\sin\alpha\sin\beta$ about the x axis. However, equilibrium about the z axis now yields:

$$M_z = T = Fd\cos\alpha + \mu F\cos\beta\sin\alpha \qquad (5)$$

so the total maximum bending moment is now given by:

$$\frac{\hat{M}_x}{T} = \frac{b + \mu d\sin\alpha\sin\beta}{d(\cos\alpha - \mu\sin\alpha\cos\beta)} \qquad (6)$$

where β is the angle of the slip path at the point A to the X axis which can be found by differentiation of Eq. 2 with respect to t, with the result:

$$\tan\beta = -\left(\frac{\partial v/\partial t}{\partial u/\partial t}\right)_{t=0,Y=b/2} = \frac{d}{b}\cos\alpha \qquad (7)$$

Equation (5) leads to a maximum value of $\hat{M}_x/T$ of 0.93 for a coupling with $b/d = 0.6$ and $\mu = 0.3$, or 0.63 for $\mu = 0$. A similar treatment by Neale [6], in which it was arbitrarily assumed that the friction was in the axial direction, yields a somewhat lower maximum moment.

A moment acting around the y axis must also be applied in order to maintain sliding friction along the contact path:

$$\frac{\hat{M}_y}{T} = \frac{\mu(\cos\alpha\sin\beta + (b/d)\cos\beta)}{(\cos\alpha - \mu\sin\alpha\cos\beta)} \qquad (8)$$

2.4 Discussion

If externally applied moments exceed these values then the coupling cannot be in equilibrium with normal and sliding friction forces on the active tooth surface. This means that if the moment is gradually increased to greater than the maximum value given by Eq. (7) we should expect a sudden onset of instability, probably leading to a discontinuous displacement of the coupling through its internal clearance and subsequent contact on the reverse faces or perhaps on the crests of the teeth. The behaviour (termed "toppling" in earlier papers) is familiar from handling of splines at low torque, for example during assembly. The phenomenon is sometimes termed "self-centering" behaviour because the restoration of stability by the imposition of a larger torque restores the male

and female spline to an ostensibly collinear position. However, it is clear from the foregoing that toppling results, in essence, from loss of equilibrium rather than any lack of centrality.

Here we have assumed that with μ positive, the friction is in the direction required for clockwise rotation of the vector $\boldsymbol{u}$ corresponding to the male spline being the driver. If the rotation (or the torque) is reversed, the friction now reduces the maximum moment, suggesting that the toppling instability is more likely - especially for splines of low facewidth.

The moments found here are maxima related to the instability during normal rotation of the shaft. They do not reveal the effect of friction, geometry or other parameters on the distribution of tooth loading at lower applied moments. Here the load will be distributed more centrally on the tooth surface and over more than two teeth. Nevertheless, slip and consequent wear is still a possibility. These factors require a more detailed analysis.

3. BEM MODEL

3.1 Introduction

When loads are applied to a real coupling, elastic displacements are inevitable. Furthermore, as the maximum possible misalignment angles, θ, are very small, it is likely that elastic and slip displacements will be of similar magnitude, casting doubt upon the validity of the above analysis. In order to carry out an elastic analysis it is necessary to model the complete coupling in three dimensions because both torsional, bending and contact deflections can occur and because there is no cyclic symmetry owing to the misalignment. Appreciable friction loading can also be present and this is also incorporated.

3.2 Description of Model

The boundary element method [12] was chosen for the analysis, using a software package specially developed by Adey and co-workers for spline analysis [13]. Details of the procedure are given in earlier papers [7,9].

The model consists of a basic involute spline coupling, with a short section of shaft adjoined to the toothed section as shown in figure 5a.

The element mesh of the male teeth can be seen in the detail (5b). Each contacting tooth surface was discretised as 2×10 elements, leading to 5×21 = 105 surface nodes. This was found to be sufficient to allow a comparison between different geometries and loading conditions whilst minimising the time taken for analysis.

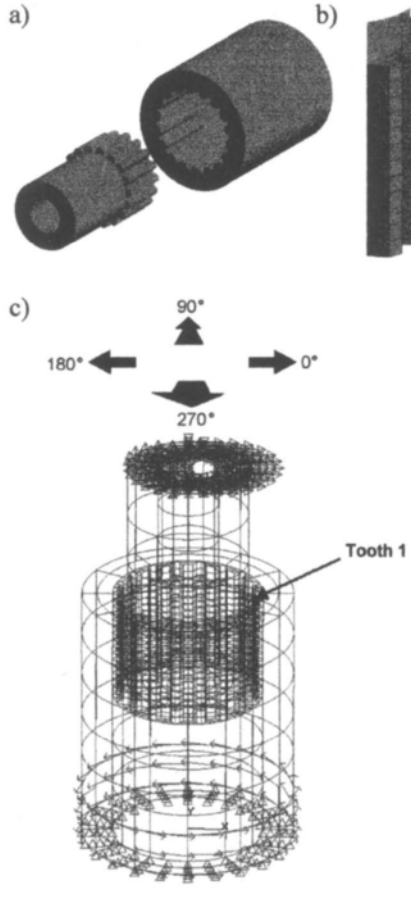


Figure 5 BEM model of spline coupling.

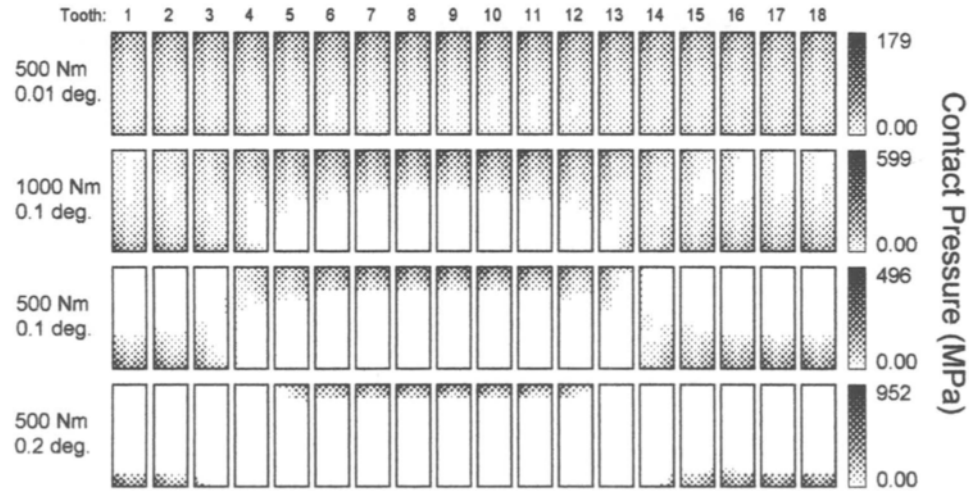


Figure 6. The effect on the distribution of pressure of changing the torque and the misalignment in a simple, 18-tooth, 2.5 module spline. The density of shading shows the local pressure on each tooth

Different mesh geometries outside of the contact zones had little effect on results. Within the contact zones, the singularity at the periphery of the contact meant that mesh size affected the maximum pressures significantly, although the form of the contact regime remained similar. Here, the use of a consistent mesh size has permitted a more reliable comparison of maximum contact pressures. The additional computational requirements prohibited the use of finer mesh densities – moreover, these would not have eliminated the problem of singularity.

The boundary and loading conditions are shown in figure 5c. The torque is applied through the lower outside edge of the female spline shaft. A misalignment angle is imposed on the coupling by displacing each point at the top of the male spline shaft (at a position 20 *m* from the end of the teeth) through the required angle about a horizontal axis through the centre of the coupling. The direction of the imposed misalignment was varied in steps of 90 degrees, giving five load conditions which are applied sequentially to simulate a complete 360° rotation of the shafts. This allowed the complete slip path and contact pressure history to be determined. The initial gap between contacting teeth was set at zero for most analyses, but by varying the initial (unloaded) gap over the tooth surface and for each tooth the effects of manufacturing errors has also been assessed.

A range of spline geometries and operating conditions have been analysed systematically using the techniques described, and the contact conditions for each case compared. The pressure and slip information has also been used to make a simplified assessment of the wear which might occur for each situation.

3.3 Results: Pressure

Typical distributions of pressure acting on each of the teeth during the rotation of the coupling have been given in earlier papers [7, 9] and examples are shown in figure 6. The results confirm that the centre of pressure oscillates up and down the face of the coupling

as the shaft rotates. As the misalignment angle, θ, is increased, not only does the pressure become more concentrated at the tips of the teeth but the maximum tooth load increases due to a concentration of load on teeth in the position shown on the right of figure 4; that is where the misalignment causes the greatest normal displacement. This is shown in figure 7.

The model has also been used to determine the bending moments M_x, for a range of geometries and conditions. All these were within the limit given by equations (6) and (7); for example, the largest value for $b/d = 0.6$ and $\mu = 0.3$ was 0.67 compared to the rigid body limit of 0.93.

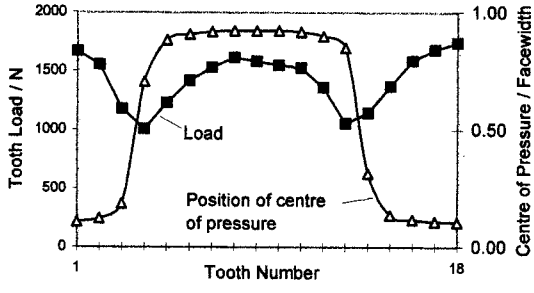


Figure 7. The load on each mating tooth pair and the axial position of the centre of pressure. Conditions as in Fig. 6, 500 N m, 0.1 deg.

Approximate equations governing the main parameters of the pressure distribution and bending moment are given below:

$$P_{max}/P_{av} \approx 1 + 0.45\,(E^*\theta)^{1.3} \qquad (9)$$

(standard deviation 12% of computational results)

$$C_{max} \approx \frac{1}{2}\left(1 - \frac{1}{1 + E^*\theta b^*(z/18)^3}\right) \quad (\pm 4\%) \quad (10)$$

$$M/T \approx \frac{(C_{max} - 0.18)(b^*)^{1/2}}{1.4(z/18)} \quad (\pm 15\%) \quad (11)$$

θ is in degrees and Eq. 11 refers to $\mu = 0.3$ only.

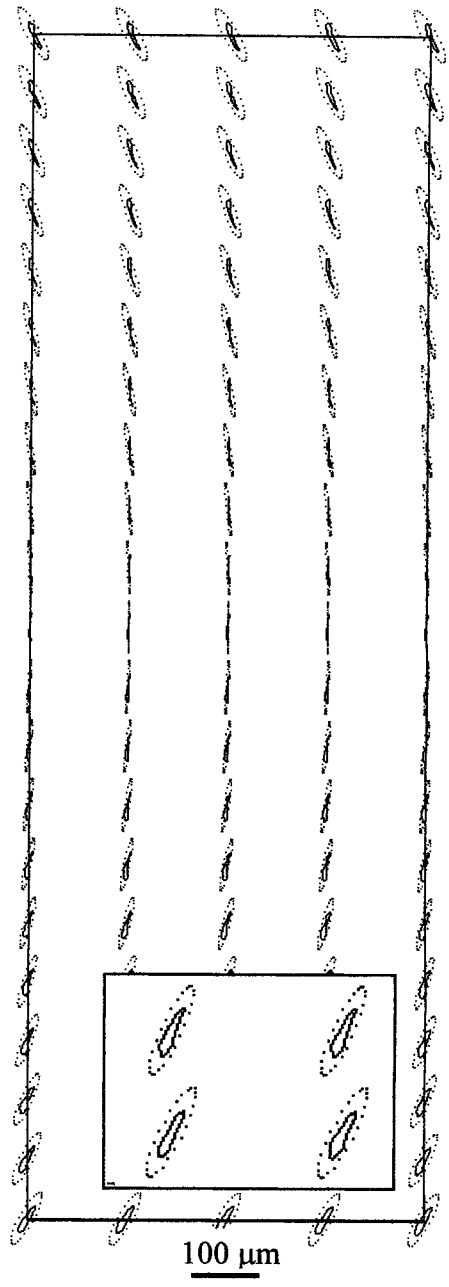


Figure 8. Locus of displacement traced out by each point on the female tooth relative to a point on the male tooth. Comparison of BEM (solid) and complete slip theory (dotted) calculations. Inset shows magnified view of region at lower right.

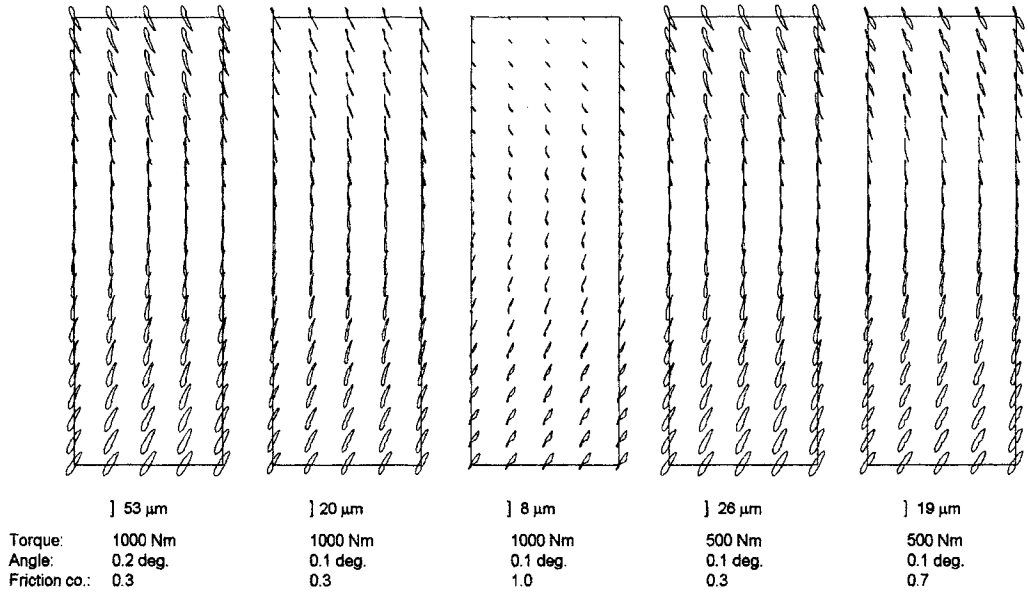


Figure 9 Effect of the coefficient of friction on the slip path.

3.4 Results: Slip

The locus of the relative tangential displacement calculated from the computational results is shown in figure 8. The rigid body predictions (equations (2)) are also shown. This relative displacement takes place both when the two surfaces are in contact (and the pressure is therefore non-zero) and also during the period when the surfaces are separated. The former comprises sliding displacement ("slip") and may contribute to wear, whereas the latter has no such effect.

Comparison of the two results shows that the effect of friction and elasticity is to reduce the amplitude of slip resulting in a much reduced locus contained entirely within that found by the equivalent rigid analysis. Consideration of the instantaneous local pressure again shows a pattern similar to that of the overlap found from the rigid analysis. More than half the displacement path occurs under *unloaded* conditions. Increasing the coefficient of friction or reducing the misalignment reduces the slip amplitude (Figure 9). At sufficiently high coefficients of friction, the slip declines to a very small value as stick behaviour predominates during the loaded part of the cycle.

3.5 Results: Wear

The results may be used to make a simple estimate of how different designs and operating conditions affect wear. In the sliding wear of metals it is usually found [e.g. Archard and Hirst, 14] that wear volume is proportional to the load and the sliding distance.

$$V = k\,W\,s \qquad (12)$$

If this applies to each small increment of area, dA, on the tooth surface, (12) may be expressed as:

$$\frac{dV}{dA} = k\frac{dW}{dA}s \quad \text{or} \quad z_w = kps \qquad (13)$$

where z_w is the local depth of wear and p is the pressure.

Provided that wear does not occur during stick and providing that the wear constant k, is not affected by the changing conditions during the cycle Eq. 13 may be expressed in differential form:

$$dz_w = \left(\frac{\partial z_w}{\partial s}\right)_p ds = kpds \qquad (14)$$

Over N cycles of periodically varying pressure and slip, the total wear depth would therefore be:

$$z_w = Nk\oint pds \qquad (15)$$

where the integral is performed over a single cycle of slip and pressure.

In the present paper, the values of the integral $\oint pds$ were determined for each node on the tooth surface and the maximum of these found for each run. For the parameter study, the results were non-dimensionalised in the form of a wear depth parameter (Table 2):

$$Z_w = \frac{\left[\oint pds\right]_{max}}{Em} \qquad (16)$$

The actual wear depth associated with the worst worn area of surface is therefore given by:

$$z_w = kN Z_w Em \qquad (17)$$

where N is the number of revolutions of the shaft and k is the (dimensional) Archard wear coefficient. Values of k for metals are given in [15]; for the fretting regime, they range from $(3 \times 10^{-7})/H$ to $(3 \times 10^{-4})/H$ where H is the hardness, depending on lubrication state and material compatibility.

Figure 10 shows how the wear depth parameter, Z_w, varies for a typical coupling in response to the elasticity parameter and the misalignment angle. Also shown is the approximate toppling boundary.

An approximate formula for the wear depth derived from the BEM results is:

$$Z_w \approx 0.0109 \frac{\theta^{1.9}(z/18)}{(b^*)^{1/2}} \quad (\pm 19\%) \qquad (18)$$

This applies to $\mu = 0.3$ only.

3.6 Discussion

The BEM model permits a more rigorous consideration of the contact conditions present in the spline coupling but, in many respects, the findings are consistent with the expectations of the rigid model presented in section 2.

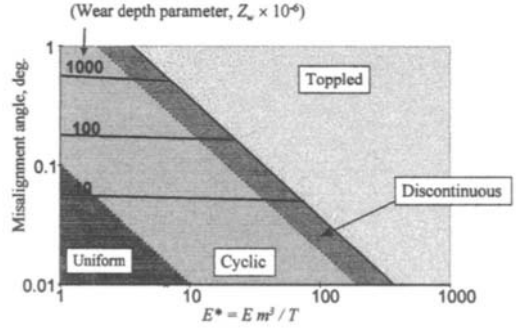


Figure 10. Effect of torque and misalignment on wear parameter.

In particular, both analyses show the existence of a maximum applied moment, above which stability is lost and "toppling" occurs; both show slip paths of similar shape and the BEM pressure distribution at high misalignments is intermittent, and hence qualitatively consistent with the regions of overlap of the surfaces in the rigid model.

However, the BEM treatment has additionally permitted the calculation of a simple wear depth parameter which can be used to compare designs. In practice a complication is that many couplings operate under mixed boundary conditions where neither the misalignment angle nor the applied moment is known in advance. Simultaneous solution of equations (12) and (18) is then required.

In an earlier study it was shown that the distribution of wear depth parameter across the tooth correlated with the distribution of

damage on a test spline tooth [9]. However, the damage recorded in most spline tests is very complex due to the localised effects of debris. This means that quantitative wear measurements are difficult so it cannot be claimed that a reliable predictive tool for spline wear has been achieved.

A further level of complexity arises because fretting wear tests show wear constants, k, which appear to *vary* with slip amplitude. However, it seems possible that, in fact, k is independent of the true slip amplitude but that workers may have erroneously neglected the effects of friction and elasticity on the actual slip within the contact. This would give the trend shown by Vingsbo and Soderberg [16] and by Baker and Olver [17]. Further experiments would be needed to resolve this point.

4. LOCKED SPLINES

4.1 BEM Model

As an example of the application of the technique to a more realistic spline coupling, we now consider the situation shown in Fig. 11. Here, the spline coupling connects two shafts and is axially located against a shoulder by a nut. This situation (a "locked" spline) is common in rotating machinery. In the model, the nut is represented by a solid extension to the male shaft so that only the clamping face is modelled explicitly.

The presence of a locking nut will have three major consequences: the relative displacement between the components and thus the slip at the contacts may be altered; the shafts and teeth will be in compression or tension, leading to a change in deformation; and, the friction at the locking mechanism may provide an alternative site for torque transmission thus reducing the loading of the teeth. It is indeed possible that, with sufficient friction from a high locking force, the entire torque could be transmitted through this means rather than through the teeth.

The boundary conditions applied to the model were similar to those of Fig. 5. Additional contact interfaces were specified at the appropriate locking points. A torque of 500 N m was applied with a misalignment of 0.05° between the bottom of the female spline component (the outer rim) and the top of the male spline. Different interface boundary conditions were used to represent different locking conditions (such as nut tightening torque).

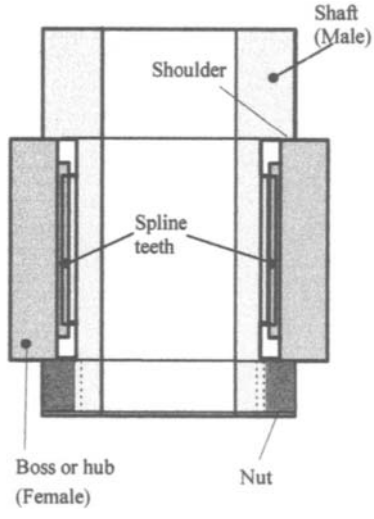


Figure 11. Locked spline coupling.

4.2 Results

With no locking action, modelled as a large initial gap at the interface, the pressure distribution is as shown in Figure 12a, and is similar to previous results. With a large locking torque (modelled as an initial interference of 100 μm) and frictionless contact, the pressure distribution is transformed to that of Fig. 12(b), with contact at the extreme edges only. This is believed to be caused by a radial displacement of the female component ("bowing") due to the large axial compression, leading to separation of the central portion of the tooth surface.

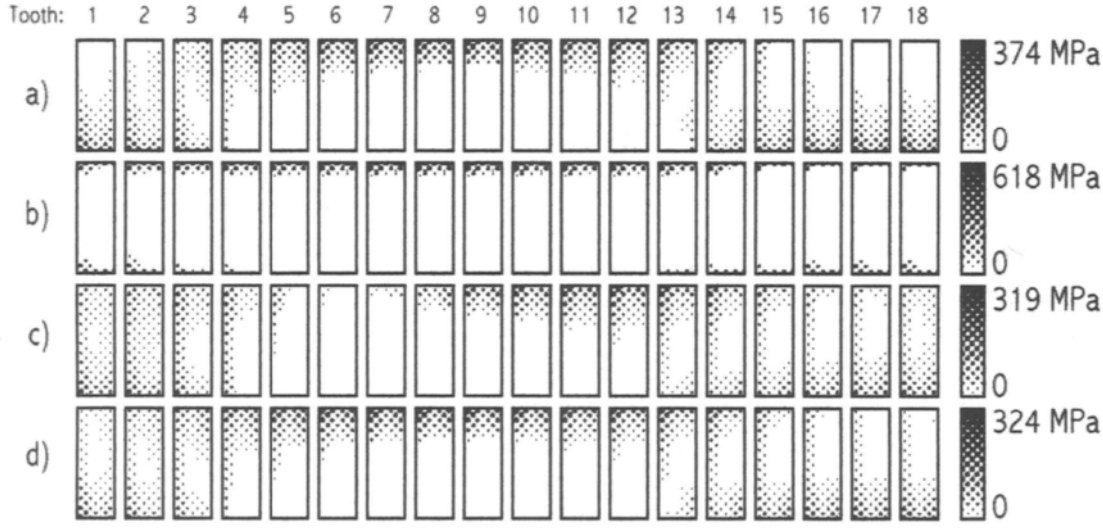


Figure 12 - Pressure distributions for locked splines (a) unlocked (b) locked (large locking load) (c) locked (small locking load / with friction) (d) locked (small locking load / frictionless)

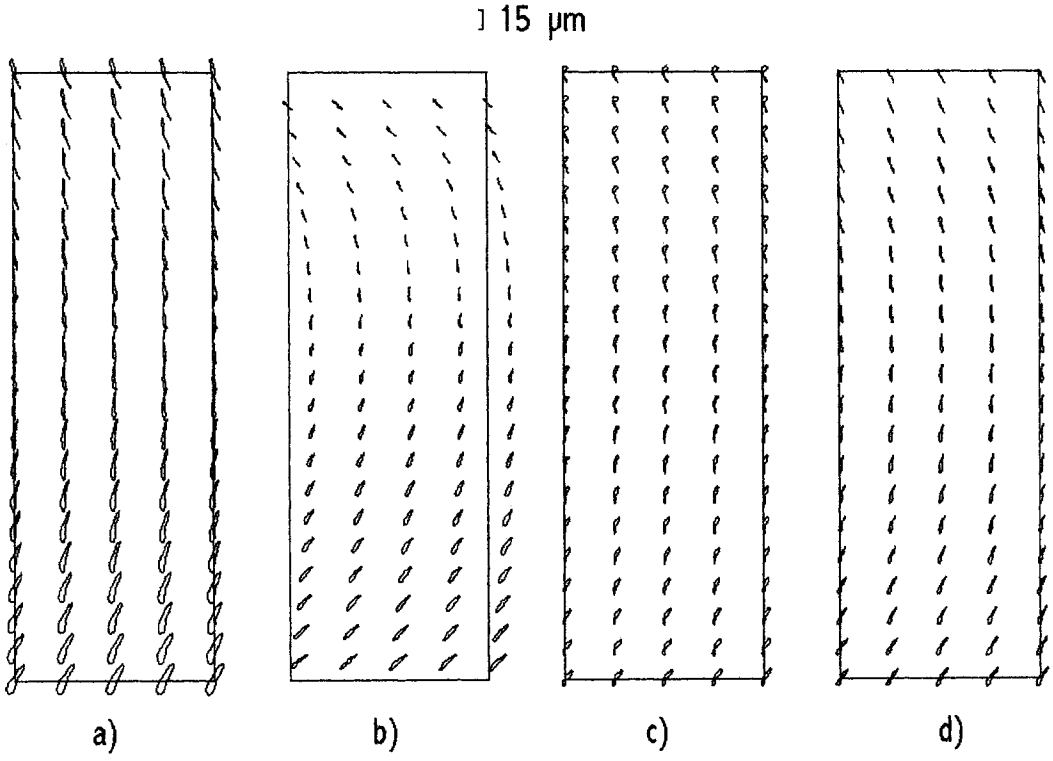


Figure 13 - Slip paths for locked splines (a) unlocked (b) locked (large locking load) (c) locked (small locking load / with friction) (d) locked (small locking load / frictionless)

A second analysis, in which friction was included, produced a situation where the entire torque was transmitted through friction at the locking interfaces, and the teeth had no contact load. This is clearly an extreme example, but demonstrates the action of torque transmission at the nut; the tribology and wear of the nut and locking interfaces then become of more importance than that of the teeth.

A more practical example of a locked spline pressure distribution is shown in Figure 12 (c) and (d), which represents a lower locking torque modelled by contact at the interface but no interference (initial gap of 0.0 μm). The former is with frictional contact; the latter is with a frictionless nut/shoulder. The pressure distribution is altered once again, but to a lesser extent.

The effect of a locking mechanism on the slip between the teeth is of greater significance, particularly for reducing wear. The slip paths for the various analyses are shown in Figure 13. (Note that the tooth outline is not to scale; the deformation is not as great as may appear in (b).) In all cases the slip amplitude is reduced. The altered position of the slip path centres of (b) indicate the tooth deformation caused by the large axial compression – the *average* relative displacement at each node has moved away from the unloaded position.

The reduction in slip has contributed to a reduction in maximum wear parameter of 50% between the unlocked spline and the locked spline (c). The wear parameter of (b) was slightly increased due to the greater contact pressures.

4.3 Discussion

Locked splines are perhaps more common than free splines in real machinery, a classical example being the attachment of a wheel or gear to a shaft. It is therefore of some interest that the contact conditions can be sensitive to clamping forces exerted by the nut. This clearly affects both the expected wear behaviour and the strategy to be adopted for optimum tooth correction, such as that reported in earlier work [9]. Nevertheless, the finding that the durability is improved by around one half for a locked spline is in agreement with common engineering folklore!

5. CONCLUSIONS

The analysis of spline couplings subjected to applied torque and misalignment or bending moment has been carried out by two methods: classical complete slip theory and an elastostatic model with stick-slip friction based on the boundary element method.

Both techniques show that tooth loadings are cyclic, and suggest the presence of a maximum bending moment beyond which an instability, "toppling", occurs. The BEM method also allows the calculation of a wear depth parameter which can be used to evaluate the durability of splines under defined conditions.

Analysis of a 'locked' spline in which the male shaft has a shoulder and a nut, suggests the durability is improved by around one half compared to an unlocked spline.

ACKNOWLEDGEMENT

The financial support of the Engineering and Physical Sciences Research Council and of Westland Helicopters Ltd, Yeovil, UK is gratefully acknowledged.

REFERENCES

1. Weatherford, W D, Valtierra, M L & Ku, P M, 'Mechanisms of Wear in Misaligned Splines', Trans. Am. Soc. Mech. Engrs J. Lub. Tech. 90F (1968) 42-50.

2. Newley, R A The Mechanisms of Fretting Wear of Misaligned Splines in the Presence of a Lubricant, PhD Thesis, University of London, 1978.

3. Olver, A V, Medina, S, Baker, R F, and Davies, D P, "The Application of Laboratory Fretting Data to Wear of Splined Couplings", I. Mech. E. Seminar on Advanced Materials and Processes for Aerospace Transmission Systems, Solihull, England, 21 May 1997, published I Mech E 1997.

4. Dudley, D W, 'When Splines Need Stress Control', Production Engrg, 28 (1957) pp 56-59

5. Buckingham, E, How to evaluate the effects of spline misalignments, Machinery, Vol. 67 119-128 and Vol. 68 123-130 (1961).

6. Neale M J, Needham, P and Horrell, R, Couplings and Shaft Alignment, MEP Ltd, Bury St Edmonds, UK, 1991.

7. Medina, S and Olver, A V, Regimes of Contact in Spline Couplings, Trans. Am. Soc. Mech Engrs., J. Tribology 124 (2002) 351-357.

8. Hyde, T R, Leen, S B, McColl, I R, Williams, E J and Taylor, J W, A Representative Model for Life Assessment of Spline Couplings, in Daizhong Su (ed), Proceedings of the International Conference on Gearing and Transmission Systems, Nottingham, UK, July 2000, Professional Engineering Publishing, Bury St Edmunds, UK.

9. Medina, S and Olver, A V, An Analysis of Misaligned Spline Couplings, accepted for Proc. Instn. Mech Engrs, London, Part J, J. Engineering Tribology, March 2002.

10. Johnson, K L, Contact Mechanics, Cambridge University Press, 1985, p259ff.

11. Medina, S, Fretting Wear of Misaligned Spline couplings, PhD Thesis, University of London, 2002.

12. Brebbia, C A, Boundary Element Method for Engineers, Pentech Press, 1980.

13. Adey, R A, Baynham, J, and Taylor, J W, Development of Analysis Tools for Spline Couplings, Proc. Instn Mech Engrs (London), 214, part G (2000) 347-357.

14. Archard, J F, and Hirst, W, "The Wear of Metals under Unlubricated Conditions", Proc., Roy. Soc., Series A, Vol. 236, 1956, p 397.

15. Rabinowicz, E, Wear Coefficients for Metals, in Peterson, M D, and Winer, W O, (Eds) Wear Control Handbook, Am. Soc. Mech. Engrs, 1980, 475-506.

16. Vingsbo, O and Soderberg, S, On Fretting Maps, Wear 20 (1988) 131-147.

17. Baker, R F and Olver, A V, 'Direct Observations of Fretting Wear of Steel', Wear 203-204 (1997) p 425.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)

© 2003 Elsevier B.V. All rights reserved.

Prediction of the Friction Coefficient Between Gear Teeth Using a Mixed Film Lubrication Model

Castro, J. [a], Sottomayor, A. [a], Campos, A. [a] and Seabra, J. [b]

[a] Departamento de Engenharia Mecânica, Instituto Superior de Engenharia, Instituto Politécnico do Porto, Portugal.

[b] Departamento de Engenharia Mecânica e Gestão Industrial, Faculdade de Engenharia, Universidade do Porto, Portugal

The experimental results found in literature for the friction coefficient between gear teeth cannot be predicted from the lubricant traction curves obtained in twin disc machine tests under full film lubrication regime.

A numerical model was developed to simulate the linear contact between rough surfaces in mixed film lubrication. The model is based on the pressure distributions of a full film EHD rough contact and of the corresponding dry contact, from which the pressure distribution in the mixed film lubrication regime is obtained using a *load share* function. This *load share* function is based on failure probability curves for lubricated gears, depending on the specific lubricant film thickness and on the gear pitch line velocity.

The mixed film pressure distribution is used to evaluate the corresponding contact deformed geometry, the lubricant film shear stresses, the lubricant and surfaces temperatures and the contact friction coefficient. The model admits the non-Newtonian behaviour of the lubricant and a local distribution of the temperature inside the contact. It was used to predict the friction coefficient along the meshing line of a FZG type C gear.

The numerical results obtained are compared with results extrapolated from twin disc machine lubricant traction curves and with results obtained from real gear tests. The correlation between numerical and experimental gear results is quite good, putting into evidence the need to consider mixed film lubrication, surface roughness, lubricant non-Newtonian behaviour and contact temperature in order to predict realistic values of the friction coefficient between gear teeth.

1. INTRODUCTION

The analysis of the lubricated contact between gear teeth is very complex. It involves two rough surfaces in relative motion separated by a lubricant film. This film when submitted to high contact pressures and to sliding, due to the differences between the surfaces velocities, increases its temperature, modifying significantly its physical properties, namely its viscosity.

Under typical gear operating conditions, the lubricant film separating the contacting teeth is very thin, usually of the same order of magnitude of the surface roughness, which may cause local lubricant film breakdown.

Full film lubrication, usually assumed in EHD contact modelling, is not compatible with this type of problem where local metal to metal contact between the contacting surfaces may occur, as considered in the mixed film lubrication regime.

A possible solution to overcome this difficulty consists in obtaining the pressure field in mixed film lubrication through the conjugation of full film and dry contact (elastic) solutions.

The model developed was applied to a FZG type C gear, 20 mm wide, lubricated with an ISO VG 150 mineral oil. The aims were to calculate numerically the average friction coefficient along the gear meshing line and compare those values with the ones obtained experimentally by Hohn [1].

2. EXPERIMENTAL RESULTS

2.1. Contact Roughness

The roughness of active flanks of pinion and wheel teeth was measured using adequate measure length and "cut-off". Afterwards, the roughness profiles were numerically treated in order to eliminate the typical flank curvature of gear teeth. The Fourier transform was applied to each roughness profile (pinion and wheel) in order to identify wavelengths and amplitudes of the tooth profile components.

The numerical reconstitution (Fourier series) of each roughness profile was obtained neglecting all wavelengths smaller than 12 μm. This procedure enables the elimination of the measuring noise that may have been induced during roughness measurement, which influences significantly the dry contact pressure distribution.

The dry contact model used reduces all problems to the contact between a rough surface and a smooth surface, and so the gap $H^0(x)$ between the contacting rough surfaces of the pinion and wheel flanks must be determined.

Considering that $H_1(x)$ and $H_2(x)$ are, respectively, the roughness profiles for the pinion (upper surface) and for the wheel (lower surface), the sum of both profiles gives

$$H(x) = H_1(x) - H_2(x) \qquad (1)$$

$H(x)$ may contain either negative values, meaning the two surfaces penetrate each other, or only values greater than zero, meaning the two surfaces do not touch. The minimum value of $H(x)$ must be determined,

$$H_{Min} = MIN\,[H(x)]\,, \qquad (2)$$

and added to $H(x)$, in order to obtain the gap between the two rough surfaces,

$$H^0(x) = H(x) - H_{Min}\,. \qquad (3)$$

In this way both roughness profiles touch in a single point before load is applied to the contact.

Figure 1 shows a partial representation of the FZG pinion and wheel roughness as well as the composed roughness.

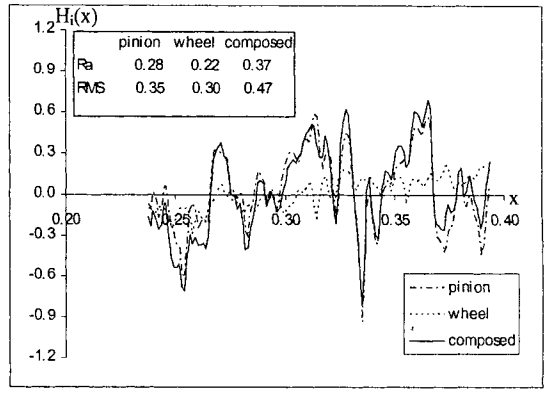


Figure 1. Roughness of a FZG gear: pinion, wheel and composed.

2.2. Lubricant Rheology

Figure 2 presents the lubricant traction curves, obtained from tests preformed in a twin-disk machine [1]. From those results (48 experimental points) the lubricant rheological parameters were calculated, using the model presented in paragraph 3.4.

Figure 2 also shows the lubricant traction curves obtained using the calculated rheological parameters. The correlation between experimental and numerical results is satisfactory (R^2=0.85), particularly for high sliding rates.

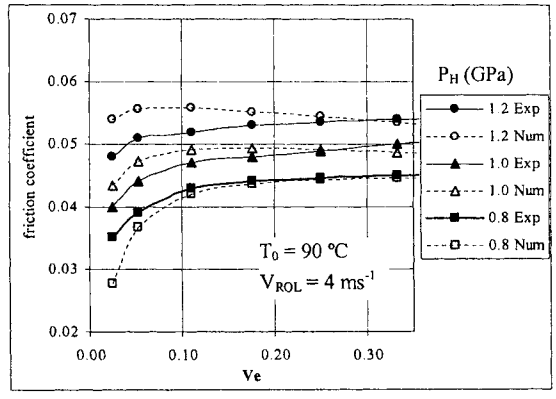


Figure 2. Lubricant traction curves (oil ISO VG 150) from twin disk machine tests [1]: Exp– experimental results, Num– numerical results.

2.3. Application to FZG gears

Hohn et al [1] performed several FZG tests with type C gears, using speeds going from 350 to

3500 r.p.m. (pitch line velocity between 2 to 20 m/s) and 4 FZG load stages (5, 7, 9 and 11), corresponding to maximum Hertzian pressures at pitch point going from 0.8 to 1.7 GPa. The oil bath temperature was kept constant at 90°C.

Figure 3 shows the average friction coefficients between gear teeth measured in those tests, which are significantly smaller than the ones extrapolated from twin-disk machine traction curves shown in the same figure. That difference is quite significant at high loads.

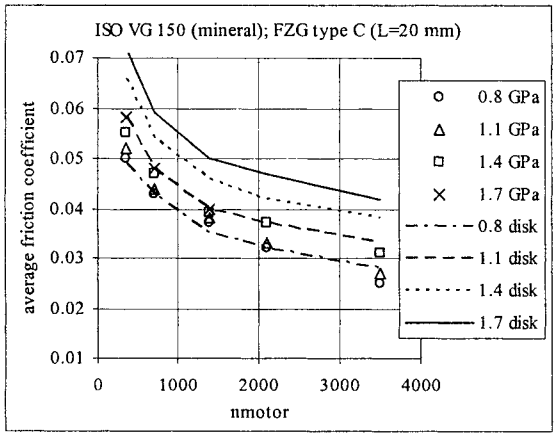


Figure 3. Average friction coefficients (experimental and extrapolated from disk machine tests) for FZG type C gears [1].

3. NUMERICAL MODEL

The numerical model developed may be divided in three stages. The first one is related to the input data, to the pinion and wheel surface roughness and to the evaluation of lubricant parameters. On the second stage the pressure field and deformed geometry of the mixed film contact are obtained from dry rough contact and rough full film EHD contact, through a load share function. The third stage corresponds to the evaluation of the lubricant film shear stresses and temperatures, as well as the contact global friction coefficient.

3.1. Dry contact between rough surfaces

The dry contact model applies elastic half-space theory to the contact between rough surfaces [2 - 4].

The contact kinematics and lubricant presence are not considered in this model. It consists of putting into contact the two real rough surfaces from the pinion and wheel teeth.

The Hertz model for normal contact between two elastic solids, submitted to a normal load (*Fn*), establishes the following conditions for the contact zone:

$$u_n - \delta \geq -H^0, \tag{4}$$

$$p_{dry} \geq 0, \tag{5}$$

$$\int_S p_{dry}\, dS = F_n. \tag{6}$$

Inequality (4) is a non-penetration condition between the contacting solids, where u_n represents the sum of the normal elastic displacements from the surfaces in contact, δ is the penetration and H^0 is the non-deformed gap between the contacting surfaces. The second inequality establishes that the pressures are positive inside contact area and nulls outside. Finally equation (6) establishes the balance between the applied normal load and the pressure distribution inside the contact area.

The solution of this system (inequalities 4 and 5 and equation 6) gives the pressure field inside the contact area and the geometry of the deformed surface, as shown in figure 4. The contact pressure field is usually characterized by high pressure peaks, and may be regarded as a lower limiting case of the mixed film lubricating regime – the dry rough contact.

3.2. Rough contact in full film lubrication

The model developed by Ai and Cheng [5, 6] gives the pressure field inside a rough contact under full film lubrication condition. The lubricant is considered a Newtonian fluid and the contact isothermal.

This simplified model calculates the local perturbations of a rough contact pressure field, using the FFT of the composed roughness profile. The pressure distribution is given by the following equation:

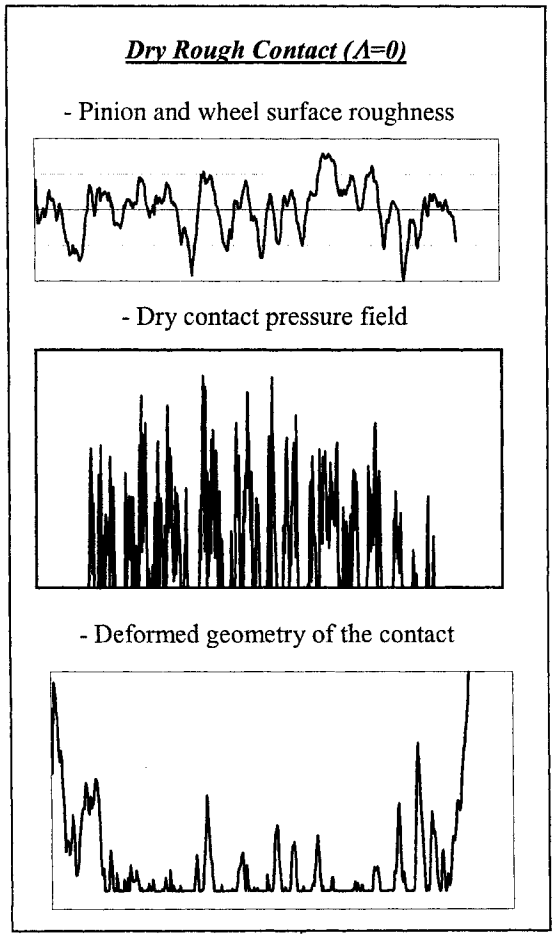


Figure 4. Dry rough contact: Pressure field and deformed contact geometry.

$$P_{ehd}(x) = P_H(x) + \\ +\varpi(x) \times L_p \sum_{i=1}^{nf} a_i^{0.768} \times \ell_i^{-0.806} \times sin\left(\frac{2\Pi x}{\ell_i} + \Phi_i\right) \quad (7)$$

The *Lp* parameter characterizes the contact operating conditions in terms of load, speed, slide-to-roll ratio, geometry and lubricant properties, being defined as:

$$L_p = 0.014095 \times W_\ell^{-0.35152} \times R_{eq}^{0.55084} \times \\ \left(E'\right)^{0.61052} \times \alpha_\eta^{0.09768} \times \left(\eta_0 \times V_{rol}\right)^{-0.16132} \times C_s \quad (8)$$

where

$$C_s = 0.507 \times \delta \times |V_e|^{0.899} + 0.074 \quad (9)$$

with $\delta = 1$ *if* $V_e < 0$ and $\delta = 0.83$ *if* $V_e \geq 0$.

The *Lp* parameter, defined by equation (8), was determined under full film lubrication conditions with specific lubricant film thickness (Λ) values greater than 3, that is, for surfaces with excellent finishing, where the *Rq* roughness parameter values are around 0.120 μm.

The a_i amplitudes of each component of the composed roughness profile (for the contacting pinion and wheel surfaces) have a significant influence in the amplitude of pressure variations.

In this case the Ai & Cheng model is applied considering that $\Lambda = 3$ and $V_e = 0$. This means the centre film thickness is given by $h_0 = 3 \times R_{qC}$, where R_{qC} is the RMS roughness parameter of the composed roughness profile $H^0(x)$. The sliding speed $|U_1 - U_2|$ is neglected, and thus $V_e = 0$. The roughness of gear teeth flanks is very high when compared to those used by Ai & Cheng to correlate their model, and in some cases may generate negative contact pressures that are physically impossible. Neglecting the sliding speed avoids this problem.

Figure 5 shows the Ai & Cheng model. The pressure field obtained may be regarded as an upper limiting case of the mixed film lubrication regime – the full film EHD rough contact.

3.3. Load share function

In mixed film lubrication the normal contact load is supported both by the lubricant film and by the direct contact between the surfaces roughness. The *load share function* establishes the amount of load supported by the dry contact (lower limit) and the amount of load supported by the full film (upper limit).

Defining $f(\Lambda)$ $(0 \leq f(\Lambda) \leq 1)$ as the percentage of load supported by the lubricant film, the pressure distribution in mixed regime is determined by the following equation:

$$P(x) = (1 - f(\Lambda)) \times P_{dry}(x) + f(\Lambda) \times P_{ehd}(x) \quad (10)$$

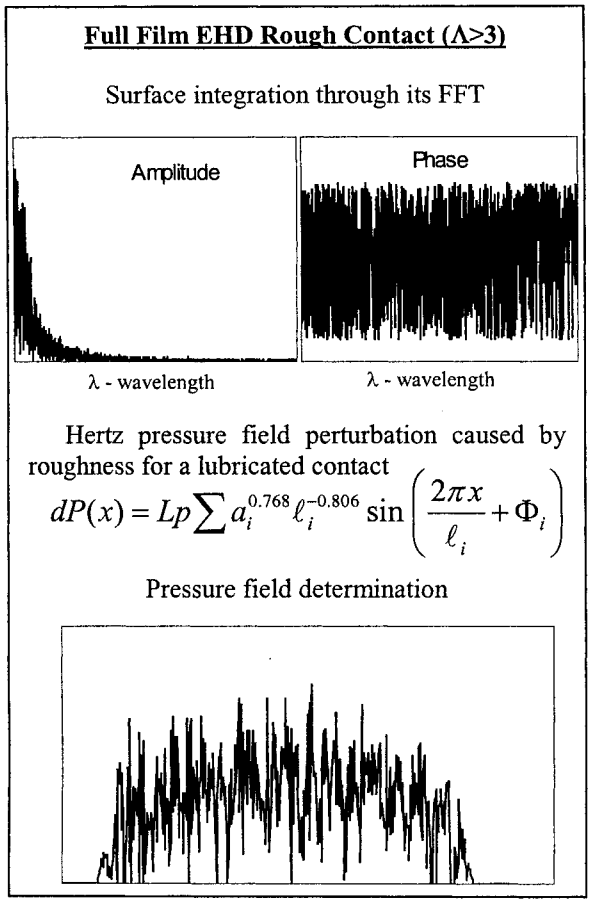


Figure 5. Full film pressure distribution [5].

Zhu and Hu [7] proposed the following definition of the load share function, in which K takes the value of 5.7:

$$f(\Lambda) = \frac{K\Lambda^{1.3}}{1 + K\Lambda^{1.3}}. \qquad (11)$$

However, similar expressions used for contact fatigue failure probability [8] depend on other parameters besides the specific lubricant film thickness, Λ.

In this mixed film lubrication model it's considered that the load share function depends not only on Λ but also on gear pitch line velocity Vt, and the values of K in $f(\Lambda)$ are determined from the gear failure probability curves from Wellauer and Holloway, presented in figure 6 [9].

These curves, based on experimental gear failure results, establish the specific lubricant film thickness (Λ) for which the probability of tooth contact failure is 5% and 80%, as a function of the gear pitch line velocity.

The load share function was developed postulating that, for a given gear pitch line velocity (v_t), when $\Lambda = \Lambda_5$, $F(\Lambda) = 0.8$. So, using equation (11) it can be found that

$$K(v_t) = \frac{0.8}{(1-0.8)\Lambda_5^{1.3}} \qquad (12)$$

Figure 7 shows the $f(\Lambda)$ function for different values of gear pitch line velocity and as a function of the lubricant specific film thickness.

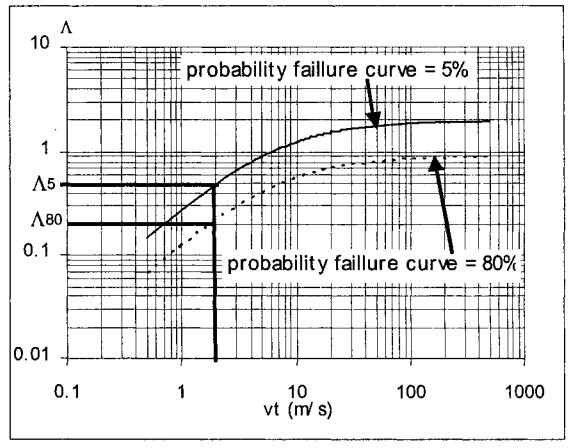


Figure 6. Gear failure probability curves (5% and 80%) from Wellauer e Holloway [9].

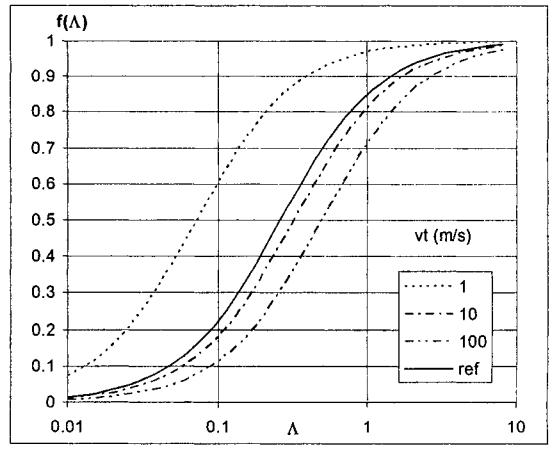


Figure 7. Load share function $f(\Lambda)$.

3.4. Shear stresses in the lubricant film

The local shear stresses in the lubricant film depend on the normal pressure distribution, on the lubricant film geometry and on the operating conditions, taking into account the local temperature variations inside the contact. These shear stresses are determined using a model developed by Sottomayor et al [10].

This model assumes that lubricant behaves as a viscoelastic-plastic fluid, in agreement with the rheological law proposed by Bair and Winer [11]

$$\dot{\gamma} = \frac{\dot{\tau}}{G} - \frac{\tau_L}{\eta} \times ln\left(1 - \frac{\tau}{\tau_L}\right) \tag{13}$$

The variation of the lubricant viscosity with pressure and temperature follows the Roelands law

$$\eta = \eta_0 \times exp\left[\alpha_\eta^* P\right], \tag{14}$$

where

$$\alpha_\eta^* P = \left(ln(\eta_0) + 9.67\right) \cdot \left[\left(1 + 5.1 \times 10^{-9} P\right)^Z \left(\frac{T - 138}{T_0 - 138}\right)^{-S} - 1\right] \tag{15}$$

and

$$Z = \frac{0.196 \times 10^9 \alpha_\eta}{ln(\eta_0) + 9.67} \tag{16}$$

$$S = \frac{\beta_\eta (T - 138)}{ln(\eta_0) + 9.67}. \tag{17}$$

The other rheological parameters, the lubricant shear modulus G and limiting shear stress τ_L, present in the Bair and Winer law, are affected by pressure and temperature according to exponential laws, as

$$\Theta = \Theta_0 \times EXP\left[\alpha_\Theta P + \beta_\Theta \left(\frac{1}{T} - \frac{1}{T_0}\right)\right] \tag{18}$$

The film and surface temperatures are obtained from the solutions of the energy equation applied to the lubricant film and contact surfaces, using a method proposed by Tevaarwerk [12].

The shear stresses in the lubricant film depend on the balance among all the physical phenomena involved [10], mainly the thermal effects that are preponderant when the contact operates at significant slide –to-roll ratios, which is the case in gears away from the pitch point.

4. NUMERICAL AND EXPERIMENTAL RESULTS

Figure 8 shows the pressure fields in dry contact, full film and mixed film lubrication, for point A of the meshing line (initial engagement point – wheel head / pinion root, Ve=0.65) of the FZG type C gear. The wheel speed is 2100 rpm (pitch line velocity of 12 m/s) for load stage 9 (maximum Hertzian pressure of 1.4 GPa at pitch point).

These pressure fields were determined using the mixed film lubrication model developed and the lubricant rheological parameters referred above.

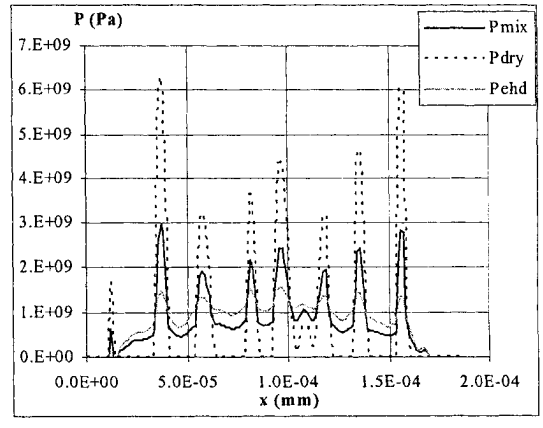


Figure 8. Pressure fields (dry, EHD and mixed) at point A of the meshing line for the FZG type C gear.

The value of the load share function in this case is 0.69, meaning that the mixed film pressure field corresponds to 69% of the full film EHD pressure field and 31% of dry pressure field.

Figure 9 shows the lubricant film shear stress field, $\tau(x)$, and local friction coefficient $\mu(x)$, calculated in at same point of the gear meshing line (point A) and for the same operating conditions. The global friction coefficient at point A (μ_{gA}) is about 0.03.

Figure 10 shows the global friction coefficients along the gear meshing line (12 points). For the operating conditions considered (2100 rpm and FZG load stage 9) the average friction coefficient (μ_{AVG})

between the active flanks of the pinion and wheel teeth (average of the global friction coefficients along the gear meshing line) is 0.034.

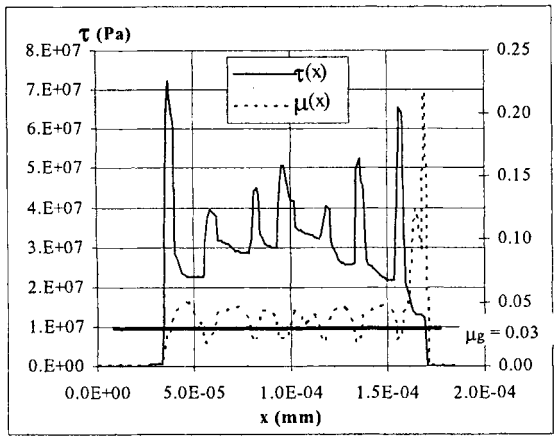


Figure 9. Lubricant shear stress field and local friction coefficients, at point A of the meshing line.

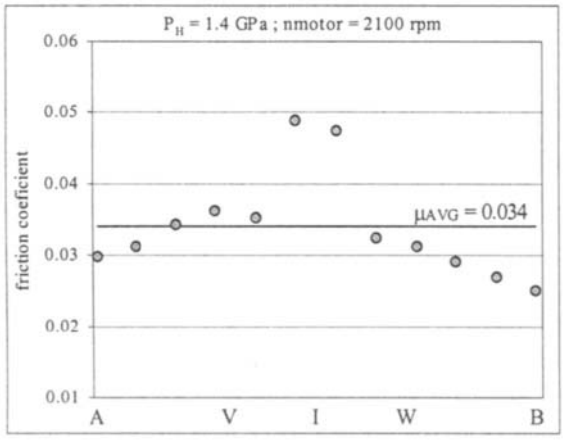


Figure 10. Global friction coefficients along the meshing line for FZG type C gear.

In the neighbourhood of the pitch point the global friction coefficient is still quite high. In fact, the smaller slide-to-roll ratio occurring in that zone generates smaller heat dissipation. Consequently, the lubricant temperature is lower and the lubricant viscosity remains quite high, producing high tangential stresses as confirmed by the traction curves shown on figure 2. Only at pitch point (or very close to it), where $Ve \approx 0$, the tangential shear stresses and the global friction coefficient are very small, but they have a very limited influence on the average friction coefficient over the meshing line.

Figure 11 shows the average friction coefficients calculated using the mixed film lubrication model and compares them with the experimental results, already shown in figure 3, for the FZG type C gears tested.

It's clear that the mixed film lubrication model calculates more realistic values of the average friction coefficient between the active flanks of the gear teeth.

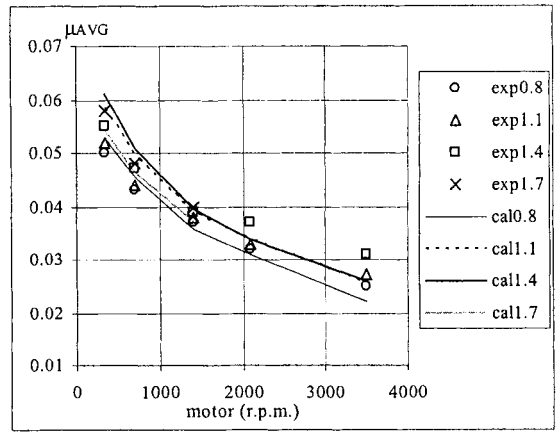


Figure 11. Experimental and numerical average friction coefficients between gear teeth.

5. DISCUSSION OF THE RESULTS

As shown in figure 3, the comparison between average gear friction coefficients measured in gear tests and those extrapolated from lubricant traction tests, performed in the twin disk machine, reveals some inconsistency, mainly concerning the influence of load or contact pressure.

The extrapolation of friction coefficient results, from twin disk machine tests to gears, neglects several important aspects that contribute to the discrepancies found:

- One of the main aspects is the significant difference in surface roughness between the disks and the FZG type C gears, which is 4 times higher for the gears, besides its different orientation relatively to the rolling speed.
- In the twin disk tests, an increase of the maximum Hertzian pressure from 0.8 to 1.7 GPa, is equivalent to multiply the maximum real contact pressure by a factor of 2, since the disks roughness is very small. In gears, the same increase in the maximum Hertzian

pressure, corresponds to an increase of the real contact pressure at roughness peaks from 2.5 to 4.0 GPa, corresponding to a multiplying factor of 1.6, which is smaller then the one obtained for the disks.

- At high loads the thermal behaviour of the contact is of major importance. In fact, in the presence of significant sliding, higher contact pressures cause higher operating temperatures, lower lubricant viscosities and, consequently, smaller lubricant film shear stresses and friction coefficients. In such conditions, the viscosity increase caused by the pressure increase, the piezoviscous effect, becomes less important than the thermal effect.

These arguments justify the differences shown in figure 3 between friction coefficients measured in gear tests and those extrapolated from the twin disk machine tests.

Local pressures up to 4 GPa are reasonable in high loaded rough contacts, which is the case of gears. The model used is not able to evaluate the plastic deformations at the roughness peaks that naturally occur under those conditions. However, it has been shown [13] that such plastic deformation only occurs at the roughness peaks and has a small influence on the real pressure distribution.

The mixed film lubrication model, being more efficient in modelling the effect of surface roughness on the lubricant film shear stresses, gives closer results to the experimental values, as explained above. These results also demonstrate how important surface roughness is for the evaluation of the friction coefficient in mixed lubrication regime.

Analysing figure 11, which compares experimental and numerical friction coefficients, it is clear that the influence of load (or maximum Hertz pressure) decreases when the rotational speed increases, particularly for high speeds. This tendency is very significant for the numerical values.

Although unexpected, the numerical friction coefficients obtained for the highest contact pressure (1.7 GPa) are smaller than those calculated for the contact pressures of 1.4 and 1.1 GPa. A possible justification for this behaviour is related to local lubricant film breakdown that inevitability occurs at the roughness peaks for high contact pressures and speeds. Lubricant film breakdown is not yet taken into account by the mixed film lubrication model developed.

The local lubricant film breakdown, promotes an increase in the local friction coefficient caused by metal/metal contact, which is extended to the global friction coefficient. Being so, the probability of lubricant film breakdown is strong for high loads and speeds. It's clear that local criteria for lubricant film breakdown must be developed and introduced into the mixed film lubrication model.

CONCLUSION

A model for mixed film lubrication was developed taking into account the real surface roughness of gear teeth in line contact.

The numerical results obtained for the average friction coefficient between the active flanks of gear teeth in a FZG type C gear, are similar to the corresponding experimental values.

The strong influence of surface roughness on the friction coefficient in mixed film lubrication regime is justified.

REFERENCES

1. Hohn, B. R., Michaellis, K., Doleschel A., "Frictional Behaviour of Synthetic Gear Lubricants", Tribology Research: From Model Experiment to Industrial Problem, G. Dalmaz et al. (Editors), Elsevier Science B. V., pp 759-768, 2001.

2. Seabra, J., Berthe, D., "Influence of surface waviness and roughness on the normal pressure distribution in a Hertzian contact", Journal of Tribology, Vol 109, pp. 462-470, 1987.

3. Seabra, J., "Influence de l'ondulation des surfaces sur le comportement des contacts Hertziens secs ou lubrifiés", PhD Thesis, n°8ISAL0018, INSA, Lyon, 1988.

4. Campos, A., "Determinação da geometria do filme lubrificante, da distribuição de pressão e das tensões de corte num contacto elastohidrodinâmico. Aplicação a rolamentos de rolos.", MSc thesis, DEMEGI, FEUP, 1994.

5. Ai X., Cheng, H. S., "A transient EHL analysis for line contacts with measured surface

roughness using multigrid technique", ASME Journal of Triboligy, Vol.116, Julho 1994.

6. Barral, R., "Modélisation du champ de pression, du film de lubrifiant et des contraintes tangentielles dans un contact EHD rugueux. Premières études des contraintes engendrées dans les massifs", Rapport technique T03, Etude du pitting de matérieux d'engrenages de boites de vitesses manuelles – approche théorique et pratique, RVBV 10 (RENAULT – DIMAT) e 98.21.01 (INEGI – CETRIB), 1999.
7. Zhu, D. and Hu, Y., "A Computer Program Package for the prediction of EHL and Mixed Lubrication Characteristics, Friction, Subsurface Stresses and Flash temperatures Based on Measured 3-D Surface Roughness", STLE Tribology Transactions, Vol.44, N° 3, pp. 383-390, 2001.
8. Gohar, R., "Elastohydrodynamics", Ellis Horwood Limited, England, 1988.
9. Henriot, G., "La Lubrification Industriele - La Lubrification de Engrenages" Tome 1 - Transmissions, Compresseurs, Turbines, Publications de l'Institute Français du Pétrole,Éditions Technip, pg. 297 - 385, 1984.
10. Sottomayor, A., Seabra, J., Campos, A., e Ferreira, L., "Thermal and non-Newtonian analysis of EHD line contacts", Proceedings of the 21th "Leeds-Lyon Symposium" on Tribology, Leeds, UK, "6 a 9 Septembro de 1994.
11. Bair, S., Winer, W., "A Rheological Model for Elastohydrodynamic Contacts Based on Primary Laboratory data", ASME Journal of Lubrication Technology, vol.101, July 1979.
12. Tevaarwerk, J. L., "Traction calculations using the shear plane hypothesis" Proceeding of the 6th Leeds-Lyon Simposium on Tribology, Lyon, Septembre 1979.
13. C. Mayeur, P. Sainsot and L. Flamand, "A numerical elastoplastic model for rough contact" Technical report ELAB T017, from BRITE-EURAM II project ELABOMM (Endurance Limit of Actual Bearings depending on Operating conditions, Materials and Manufacturing), March 1966, 25pp.

ACKNOWLEDGMENTS

The authors would like to acknowledge the "Fundação para a Ciência e Tecnologia" from the "Ministério da Ciência e Tecnologia" of Portugal, for their support to this work through project PCTI/1999/EME/35957 (PRAXIS XXI).

NOMENCLATURE

- a_i – amplitude of the component i of the roughness profile
- b – Hertz contact half-width (rolling direction)
- $dP_i(x)$ – Pressure variation of the rough EHD contact, for component i of the roughness profile
- E' – equivalent Young's modulus for the contact
- G – lubricant shear modulus
- F_n – normal contact load
- $H^0(x)$ – non-deformed gap between contacting rough surfaces.
- H_0 – lubricant film thickness at contact centre
- λ_i – wavelength of the component i of the roughness profile
- nf – number of terms used in profile
- $P(x)$ – mixed film contact pressure
- $P_{EHD}(x)$– full film contact pressure
- $P_H(x)$ – Hertzian contact pressure
- P_H – maximum Hertzian contact pressure
- $P_{DRY}(x)$– dry contact pressure
- Req – equivalent curvature radius for the contact
- T_0 – oil temperature
- U_1 – tangential speed surface 1
- U_2 – tangential speed surface 2
- V_e – slide-to-roll ratio $V_e = \frac{|U_1 - U_2|}{U_1 + U_2} = \frac{|V_{SLI}|}{V_{ROL}}$
- V_{ROL} – rolling speed
- V_{SLI} – sliding speed
- W_λ – load per unit length
- x – x coordinate of the roughness profile
- α_θ – generic coefficient for variation with pressure
- α_η – lubricant piezoviscosity coefficient
- β_θ – generic coefficient for variation with temperature
- β_η – lubricant termoviscosity coefficient
- Φ_i – phase of component i of the roughness profile

Φ_T – thermal correction factor for the mean film thickness

$\dot{\gamma}$ – shear strain rate

η_o – dinamic viscosity

Λ – specific lubricant film thickness $\Lambda = \dfrac{h_o \Phi_T}{\sigma}$

μ_{AVG} – average friction coefficient along the gear meshing line

μ_g – global friction coefficient inside the contact

$\mu(x)$ – local friction coefficient inside the contact

Θ – generic rheological parameter

σ – RMS roughness parameter

$\tau(x)$ – local lubricant shear stress

τ_L – lubricant limiting shear stress

$\varpi(x)$ – factor for equation (7):

$$\varpi(x) = \begin{cases} 1 & |x/b| \le 0.9 \\ \dfrac{1}{2}\{1+\cos[10\pi(|x/b|-0.9)]\} & 0.9 \le |x/b| \le 1 \\ 0 & |x/b| > 0.9 \end{cases}$$

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Influence of Temperature Distribution on the Tribological Performance of Automotive Synchronisers

Prof. Dr.-Ing. Gerhard Poll and Dr.-Ing. Marcus Spreckels
University of Hannover,
Am Welfengarten 1, D-30167 Hannover, Germany

An essential component of manually operated vehicle transmissions is the synchronizer. Synchronizers have the task of minimizing the speed difference between the shifted gearwheel and the shaft by means of frictional torque before engaging the gear. Proper operation requires a sufficiently high coefficient of friction. It is common practice to investigate the friction and wear behaviour under various loading conditions on test rigs or in vehicle tests. An optimised design of the system with regard to appropriate function and durability on the one hand as well as low cost, low mass and compact over-all dimensions on the other hand requires extensive testing. According to the present state of knowledge, derived from numerous experimental investigations, temperature can be attributed the most significant influence on the tribology of synchronizing systems. Therefore, the influence of various loading conditions on contact temperature was investigated; a relation between temperature and tribological performance was established.
The Finite Element Method was applied to simulate the thermal behaviour of a synchronizing system depending on different operating conditions. A result of this simulation is the spatial and time-dependent temperature distribution in the area of contact.
Characteristics of the tribological performance of a molybdenum coated synchromesh ring in contact with a steel cone were derived from extensive experimental investigations. Significantly different friction and wear patterns can be distinguished. At heavy loading conditions the coefficient of friction is quite high and continuously severe wear occurs; light operating conditions result in a low friction coefficient, whilst no more wear is observed. Between those two extremes an indifferent regime exists, in which both patterns of tribological behaviour occur.
A reason for this characteristic behaviour of the system described here was found by means of the Finite Element simulation. Apparently, the friction and wear pattern depends on the temperature in the contact area; for mild wear and low friction coefficients the contact temperature must not exceed a critical value in order to avoid severe wear. Within limits the predicted tribological behaviour and the test results are in good agreement.
The calculation of the temperature in the contact area provides a basis for a classification of the load conditions in terms of their thermal and tribological effect, a practically applicable estimation of service life and a design procedure based on numerical simulation rather than on testing. Thus, a reduction of development efforts will be possible.

1 Introduction

One important component of manually operated vehicle transmissions is the synchromesh gear. It is the task of synchromesh gears to reduce the speed differences between the shaft and the idler gear during the gear shift by means of friction before positive engagement. To ensure proper functioning, a sufficiently high coefficient of friction is essential. In addition to the geometry and material properties of the synchromesh as well as lubricant properties, tribological stress conditions also have an influence on its operation (Figure 1). The frictional and wear behaviour of a synchromesh subject to various stresses is determined in practice using test rigs or in vehicle tests. Optimizing the design of the system in terms of function and service life on the one hand and space, weight and cost on the other hand involves a considerable amount of testing.

Various experimental studies have indicated that temperature has a considerable influence on friction and wear behaviour. The aim of the research presented here was thus to compute the effects of the load conditions on the temperature in the contact and to work out the relationship between the tribological behaviour as determined experimentally and the temperature.

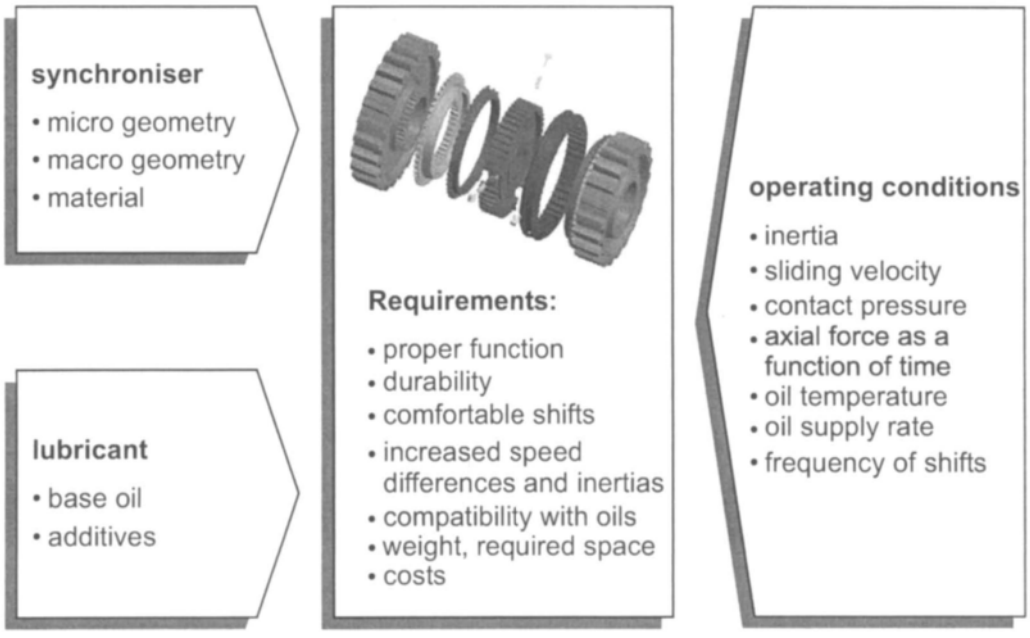


Figure 1: Synchromesh requirements and influencing factors

2 Operation of a synchromesh

Figure 2 shows the components of a synchromesh in an exploded view. The cone is positively engaged with the gear. The synchro body, the synchro sleeve and the synchro ring as well as the springs and the centering pieces of the pre-synchromesh unit are the components that rotate with the shaft. The synchro sleeve can be displaced axially during gear shift with the aid of the gearshift linkage in the direction of the gear. This presses the ring against the cone via the stop teeth. The difference in speed is reduced by friction in the conical friction surface of ring and cone.

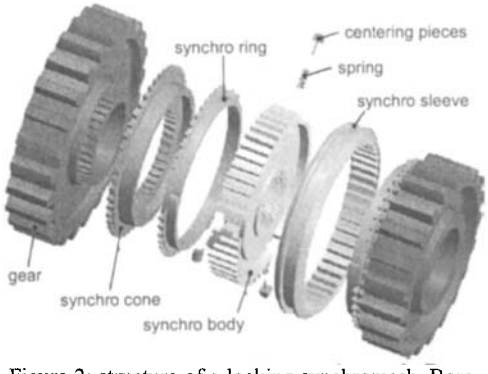


Figure 2: structure of a locking synchromesh. Borg-Warner system

When the speeds have been equalised, positive engagement is created by shifting the synchro sleeve teeth into the spaces between the clutch teeth.

Proper functioning is only ensured if the frictional torque in the conical surface between the ring and the cone is sufficiently large. The frictional torque is created by the axial pressing of the ring against the cone (Figure 3). The normal force in the contact results from the axial pressing force intensified by the cone angle. A frictional torque arises corresponding to the coefficient of friction.

Material removal in the conical surface due to wear can result in the planar flange facings coming into contact. The axial force is no longer solely transmitted via the conical surface. The intensifying effect of the cone partly vanishes and the frictional torque is reduced. This damaging process is generally referred to as axial wear. Gear-change times then rise beyond the permissible level, and in extreme cases it is no longer possible to equalise the speed at all.

A further reason for the reduction in frictional torque may be a decreasing coefficient of friction in the conical surface due to altered tribological conditions. This impairing effect is referred to as decrease in the coefficient of friction.

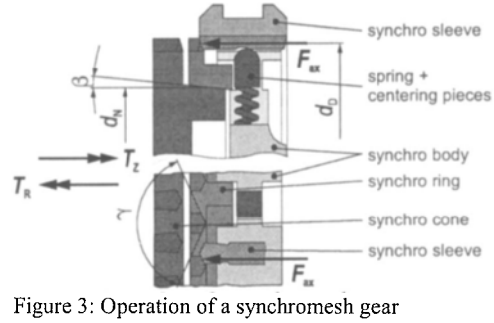


Figure 3: Operation of a synchromesh gear

An additional function of synchromesh gears is that they are intended to prevent positive engagement until speeds have been completely equalised. For this to happen, the locking condition must be fulfilled. The frictional torque T_R must be greater than a restoring moment T_Z which occurs as a result of the pitch angle at the teeth faces. In this case, the opposing teeth block further shifting of the synchro sleeve. In addition to axial wear and the decrease in the coefficient of friction as causes, pitch wear may thus result in the locking condition being violated if the aligning torque rises as a result of a reduction in the pitch angle.

Any malfunction in the system makes a noticeable grating sound when the teeth of the sliding sleeve and the cone meet whenever there is a speed difference.

Therefore, in addition to constant tribological properties, it must also be ensured that the axial gap between the ring and the cone – the wear reserve – is greater than the axial displacement of the ring due to material removal as a result of wear.

3 Experimental investigations

The experimental investigations were conducted on a simple synchromesh from a lorry gearbox with a molybdenum coating. The coating was applied using thermal spraying. Oil drainage grooves allow for removal of the oil from the gap. The lubricant used in both cases was a standard gear oil.

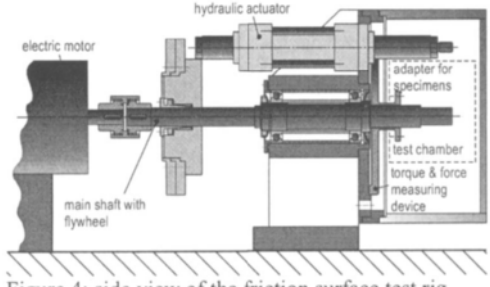


Figure 4: side view of the friction surface test rig

The experiments were conducted on a special test rig (Figure 4). The basic principle of the test rig is that the test-rig shaft is coupled with a flywheel and accelerated by an electric motor to a predetermined speed. Then the shaft is braked over the friction surface against the stationary measuring device. Gear shift is by means of a hydraulic cylinder. The force of the hydraulic cylinder is regulated. Lubrication is by oil circulation with temperature control. The test specimen is connected with the test-rig shaft and with the sensor.
An inductive displacement sensor measures the axial displacement of the ring with regard to the cone. When the axial force is constant, this shift corresponds to the wear in the friction surfaces. The sensors specially developed for this test rig measure axial force and frictional torque at the test piece, so that it is possible to precisely determine the coefficient of friction in the friction surfaces. Also, the temperatures were measured in the component and the finite-element model described below was compared with the measurements.

4 Finite-element computation

In addition to the experimental investigations using the test rig, the thermal behaviour of the components was analysed using a finite-element model. The synchro ring and cone were represented by a rotationally symmetrical two-dimensional model (Figure 5). The elements used have both mechanical and thermal degrees of freedom. In line with the materials used, the elements were assigned the mechanical and thermal properties of a steel alloy. The molybdenum coating on the synchro ring was modelled using a series of elements with the material properties of molybdenum. In the area of the potential contact surface between ring and cone, special contact elements were placed which detect any contact between the components and compute the deformations, axial loads and heat fluxes accordingly.

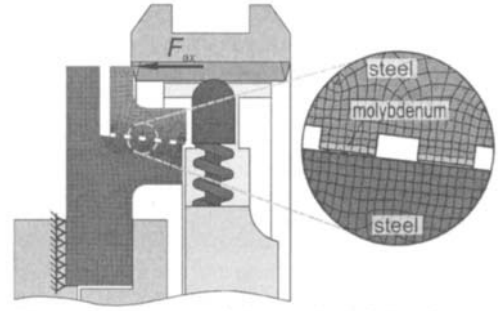


Figure 5: Diagram of the meshed finite-element model

The interactions with the other components were replaced by suitable boundary conditions. Thus in the area where the cone is mounted, the model was defined mechanically in all the co-ordinate directions. At the teeth of the synchro ring, the contact with the sliding sleeve is replaced by an external force. The friction power was simulated in the form of a heat flux input at the contact elements. The heat was dissipated by convection at the edges of the components.

The heat flux is produced in line with the change in stress in the test-rig experiment (Figure 6). The change in axial load over time in the test rig is achieved in such a way that after a linear increase in axial load with defined gradient there is a phase of constant axial load up until the end of the gear shift. The sliding velocity results from the initial sliding velocity before the gear shift and a deceleration as a result of the frictional torque. Apart from the

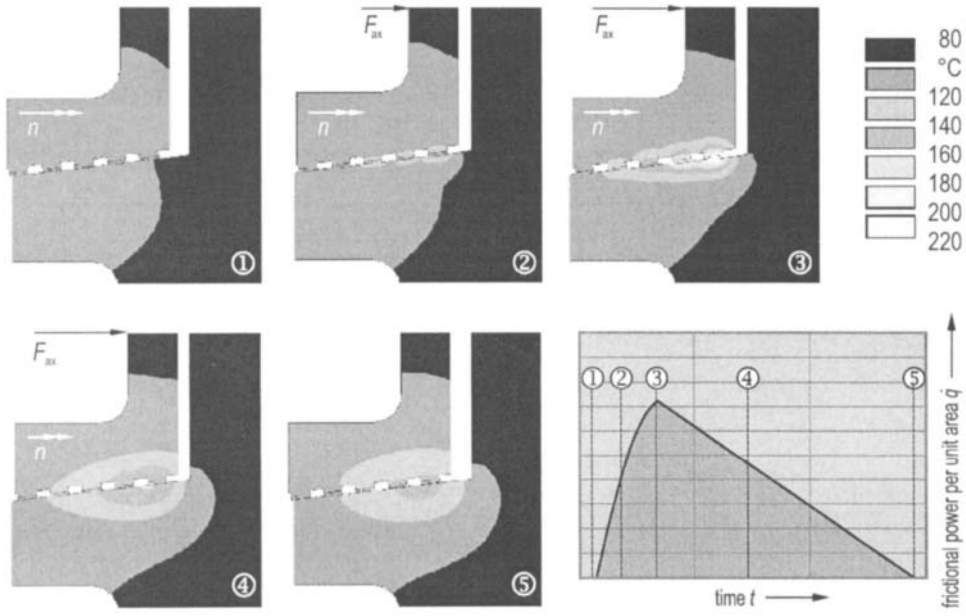


Figure 7: Temperature distribution in the synchro cone and ring at various points in time during a shift cycle according to the simulation

geometrical dimensions, other factors are the coefficients of friction, the mass moment of inertia and the change in axial load. The frictional power per unit area – the product of axial load, sliding velocity and coefficient of friction – is an important factor concerning the thermal stress of the friction surface. Assuming that the kinetic energy is completely converted into heat, the frictional power per unit area corresponds to the heat flux density supplied to the component. It is thus decisive for the temperature distribution. These relationships are employed as part of the finite-element calculation for the computation of heat fluxes in the contact. The specific frictional work per gear-shift corresponds to the area beneath the curve of frictional power per unit area.

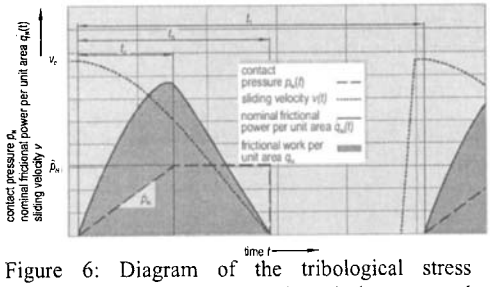


Figure 6: Diagram of the tribological stress parameters as a function of time during repeated gear shifts with a cycle time t_T. (t_A: axial force build-up time; t_s: speed equalization time)

As the parameters axial load, sliding velocity and mass moment of inertia have a direct influence on the thermomechanical conditions in the contact, they will referred to as primary stress parameters. Secondary stress parameters are the frequency of gear shifts – the inverse corresponds to the time interval between two gear shifts – as well as the oil temperature and oil supply rate. The frequency of gear shifts is thus a measure of how much time is available for heat dissipation between two gear shifts. The average frictional power per unit area is the frictional work per shift with respect to the cycle time t_T. One can visualise the average frictional power per unit time as a stationary heat flux which involves the same thermal energy over a shift cycle as the discrete individual gear shifts. The oil temperature and oil supply rate of the injected lubricant are primarily responsible for the heat dissipation by the lubricant.

The result of the finite-element computation is shown in Figure 7. It shows the temperature distribution as a function of time in the test component over a shift cycle at five different times. The zone thermally influenced by the friction is restricted to a narrow zone around the contact.

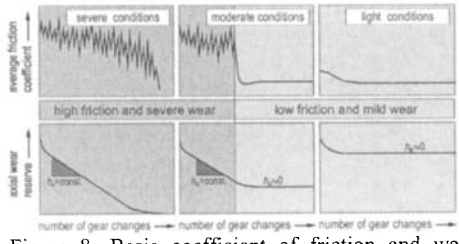


Figure 8: Basic coefficient of friction and wear behaviour at different externally applied tribological stress levels

5 Results of the experimental investigations

The division of tribological behaviour into three regimes /1/, which has already been used with other systems with a molybdenum coating, can also be

applied to the system with a thin molybdenum coating (Figure 8). In the range with low tribological stresses (regime I) the system stabilises after an initial phase with slightly elevated coefficients of friction and mild wear, in a state with non-measurable wear and constant coefficients of friction. Within the around 100,000 gear shifts up to the end of the experiments, no further changes in the tribological behaviour occur. This behaviour will be referred to as low-level friction and wear.

Tribological behaviour at high stresses (regime III) is characterised over the entire service life by strongly scattered, high coefficients of friction and severe wear. The service life is limited by axial wear. The wear remains constant, apart from the initial phase and the final phase of the experiment, and is characterised by the wear gradient, i.e. the slope of the straight line. As shown in previous studies, this gradient is primarily proportional to the converted frictional work /1/. Another, more accessible description of the wear gradient is "wear per shift". This behaviour will be referred to as a high level of friction and wear.
In between there is a transitional regime (II). After a high level of friction with recognisable wear the system moves without any change in the external stress into a low level of friction and wear. The wear gradient in the high level of friction and wear is constant apart from the initial phase, and afterwards no wear is measurable any longer. The end of the experiment is also reached without failure.

Table 1 (left) shows the tribological behaviour of the parameter combinations investigated. Oil temperature, oil supply rate and frequency of gear shifts were kept constant.
One indicator that the temperature causes the observed changes in behaviour is provided by the experiments where the oil temperature and the mean frictional power are varied. Here, the oil temperature was increased for a tribological stress level which under standard conditions had caused a behaviour from the transitional regime II (framed parameter combination with p_N = 3.0 N/mm² and v_0 = 7.5 m/s). The result is that when the temperature is increased there is no longer a transition to a low level of friction and wear, corresponding to a change to a behaviour from regime III. Accordingly, failure as a result of axial wear ensued.

On the other hand, with a stress causing a constant high level of friction and wear from regime III the mean frictional power was reduced (Table 1, right).

	"Standard conditions"	
v_0 [m/s]	$\hat{p}_N$ [N/mm²] 3,0	5,0
4,0		
5,0		D 250
6,0		D 750
6,5	D 250	A 7.600
7,0		
7,5	D 500	A 4.500
8,0	A 4.900	A 3600
9,5	A 3.650	

	increased temperature	reduced frictional power
v_0 [m/s]	$\hat{p}_N$ [N/mm²] 3,0	5,0
4,0		
5,0		
6,0		
6,5		
7,0		
7,5	A 4.500	D 2.000
8,0		
9,5		

Behaviour according to
regime I
regime II
regime III

Reasons for failure:
A: axial wear
R: decrease in the coefficient of friction
D: no failure
(duration of experiment: 100,000 gear shifts)

The figures indicate the number of gear shifts in the high level of friction and wear. For regime III, this figure is equal to the total number of gear shifts

Table 1: Correlation between the parameters governing the tribological stresses, temperature and the classification of performance according to the three characteristic regimes

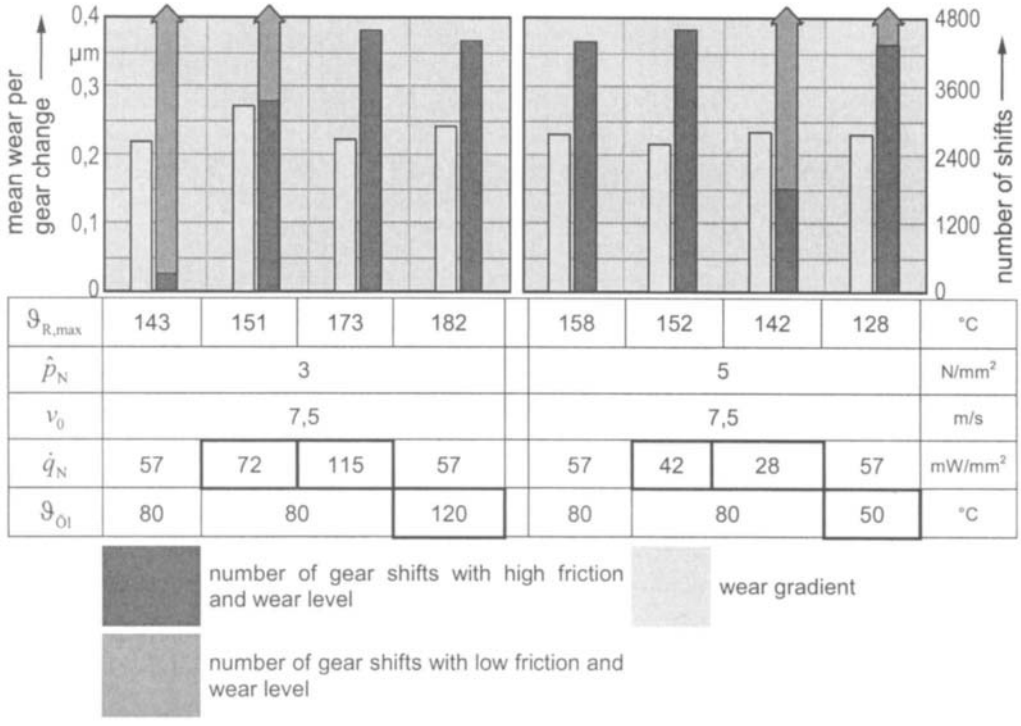


Figure 9: Number of gear shifts with high and low friction and wear level at different friction surface temperatures

A reduction in the mean frictional power at a constant mass moment of inertia is equivalent to increasing the cycle time. Instead of a steady high level of friction and wear we now have more favourable tribological behaviour from regime II with a transition to the low level of friction and wear.

Since the mechanical stress in the contact does not alter when the secondary stress parameters vary, the temperature must therefore be the cause of the altered tribological behaviour. It is conceivable that there is a limiting temperature below which behaviour from the low level of friction and wear occurs and above which behaviour from the high level occurs.

One indicator of this limiting temperature is provided by further experiments where the secondary stress parameters are varied (Figure 9). The bar diagram shows the number of gear shifts at high and low levels of friction and wear under stress levels corresponding to regime II. The dark grey bar segment symbolises the high level, and the light grey bar segment shows the values of the low level. On the left the results are shown of a contact pressure of 3 N/mm^2 at increased temperatures (from left to right). The increase was brought about by an increase in the mean frictional power and the oil temperature. The top line shows the computed maximum friction surface temperatures. When the friction surface temperature rises, the number of gear shifts in the high level of friction and wear rises initially. When a limiting temperature is exceeded, a behaviour with a steady high level of friction and wear is exhibited. One possible limiting temperature is thus between 151° and 173°.

The right-hand diagram shows the results at an a contact pressure of 5 N/mm^2 at reduced temperatures (from left to right). The increase in temperature is thus in the opposite direction. The variational parameters are highlighted in the table. When the friction surface temperature drops, a more favourable tribological behaviour appears with a transition into the low level of friction and wear. The alteration in behaviour takes place at maximum friction surface temperatures between 140° and 150°C.
The wear gradient (empty bar) in the high level of friction and wear is apparently independent of the friction surface temperature when the initial sliding velocity is constant.

6 Comparison with results from the finite-element computation

As shown, the friction surface temperature could be a criterion for the classification of tribological behaviour given different secondary stress parameters. Accordingly, the transition from the high level of friction and wear to the low level in the course of an experiment at constant stress might

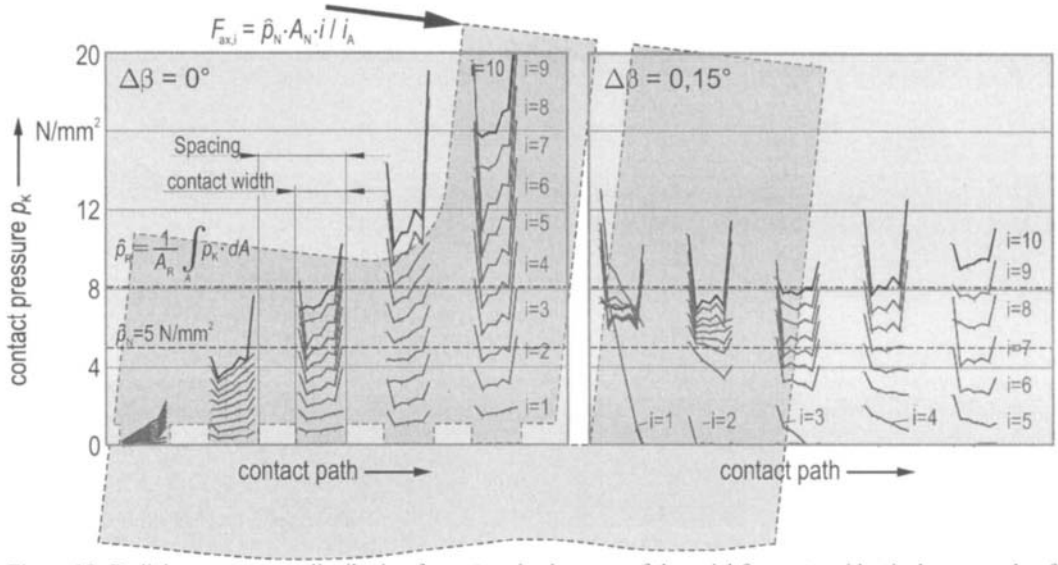


Figure 10: Radial contact stress distribution for a stepwise increase of the axial force at an identical cone angle of the ring and cone (on the left) as well with a slightly larger cone angle of the ring (on the right). "i" characterises the number of loading steps until the respective load level is reached.

be a consequence of changing friction surface temperatures.

An analysis of the radial contact stress or surface pressure distribution shows that this distribution may change as a result of material removal caused by wear. Figure 10 shows the radial contact stress with increasing axial force. The maximum axial load corresponds to a nominal contact stress of 5 N/mm^2. The cone angle of the ring and cone in the left-hand picture have exactly the same values, and so the angular difference between the ring and cone is 0°. It is apparent that there is virtually no mechanical load on the friction surface at the smallest diameter. The contact load is 0 there and then rises in a virtually linear fashion along the line of contact towards the larger diameter. The reason is a distortion of the ring due to the external axial load and the resulting normal force in the friction surface. Factors that have an effect on the axial load distribution in the contact are thus the stiffness of the ring and cone and the elasticity of the contact surface. However, the angles of the conical surfaces of the ring and cone are also involved. Even at an angular difference as small as 0.15° the contact pressure distribution is already completely different, as can be seen here in the right-hand diagram. An angular difference of 0.15° means that the angle at the ring is greater than the angle at the cone by this amount. Then the axial load first starts to increase rapidly at the smaller diameter, while at the larger diameter there is initially no stress on the friction surface.

When the axial force continues to rise, the axial load distribution becomes more uniform up to the point where at a particular force a virtually uniform distribution occurs. Since the frictional power per unit area over the line of contact is virtually proportional to the axial load, the temperature distribution results accordingly

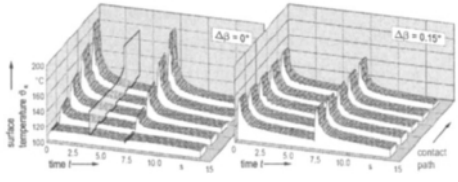


Figure 11: change in friction surface temperature over two shift cycles with the same load but differing cone angle differences according to the finite-element computation.

The temporal and spatial temperature distribution in the friction surface over a shift cycle is shown in Figure 11 at the cone angle differences mentioned above. As with the axial load distribution, when the cone angle of ring and cone is identical there is virtually no temperature increase at the smallest diameter. The maximum temperature that occurs is about 190°C. On the other hand, the maximum temperature at a slightly greater cone angle is around 150°C. The temperature distribution over the line of contact is then virtually uniform.

The left-hand diagram in Figure 12 shows the temporal maxima as a function of the axial location on the line of contact for various angular differences but nominally equal tribological stresses. It can be seen that at a contact pressure of 5 N/mm^2, as was used as the basis for this computation, there is a practically uniform temperature distribution at an angular difference of 0.15°. The minimum friction surface temperature, which always appears just before the beginning of the shift, is almost independent of the position in the contact surface and of the angular difference.

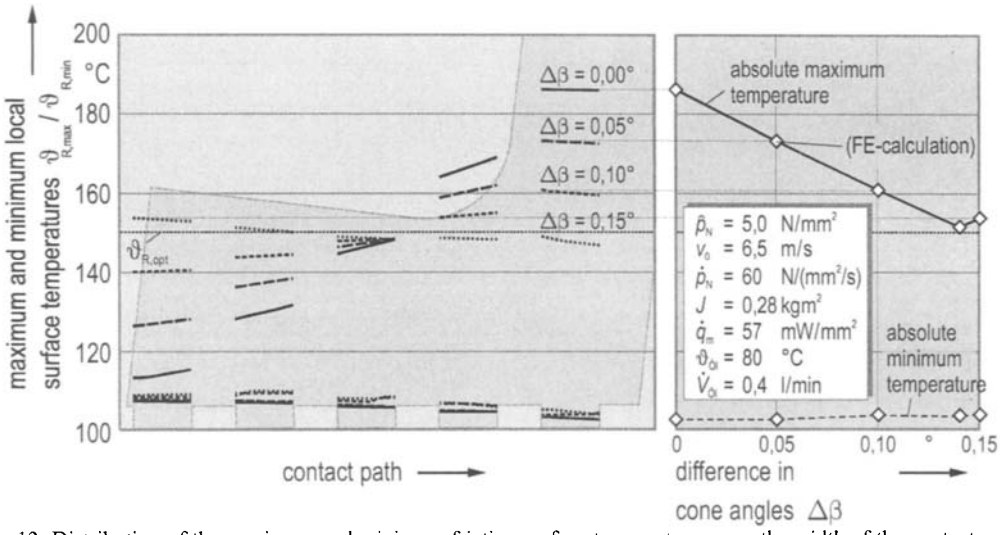


Figure 12: Distribution of the maximum and minimum friction surface temperatures over the width of the contact (left) and the absolute maxima and minima of temperature as a function of the cone angle difference (right).

The integral mean of the temperature over the contact surface is the same for all angular differences and represents the friction surface temperature occurring in the most favourable case with an optimum uniform temperature distribution. The right-hand diagram shows the dependence of the absolute maximum on the angular difference. The absolute maximum is the maximum both temporally and spatially. With increasing angular difference, the maximum friction surface temperature decreases linearly to an optimum value. This value is just below the limiting temperature determined in Figure 9 for moving out of the steady high level of wear. When the optimum value is reached, the temperature rises again. The maxima now appear at the smaller diameter instead of at the larger diameter. It is conceivable that the angles of the ring and cone alter during the course of the experiment due to wear and thus bring about considerably differing mechanical and thermal conditions in the friction surface. After manufacture, the ring and cone have the same angles within manufacturing tolerances. There is thus initially a strongly non-uniform temperature distribution over the friction surface with a high maximum. If wear particles now tend to be removed in the range of high thermal and mechanical stress, then the angular difference will increase and the maximum friction surface temperatures will fall. Below a limiting temperature there is then a transition from the high level to the low level of friction and wear.

Figure 13 shows that classifying tribological behaviour on the basis of a limiting temperature is possible, even when there are differing primary stress parameters. These diagrams show the maximum friction surface temperatures at different axial loads and sliding velocities. The right-hand diagram shows the situation at the beginning of the experiment with the cone angles the same. On the left we see the temperatures when the cone angle difference is optimum, i.e. for the state of uniform temperature distribution. The tribological behaviour is indicated by colours. Black means the high level of friction and wear (regime III), white is the low level (regime I), and grey indicates the transitional regime (II). If one wanted to draw a limiting temperature curve, it would have to begin between the low level and the transitional regime. Under optimised conditions with adapted cone angles, the same limiting curve would have to separate the high wear and friction level regime from the transitional regime. Also, we know from our previous discussions that the limiting temperature should be in the area of 150°C.

It is possible to draw a limiting temperature curve of this kind under the stated conditions, as can be seen here. There are also other ways of drawing a limiting curve, but detailed information is not available from the experiments. It thus seems to be basically possible to predict tribological behaviour on the basis of the friction surface temperatures we have computed.

7 Summary

With the aid of the experimental and theoretical investigations presented here the temperature can be identified as a controlling factor influencing frictional and wear behaviour. It determines the proportionality factor between friction and wear. It was possible to compute the effects of the stress parameters of a synchromesh on the temperature in the contact and to show that there is a relationship between the tribological behaviour as determined experimentally and the temperature.

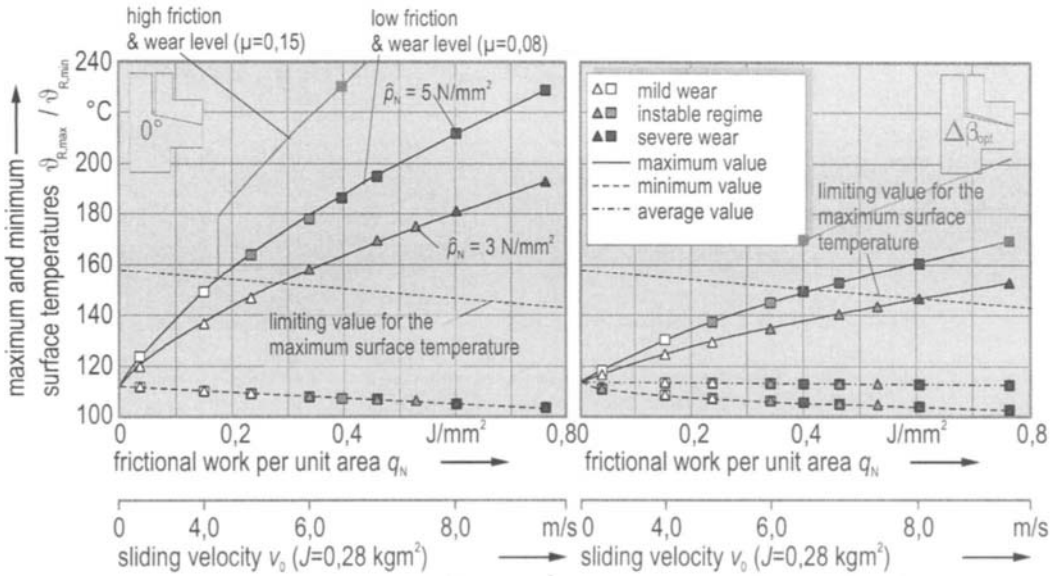


Figure 13: Maximum and minimum computed friction surface temperature as a function of the frictional work per unit area at the contact, with the cone angle the same (left) and adjusted to the optimum value (right).

In addition, the thermal behaviour of a synchromesh was analysed using a model computation on the basis of the finite element method. One result of the computation is the temporal and spatial distribution of temperature in the contact surface.

Information concerning the tribological behaviour of a synchromesh with molybdenum-steel friction pairing under different stresses has been obtained in extensive test-rig experiments. The friction and wear behaviour of this pairing is characterised by three regimes: a steady high level of friction and wear, a steady low level without measurable wear and an initial high level with a transition to the low level in the course of the experiment.

A simulation allowed us to find the cause of this characteristic behaviour: for low-level friction and wear behaviour a limiting value of the friction surface temperature may not be exceeded. Within limits the tribological behaviour predicted in this way correlates very well with the experimental results.

Computation of the friction surface temperatures can thus allow assessment of tribological stresses with regard to their thermal or their general tribological effect. It also allows us to estimate service life in a more practical manner; test-rig experiments can be replaced by simulations, thus ultimately allowing the reduction of development costs.

More details are included in Spreckels' PhD thesis /2/.

References

/1/ Lösche, T; Paland, E.-G.; Poll, G.: Wear Behaviour of Synchronisers in Relation to a Duty Parameter; Tribology for Energy Conservation, Tribology Series, 34 (ISBN 0-444-50033-2, Vol. 34), ed. Dowson, D., Elsevier Science, Amsterdam (1998)

/2/ Spreckels, M.: Einfluss der Temperaturverteilung auf das tribologische Verhalten von Synchronisierungen, PhD thesis, Hannover (2001)

This Page Intentionally Left Blank

SESSION XVII
FRICTION, WEAR AND LIFE OF CONTACTS IN MACHINE ELEMENTS (2) - WEAR

Chairman : Professor T H C Childs

Paper XVII (i) **Feasibility of Using Electrostatic Monitoring for Oil Lubricated Ceramic to Steel Sliding Contacts**

Paper XVII (ii) **High Temperature Abrasion Behaviour of Al/ SiC_p Composites**

Paper XVII (iii) **Friction and Wear of Spheroidal Graphite Cast Iron Under Severe Sliding Conditions**

Paper XVII (iv) **Wear Particles Produced in Repeated Contact of Two Asperities**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Feasibility of Using Electrostatic Monitoring for Oil Lubricated Ceramic to Steel Sliding Contacts

L. Wang[a], R. J. K. Wood[a], T. J. Harvey[a], S. Morris[a], H. E. G. Powrie[b], I. Care[c]

a. Surface Engineering and Tribology Group, School of Engineering Sciences, University of Southampton, Southampton, SO17 1BJ, UK, e-mail: ling.wang@soton.ac.uk

b. Smiths Aerospace Electronic Systems - Southampton, School Lane, Chandlers Ford, Hampshire SO53 4YG, UK, e-mail: hep@shl.co.uk

c. Rolls-Royce plc, ML-77, PO BOX 31, Derby, DE24 8BJ, UK, e-mail: ian.care@rolls-royce.com

ABSTRACT

Previous work has shown that electrostatic charge signals can be used to detect the onset of wear in lubricated tribocontacts. Preliminary investigations have shown the viability of this system when tested on a laboratory-based pin-on-disc rig, a reciprocating laboratory wear rig [1-3] and in an FZG gear scuffing rig [4,5]. These preliminary experiments have indicated several charging mechanisms could be involved within highly stressed lubricated contacts, namely tribocharging, exo-emissions, surface charge variations and debris generation [6,7].

This paper details further studies looking at the levels of charge generated by lubricated hybrid (silicon nitride sliding on bearing steel) contacts. Correlation between the charge levels generated within hybrid contacts under fully lubricated conditions and the resulting wear rates is investigated. Results from the Pin-on-Disc tests for silicon nitride against various bearing steels contacts, lubricated by Mobil Jet II, are presented and compared to steel on steel contact performance. The materials are ranked based on their performance during the wear tests. The wear mechanisms of hybrid contacts are discussed based on the on-line multi-sensor monitoring results and the post-test wear surface observation by 3-D laser profilometry measurements.

Keywords: Wear, electrostatic sensing, ceramic, bearing steel.

1. INTRODUCTION

The outstanding properties of silicon nitride have made it very attractive to a variety of industries, such as automotive, mineral processing, metal processing and oil field. The key properties of silicon nitride are its high strength over a wide range of temperature, high fracture toughness, high hardness, good wear resistance, and low density. Thus, silicon nitride has been proposed for use in very high temperature rolling element bearings on aircraft engines [8]. Silicon nitride materials with sufficient life and reliability, have been developed for rolling bearings. Hybrid ceramic ball bearings, with Si_3N_4 balls and steel rings, have been used in machine tool spindles under high speed operation for approximately 10 years [9]. Yui et al [10] have successfully tested hybrid ball bearings under high temperature (300°C) and high speed (3.5 million DN) conditions. The results confirmed that the bearings function normally with tripentaerythritol ester oil at temperatures that exceed the limiting temperature of conventional ester-based gas turbine engine oils. It was also confirmed that the sintered powder carburising steel was well suited to bearing applications in which present levels of temperature and speed are exceeded.

The wear mechanisms of ceramics have also been widely discussed and several mechanisms have been suggested, which include microcracking [11], fracture [12,13], tribochemical reactions [13,14], transgranular and intergranular cracking [15] and polishing wear [16].

The wear states of ceramic are characterised by mild and severe conditions [17]. For mild wear, tribochemical reactions take place and mechanically form fine wear debris. The debris is normally smaller than the grain size by one tenth or less. The wear rate remains below 10^{-6} mm^3/Nm. For severe wear, debris are formed by mechanical cracking of the grains or the delamination of tribo-films (formed in the contacts) of the scale of grain size. Wear rates in this regime are larger than 10^{-6} mm^3/Nm.

By simulating loading conditions in the Si_3N_4 hybrid ball bearing system, a wear process was proposed [11]. Wear in the Si_3N_4 bearing ball was to be initiated by contact fatigue stresses. The generated debris are subsequently suspended in the lubricant and create a second wear mechanism, that of three-body abrasive wear; i.e. scratching or cutting in both Si_3N_4 and bearing steel race surfaces. Large pieces of debris in the lubricant act as sharp indenters and create larger local inelastic deformation regions in the ball or race surface due to their indentation. The lateral cracks from this process could be initiated underneath the free surface due to local inelastic deformation. Therefore, the third wear mechanism spallation starts to join in the wear process.

Considerable research has been devoted to the coefficient of friction for the hybrid contacts, and it was found that the coefficient of friction for lubricated hybrid contacts is normally under 0.1, and for dry hybrid contacts it is between 0.1 – 1.0 [17-19].

This study is concerned with the condition monitoring of hybrid contacts by electrostatic sensing technology. This technology was developed originally for the detection of the electrostatically charged debris present in the gas path of turbine engines. The electrostatic charge levels in the exhaust gas changes when the amount of charged debris changes. This change can be detected by installing electrostatic sensors in the gas path and thus is a non-contact detecting system [3].

The electrostatic sensing technology is based on the assumption that a change of the static charge levels in the system reflects certain changes in conditions of the dynamic system. The technology and charge theory has been overviewed in [7]. The fundamental principle of the technique is to detect electrostatic charge associated with tribocharging, exo-emissions, surface charge variations and debris generation. The theory behind tribocharging can also be found in [20,21]. The advantages of using this technique compared with conventional condition monitoring technologies are:

a) It uses a direct measurement of electrostatic charged debris produced by engine faults rather than secondary effects such as excessive vibration and/or temperature exceedances [2];
b) It is sensitive to most material types including ferrous, metallic and non-metallic such as ceramics [2, 22].

Tasbaz et al [1] have investigated the use of electrostatic sensors for early detection of scuffing in oil lubricated metal/metal contacts. Electrostatic events were detected prior to severe scuffing and this work demonstrated the potential of using electrostatic detection technology as an on-line machinery condition monitoring system.

In this study, the application of the electrostatic sensing technique has been applied to detect wear of hybrid contacts under lubricated pure sliding conditions. The main objective was to seek correlation between the specific wear rates of the contact materials and the electrostatic charge levels. The other objective was to compare the wear performance of a range of advanced bearing steels against silicon nitride under lubricated sliding tests on the laboratory pin-on-disc tribometer.

The wear mechanisms of silicon nitride and the bearing steels are also discussed. The tests were conducted on a fully instrumented test-rig capable of on-line disc wear track depth and profilometry, as well as contact temperature monitoring.

2. EXPERIMENTAL SET-UP AND TEST CONDITIONS

Lubricated Si_3N_4 balls were loaded against various bearing steels under pure sliding contacts on a Pin-on-Disc (PoD) wear test rig. The performance of the various Si_3N_4 balls on steel contacts was compared to the baseline case of M50 bearing steel sliding against itself.

• Pin-on-disc tribometer

Figure 1 is a schematic of the pin-on-disc test-rig. The tested pins (both M50 and Si_3N_4) were 6.35 mm diameter balls and the bearing steel discs were all 100

mm diameter and 14 mm thick. The balls were tested in the 'as received' state while the discs were ground and lapped (diamond) before testing.

To enable an extensive investigation of the wear, as well as the charge generation, additional instrumentation was installed including:

a. A tachometer: monitoring the disc rotating speed;
b. A button type electrostatic sensor: connected to a charge amplifier (ENDEVCO 133), monitoring the charge level on the wear track;
c. A strain gauge: monitoring the friction between the pin and disc;
d. An LVDT (Linear-Variable-Differential-Transformer): a displacement sensor monitoring the vertical displacement of the pin towards the disc;
e. A laser probe: monitoring the depth and profile of wear track on the disc. (Accuracy ±0.2μm, laser spot size of 65 μm diameter);
f. An infrared thermometer: monitoring the disc surface temperature (close to the exit of the contact point).

The positions of these sensors on the rig are illustrated in Figure 1.

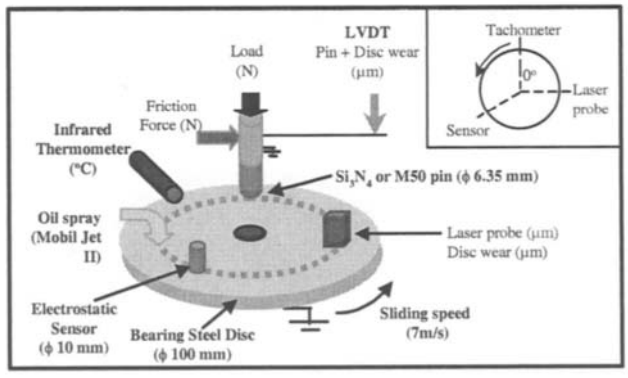


Figure 1. Schematic diagram of the laboratory pin-on-disc tribometer

- **Test materials**

The materials tested are as follows:

a. Ball materials: M50 and Si_3N_4 (commercially available);
b. Disc materials: M50 plus six other bearing materials are described in Table 1.
c. The lubricant tested was Mobil Jet II – a synthetic aircraft engine oil with viscosity of 25.3 cSt @ 40°C and 5.0 cSt @ 100°C.

Table 1 A list of the bearing steels tested on the pin-on-disc tribometer

Bearing steel	**Type/hardening details**	**Hardness ($Hv_{0.1kg}$)**
M50	VIM-VAR	784
M50 NiL	VIM-VAR, Case hardened	793
M50 DH	VIM-VAR, Duplex-hardened (through hardened and plasmanitrided)	1040
M50 NiL DH	VIM-VAR, Duplex-hardened (Case hardened and plasmanitrided)	911
XT1	VIM-VAR, case hardened stainless steel	700~800
XT2	Through hardened nitrogen-alloyed stainless steel	777
XT3	Through hardened PM high speed steel	700~800

Table 2 Details of the wear test conditions (100% sliding)

Condition	**Values**	
Sliding velocity (m/s)	7	
Maximum load (N)	60	120
Duration under maximum load (minutes)	30	90
Sliding distance (m)	12600	37800
Lubricant	Mobil Jet II	
Temperature (°C)	Ambient	
Humidity (%)	40-70	
Maximum Hertzian contact stress (GPa)	2.50 for M50/steel 2.83 for Si_3N_4/steel	3.56 for Si_3N_4/steel
Initial lamda (λ)	2.4~3.3	2.3~3.5
Surface roughness (μm)	0.06 for M50 and Si_3N_4 balls 0.1-0.2 μm for all discs	

- **Test conditions and test programme**

All the pin-on-disc tests were carried out at ambient temperature and with relative humidity of 40-70%.The tests started with no load (no contact) and no lubrication. After a few minutes running, the lubricant was sprayed on to the disc before contact allowing entrainment in to contact. The rig was run under this condition for approximately two minutes then an initial load of 10 N was applied. The load was then gradually increased to the maximum load. The rig would run for a certain period of time under the maximum load. Further details about the test conditions are listed in table 2. The disc rotation speed was varied to maintain the sliding velocity at 7 m/s for all tests. This speed was chosen to replicate the sliding conditions experienced within the outer parts of the ball/race contact, which for aircraft bearings vary between 7 and 20 m/s. Although the contact within the hybrid bearing normally experiences sliding and rolling, under low bearing loads pure sliding can occur. However, pure sliding conditions were chosen here primarily to accelerate the tests and to generate wear debris to investigate the sensitivity of the electrostatic sensor technology to these wear rates and mechanisms.

A load of 60 N was applied to the M50 on steel as well as the silicon nitride on steel contacts. This particular load was chosen to achieve a contact pressure of 2.5 GPa for the steel/steel and over 2.50 GPa for the Si_3N_4/steel. A higher load of 120 N was applied to the Si_3N_4/steel combinations to differentiate them due to the low wear rates generated at 60 N loads.

- **Wear surface observation**

After the wear tests, the worn surfaces of both the pin and disc were observed under an optical microscope and a 3-D laser surface measurements were performed to further investigate the wear mechanisms. The volume losses of both pin and disc were calculated by the measurements of wear scar diameter and wear track area.

3. DATA ANALYSIS METHODS

On-line monitoring signals were processed to look at the instantaneous changes of charge, pin/disc linear wear, friction and temperature. Correlation between the wear and charge levels, disc wear track depth, and temperature were investigated. The data acquisition and signal processing layout is shown in Figure 2. Amplifiers were used for the signals from the electrostatic sensor, the laser probe, the strain gauge and the LVDT probe. A sampling rate of 8 kHz was used for all channels. A four-channel oscilloscope was used to visualise the on-line readings from the tacho, electrostatic sensor, laser probe and temperature probe. A DAT recorder was used to record the raw data as a backup.

Post-test disc wear tracks were measured with a 2-D Talysurf profilometer. For each wear track, four measurements were carried out and the average of the four measurements was used to calculate the disc volume loss, see equation (1). Pin wear scars were measured under optical microscopy. The average of two diameters (the largest and the smallest diameter across the wear scar) was used to calculate the volume loss of the pin [23], see equation (2) while the height of the worn pin can be

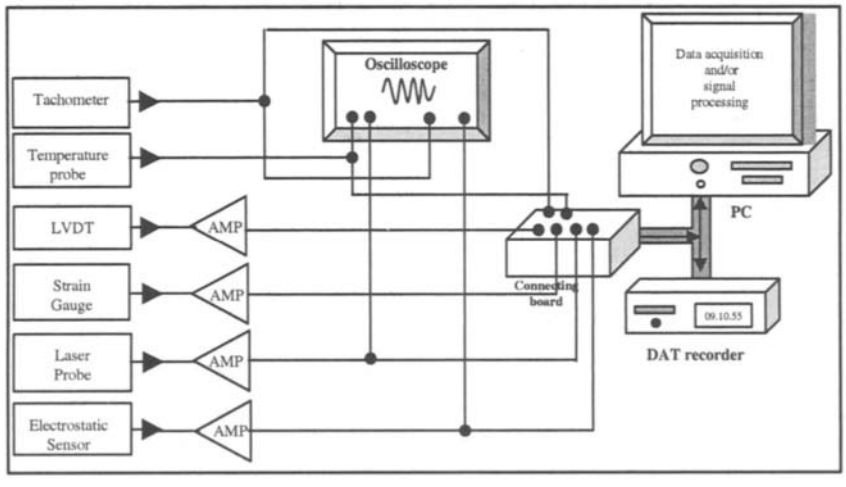


Figure 2. Block diagram of the data acquisition system for the pin-on-disc tests

calculated according to equation (3). Thus the specific wear rates of the contact pair were then estimated and the performance of the combinations was ranked accordingly, see equation (4).

$$VL_{disc} = \pi \times D \times A \quad (1)$$

$$VL_{\text{pin}} = \frac{\pi \times h^2 \times (3 \times R - h)}{3} \quad (2)$$

$$h = R - (R^2 - \frac{d^2}{4})^{\frac{1}{2}} \quad (3)$$

$$SWR = \frac{VL\,(mm^3)}{F\,(N) \times SD\,(m)} \quad (4)$$

Where D is the track diameter; A is the average track cross section area; R is the radius of the pin; h is the height of the worn pin. VL is the volume loss of disc or pin, F is the load, SD is the sliding distance and SWR is the specific wear rate.

4. RESULTS

Wear tests were carried out to compare the performance of the steels, as well as validation of the electrostatic sensing technique for the hybrid contacts. Correlation between the simultaneous specific wear rate, the wear of the disc, charge levels in the contacts, friction coefficient, and the temperature close to the contacts was found for both steel on steel and silicon nitride on steel lubricated contacts.

4.1 Instantaneous Monitoring Results for the Si_3N_4/M50 Lubricated Contact

- **Electrostatic sensor responses and the specific wear rate**

The results from the wear test of Si_3N_4/M50 contact under maximum load of 120 N with Mobil Jet II lubrication are shown in Figure 3.

From the top (a) to the bottom plot (e), the traces show the change of temperature, friction coefficient, electrostatic sensor response, the laser probe reading and the specific wear rate of the contact pin and disc versus time respectively. It should be noted that, the time axis started from 43 minutes when the constant load of 120 N (max load) was applied. The loading period is not shown in this figure.

It can be seen that, under the constant load of 120 N, the specific wear rate decreased gradually until the 68th minute, see plot (e), this is possibly due to the running-in 'polishing' wear mechanism [16]; then

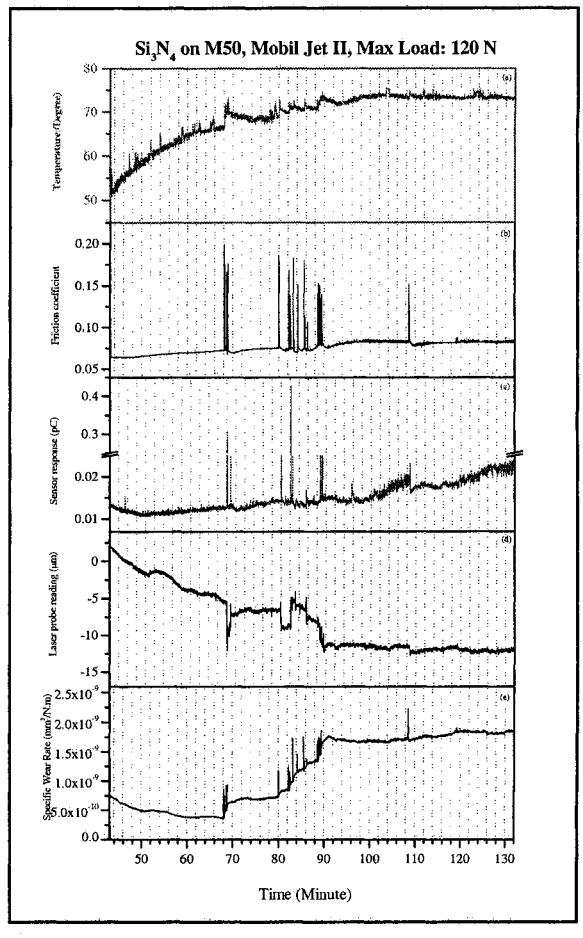


Figure 3. The test results for Si_3N_4/M50/120N contact with Mobil Jet II lubrication: (a) Contact temperature (b) Friction Coefficient (c) Electrostatic sensor response (d) Laser probe reading and (e) Specific wear rate vs. time

there are some sudden disturbances followed by a steady increase in SWR for about 12 minutes before another period of disturbances. After the 90th minute the SWR becomes steady again.

A similar pattern occurred on the electrostatic charge plot (c). While there are lot of activities, some fairly high charge levels were generated between 80 and 90 minutes as well. The increased charge levels could be caused by the increasing amount of charged debris or other sources such as tribocharging, triboemission, and contact potential difference [24]. A clear correlation between the charge level and the specific wear rate can been seen by comparing the two lines in Figure 3 (c) and (d) respectively.

Figure 3(d) shows the results from the laser probe. It shows the change of the wear track depth/profilometry during the test. While the track deepens, the value in the plot decreases. At the 68th minute, there was a large drop of the laser probe response while the wear rate increased significantly. This could be the result of the silicon nitride ball spallings and increasing cutting of the disc surface. Subsequently, the wear track was filled with the fine rubbing debris, indicated by a decrease in the wear track depth. After about a further 12 minutes steady running, the depth of the track suddenly increased again. In this case, a longer recovering time was required due to possible larger size/amount of pin debris generated. The three-body wear mechanism is likely to be involved.

- **Coefficient of friction**

From Figure 3(b), the coefficient of friction for the Si_3N_4/M50 lubricated contact by Mobil Jet II kept at a low level of less than 0.075 apart from some high level peaks (up to 0.2) appearing at the stages where large quantities of wear occurred. When 'large' and hard silicon nitride debris are generated, cutting wear is likely to be involved. Rougher surfaces will then be generated, which will increase the friction between the contacts. The relatively large increases of coefficient of friction indicate that the wear became sudden and big. Even though, the friction between silicon nitride and steel is still very low [19].

- **Disc surface temperature monitoring**

The disc surface temperature during test was monitored with an infrared thermometer. The results from the Si_3N_4/M50 lubricated contacts under 120 N maximum load is shown in Figure 3 (a). Apart from the fact that the overall temperature increases during the test, there are some sudden discontinuities/increases when the wear suddenly increases. However, the rate of the temperature increase slowed down while the worn surface was healing until the next sudden wear event happened. Some correlation between the surface temperature and the wear rate can be clearly seen. The overall increase of the temperature for this test was from ambient to about 70 °C.

4.2 Instantaneous Monitoring Results for the Si_3N_4/M50 NiL Lubricated Contact

Similar results for the Si_3N_4/M50 NiL lubricated contact are shown in Figure 4. Comparing with the

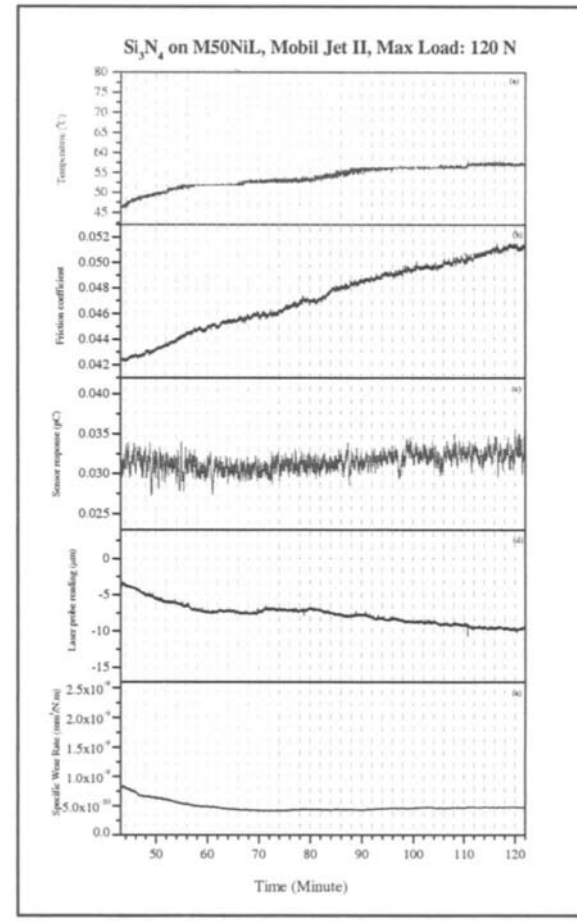


Figure 4. The test results for Si_3N_4/M50NiL/120N contact with Mobil Jet II lubrication: (a) contact temperature (b) friction coefficient (c) electrostatic sensor response (d) laser probe reading (e) specific wear rate vs. time

results of Si_3N_4/M50, the wear rates remain low for the whole test duration.

The laser probe reading shows that the wear track has not changed very much (about 5 μm) and there was no abrupt 'drops' or 'recoveries'.

The charge level changes very little during the test (see Figure 4 (c)). The overall charge level change was much smaller than that in Figure 3 (c). The random bipolar fluctuations seen on the plot are above background noise levels and indicate small differences in charge.

For the coefficient of friction (see Figure 4 (b)), it changed from about 0.042 up to 0.052, which is also lower than the friction coefficient level for Si_3N_4/M50 sliding contact. It is noted that the friction levels increase almost linearly with time during the 37.8 km test. Extended tests are planned to determine if stable friction levels are reached at larger sliding distance.

The temperature increases from about 45 °C to 57 °C (see Figure 4 (a)), which is approximately half the temperature rise seen in the higher wear contact (Si_3N_4/M50).

4.3 Bearing Steels Wear Performance Ranking

The post-test specific wear rates of the test combinations were calculated and the results are shown in Table 3. The results are also plotted in Figure 5.

It has shown that, the wear rates for the M50 sliding on bearing steel were at least 2 orders of magnitude higher than Si_3N_4 sliding on bearing steel except the Si_3N_4/XT1 and Si_3N_4/XT2 contacts under 60 N; pin wear rates were at least an order of magnitude lower than the associated disc wear rates. The SWR for M50 pin on steels lubricated sliding contacts are in the order of 10^{-7} mm^3/Nm for the pin and 10^{-6} ~ 10^{-5} mm^3/Nm for the discs except the M50/XT3 combination; for Si_3N_4 on steel lubricated sliding contacts in the order of 10^{-11} ~ 10^{-9} mm^3/Nm for the pin and 10^{-9} ~ 10^{-7} mm^3/Nm for the discs except XT1 and XT2. The rankings of the materials under the same pin material and maximum load conditions are shown in Table 3 according to the specific wear rates. The best (Rank 1) combinations among the tested materials are the Si_3N_4/M50NiL DH under 60N load and Si_3N_4/M50NiL under 120 N load with Si3N4 pin. The best combination for M50 pin is the M50/XT3 under 60N load.

It is noted that Si_3N_4/M50NiL DH performs relatively badly under higher load. This was against expectations but emphasises the sensitivity of the ranking of material combinations to the load conditions

Table 3 Specific wear rate of the wear tests and the ranking of the combinations

Combinations (Pin/Disc)	Load (N)	SWR-Pin (mm^3/N.m)	Rank	SWR-Disc (mm^3/N.m)	Rank
M50/M50	60	3.25×10^{-7}	2	1.85×10^{-5}	3
M50/M50 NiL	60	3.52×10^{-7}	3	3.07×10^{-5}	4
M50/XT3	60	3.59×10^{-7}	4	5.54×10^{-7}	1
M50/XT1	60	5.57×10^{-7}	5	6.90×10^{-5}	5
M50/XT2	60	2.15×10^{-7}	1	4.80×10^{-6}	2
Si_3N_4/M50	*60*	*3.23×10^{-9}*	*6*	*2.02×10^{-7}*	*4*
Si_3N_4/M50 NiL	*60*	*1.77×10^{-10}*	*2*	*2.49×10^{-8}*	*2*
Si_3N_4/M50 DH	*60*	*2.73×10^{-10}*	*3*	*8.69×10^{-9}*	*3*
Si_3N_4/ M50 NiL DH	*60*	*3.97×10^{-11}*	*1*	*3.08×10^{-9}*	*1*
Si_3N_4/ XT3	*60*	*1.68×10^{-9}*	*4*	*3.08×10^{-9}*	*1*
Si_3N_4/XT1	*60*	*1.90×10^{-8}*	*5*	*4.17×10^{-6}*	*5*
Si_3N_4/XT2	*60*	*1.08×10^{-7}*	*7*	*2.33×10^{-5}*	*6*
Si_3N_4/M50	120	4.59×10^{-9}	4	3.50×10^{-7}	4
Si_3N_4/M50 NiL	120	4.55×10^{-11}	1	$<1.05\times10^{-9}$	1
Si_3N_4/M50 DH	120	2.05×10^{-9}	3	5.54×10^{-8}	3
Si_3N_4/M50 NiL DH	120	8.55×10^{-9}	5	5.32×10^{-7}	5
Si_3N_4/XT3	120	3.86×10^{-10}	2	1.16×10^{-9}	2

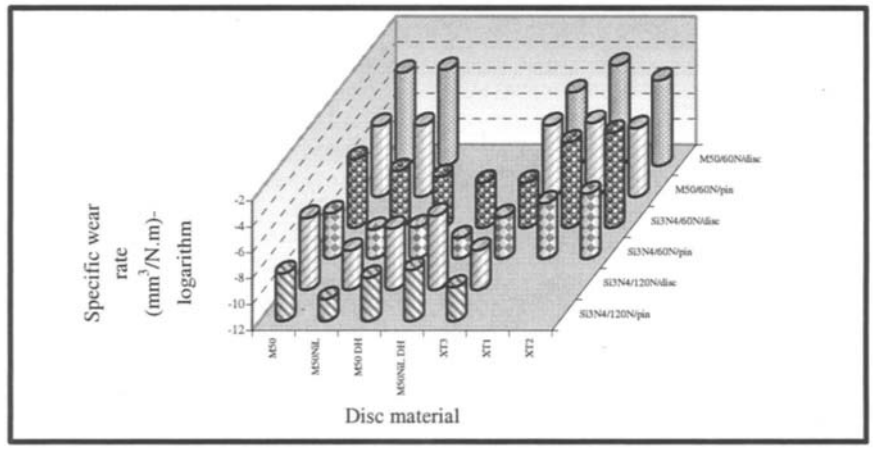


Figure 5. The specific wear rates of pins and discs under different loads for M50 or Si_3N_4 on bearing steel sliding contacts with Mobil Jet II lubrication

4.4 Inspection of the wear surfaces

3-D laser measurement has been used this study to analyse the worn surfaces of the pin and disc after being tested on the Pin-on-disc tribometer. This technology provides high-precision measurement and analysis of surface topography. It has been used for the evaluation of the form of aspheric lenses and the volumetric erosion of a nominally spherical surface. Images have been obtained for the worn surfaces of both disc and pin under lubricated sliding contacts of Si_3N_4 on M50 steel 120 maximum load and are shown in Figures 6 (a) & (b). Figures 6 (a) and (b) show the worn surfaces of the disc (M50) and pin (Si_3N_4) under 120N load. It can be seen that the wear track and the wear scar were both grooved. The depth of the wear track can be read from Figure 6 (a), which is about 14 μm. This result matches with the laser probe reading from Figure 3 (d). The wear track depth measured by the laser probe was also about 14 μm.

Compared with the Figures 6 (c) & (d), which shows the best combination of Si_3N_4/M50NiL under 120N, due to the very little wear on the disc for this contact (less than 1.05×10^{-9} mm^3/Nm), the disc surface actually shows 'invisible' or very low wear Figure (c). The pin also shows very little wear as well Figure 6 (d).

However, for the M50/M50 contact under 60N load, the depth of the disc wear track is about 60 μm, Figure 6 (e) and the diameter of the pin wear scar is about 2.5 mm, Figure 6 (f). Therefore, both pin and disc wear at a much higher rate than the lubricated hybrid contacts. The 3-D images from the pin and disc surfaces show the steel on steel wear mechanisms to be adhesive and abrasive grooving.

5. DISCUSSION

For Si_3N_4 sliding against steels under fully lubricated conditions, wear starts with a polishing action of the surfaces [14]. In this wear mode, the wear rate decreases slowly and smoothly (see in Figures 3 &4 (e)). With continued polishing and repeated contact, more and more micro-size particles pile up in the wear track. Larger debris are formed, which initiate the transition of wear mode from polishing to fracture of microstructure [12]. This may cause the increases of surface roughness and friction between the hybrid contacts. Fine debris generated would be deposited in the oil lubricant, fill in the wear track or stick on to the pin, and run in the contact. Large debris could be formed and three-body mechanism also influences the wear procedure.

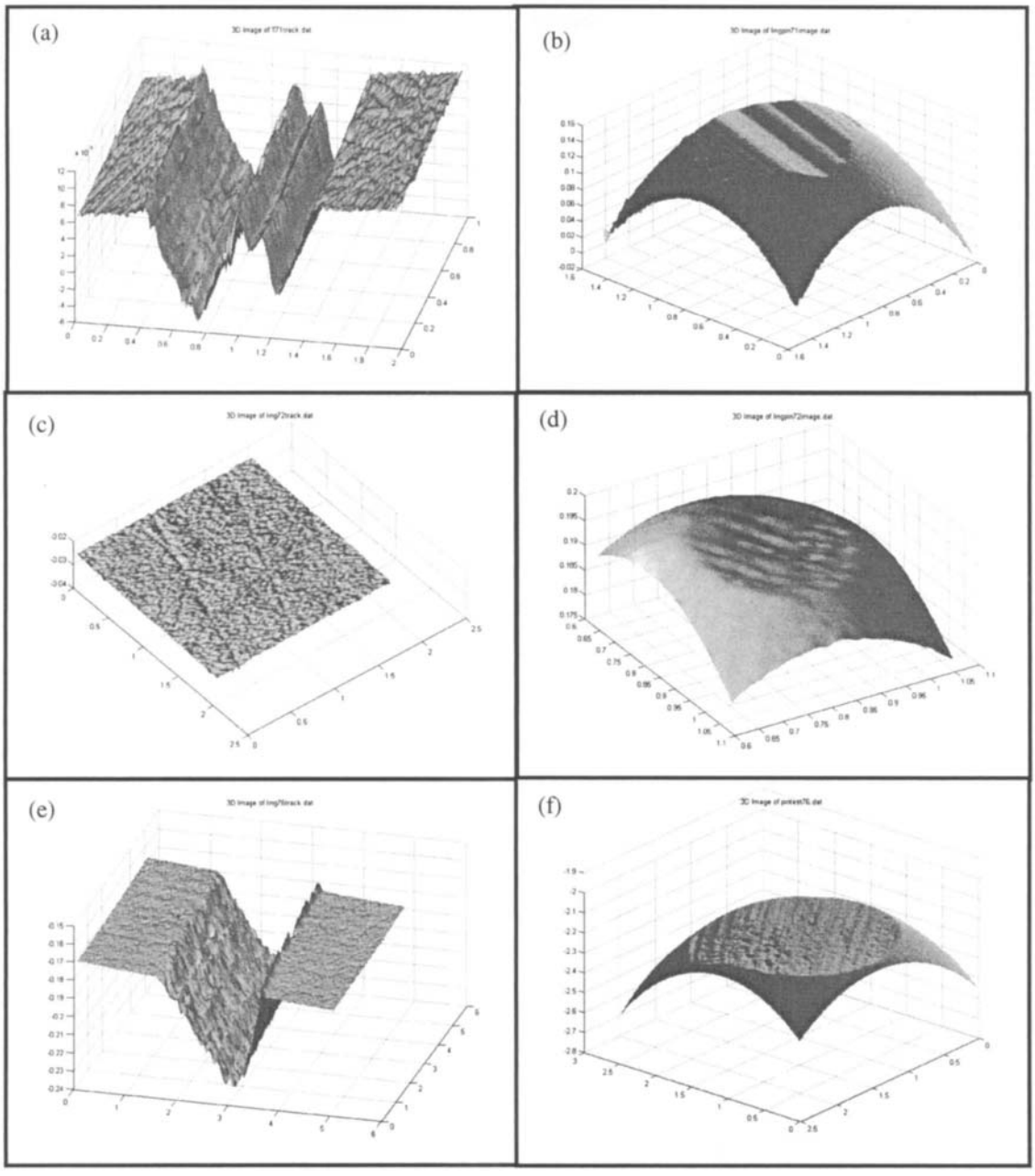


Figure 6. 3-D laser measurement images of pins (left) and discs (right) from the test combinations: (a)& (b) Si3N4/M50/120N, (c) & (d) Si3N4/M50NiL/120N, (e) & (f) M50/M50/60N. (The unit for all axes is mm)

Figure 7 shows the distribution of the debris from Si_3N_4/M50/120N Mobil Jet II lubricated sliding contact. The debris was in a large size range of between less than 1 μm up to about 50 μm. The larger debris in Figure 7 is formed by a group of small debris. 'Drops' at 68th and 80th minutes in Figure 3 (d) is possibly due to the removing of fine debris in the wear track by large debris formed in the contact. The wear track width, which can be estimated from Figure 6 (a), is over 1 mm. With a 65 μm diameter laser spot

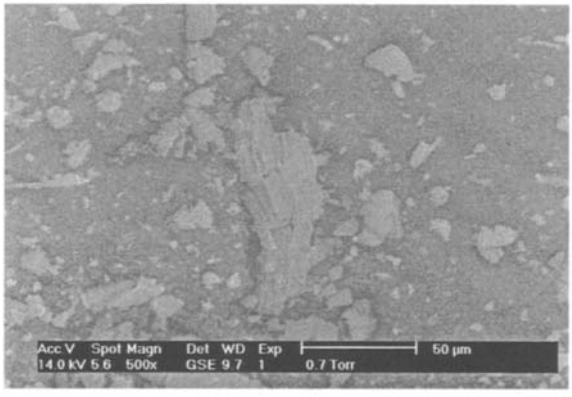


Figure 7. SEM image for debris of Si3N4/M50/120N Mobil Jet II lubricated sliding contact, sliding velocity 7 m/s

e

rotating rig, errors could also be introduced. The laser spot size is far more less than the wear track width. The measuring point might not be exactly at one specific area. This could be one of the reasons for some of the discontinuities in Figure 3 (d) and the smooth change in Figure 4 (d).

Hardness of the steels is one important factor, which affects their wear performance. Details about the relationship between wear performance under tested conditions listed in table 2 can be found in [25].

6. CONCLUSIONS

The test results have shown a very strong correlation between the wear rate of silicon nitride and bearing steel under oil lubricated contacts and electrostatic charge levels. This has established a fundamental concept for further investigation of using electrostatic sensing technology for the hybrid bearing condition monitoring. The conclusions from this study are summarised as follows.

a. Measurable electrostatic charge was generated during lubricated pure sliding wear tests for Si_3N_4 balls on bearing steels and a direct correlation was established between the wear rate and charge levels;
b. Advanced bearing steels were ranked according to their wear performance under oil lubricated sliding condition;
c. Si_3N_4 has at least 2 orders of magnitude lower wear rate than M50;
d. A correlation between the surface temperature and the wear of the contact was also indicated;
e. The wear mechanism for Si_3N_4 on steel under lubrication is suggested to start with polishing wear and followed by microstructure fracture when reasonable wear happened to the contacts. Typical sliding contact wear mechanisms were found for steel on steel.

ACKNOWLEDGEMENTS

The authors would like to thank the European Commission and Rolls-Royce for funding this work within the ATOS project under Framework V. Acknowledgements also go to the project partners in the work package Rolls-Royce Derby, FAG and Hispano-Suiza for supplying the test materials. Finally the authors would like to thank TaiCaan Technologies Limited for their help with the non-contact 3D laser measurements and Smiths Aerospace (ES-S) for supplying the electrostatic sensors and for their help with analysing and interpreting the electrostatic sensor data.

REFERENCES

[1] Tasbaz, O.D., Wood, R.J.K., Browne, M., Powrie, H.E.G. and Denuault, G., Electrostatic monitoring of oil lubricated sliding point contacts for early detection of scuffing, *Wear,* 230/1 (1999), 86-97.

[2] Tasbaz, O.D., Powrie, H.E.G. and Wood, R.J.K., Electrostatic Monitoring of Oil lubricated contacts for early detection of wear, *International Conference on Condition Monitoring, 12-15th April 1999.*

[3] Wood, R.J.K., Browne, M. and Thew, M.T., Electrostatic charging precursor to scuffing in lubricated contacts, *IMechE World Tribology Congress, 8-12th Sept 1997, London.*

[4] Powrie, H.E.G., Fisher, C.E, Tasbaz, O.D. and Wood, R.J.K., Performance of an Electrostatic Oil Monitoring system during an FZG gear scuffing test, *International Conference on Condition Monitoring, University of Wales Swansea, 12-15th April 1999.*

[5] Powrie, H.E.G., Wood, R.J.K., Harvey, T.J., & Morris, S., Re-analysis of electrostatic wear-site sensor data from FZG gear scuffing tests, *International Conference on Condition Monitoring, 25th-27th June 2001 also Condition Monitor, ISSN 0268/8050, Sept 2001, 6-12*

[6] Harvey, T.J., Wood, R.J.K., Denuault, G. and Powrie H.E.G., Effect of oil quality on

electrostatic charge generation and transport, *Journal of Electrostatics, 55 (2002), 1-23*

[7] Powrie, H.E.G., Wood, R.J.K., Harvey, T.J., Wang, L., & Morris, S., Electrostatic Charge Generation Associated with Machinery Component Deterioration, *2002 IEEE Aerospace Conference Proceedings, ISBN 0-7803-7232-8*

[8] Parker, R J., Zaretsky, E.V., 1975, Fatigue life of high-speed ball bearings with silicon nitride balls, *Transactions of ASME, J. Lub. Eng., 97(3) (1997), 350-357*

[9] Shoda, Y., Ijuin, S., Aramaki, H., The performance of a hybrid ceramic ball bearing under high speed conditions with the under-race lubrication method, *Tribology Transactions, 40 (4) (1997), 676-684*

[10] Yui, H., Aihara, S., Yamamoto, S., The performance of hybrid ball bearings under high-temperature and high-speed conditions, *Proceedings of the International Gas Turbine Congr,(1999)*

[11] Chen, Z., Cuneo, J., Mecholsky, J.J., Damage processes in Si_3N_4 bearing material under contact loading, *Wear, 198 (1996) 197-207*

[12] Fischer, T.E., Zhu, Z., Kim, H., Genesis and role of wear debris in sliding wear of ceramics, *Wear, 245 (2000), 53-60*

[13] Fischer, T.E., Tomizawa, H., Interaction of tribochemistry and microfracture in the friction and wear of silicon nitride, *Wear, 105 (1985), 29-45*

[14] Gomes, J.R., Miranda, A.S., Vieira, J.M., Sliding speed-temperature wear transition maps for Si_3N_4/iron alloy couples, *Wear, 250 (2001), 293-298*

[15] Chao, L.Y., Lakshminarayanan, R., Iyer, N., Transient wear of silicon nitride in lubricated rolling contact, *Wear, 223 (1998), 58-65*

[16] Braza, J.F., Cheng, H.S., Fine, M.E., Silicon nitride wear mechanisams: rolling and sliding contact, *Tribology Transactions, 32(4) (1989), 439-446*

[17] Adachi, K., Kato, K., Chen, N., Wear map of ceramics, *Wear, 203-204, (1997) 291-301*

[18] Zhao, X., Liu, J., Fischer, T.E., Effects of lubricant rheology and additive chemistry in the wear of Si_3N_4 sliding on steel, *Wear, 223 (1998), 37-43*

[19] Andersson, P., Ylostalo, O., The influence of lubrication on ceramic and steel sliding contacts, *Materials Science and Engineering, A109, (1989), 407-413*

[20] Kajdas, C., Furey, M.J., Ritter, A.L., Molina, G.J., Triboemission as a basic part of the boundary friction regime, *12th International Colloquium, Tribology 2000 – Plus, Volum 3, January11-13, 2000, 2075-2096*

[21] Kajdas, C., Tribochemistry, Tribology 2001, Scientific Achievements Industrial Applications Future Challenges, 2nd World Tribology Congress, Vienna, Austria, 3-7 September 2001, ISBN 3-901657-07-X, 39-46

[22] Powrie, H.E.G., McNicholas, K., Gas path condition monitoring during accelerated mission testing of a demonstrator engine, *AIAA Paper 97-2904*

[23] Gee, M.G., Guidelines for Unlubricated sliding wear tests: Part 1 & 2., National Physical Laboratory, 1993

[24] Morris, S., Wood, R.J.K., Harvey, T.J. and Powrie H.E.G., Use of electrostatic charge monitoring for early detection of adhesive wear in oil lubricated contacts, *Transactions of the ASME: Journal of Tribology, 124 (2002), 288-296*

[25] Wang, L., Wood, R.J.K., Powrie, H.E.G., Care, I., Wear performance of oil lubricated silicon nitride sliding against various bearing steels, *The 14th International Conference on Wear Of Materials, Renaissance Hotel, Washington DC, USA, March 30 - April 3, 2003 (Paper is accepted)*

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

HIGH TEMPERATURE ABRASION BEHAVIOUR OF Al/SiC_p COMPOSITES

Hayrettin AHLATCI[1], Ercan CANDAN[2], Hüseyin CIMENOGLU[1,]

[1] Istanbul Technical University, Metallurgy and Materials Engineering Department, 80626 Maslak Istanbul/Turkey

[2] Zonguldak Karaelmas University, Technical Education Faculty, Division of Metal Casting, 78100 Karabük/Turkey

ABSTRACT

In this study, high temperature wear behaviour of commercially pure Al matrix and Al-8 wt.% Mg alloy matrix composites reinforced with 60 vol.% SiC_p were compared. Composites were produced by pressure infiltration technique and their tribological performance were examined by abrasive wear tests at various temperatures up to 500 °C. Wear tests were carried out under normal stress of 0.7 N/mm^2 by rubbing on 125 μm Al_2O_3 abrasive grains bonded to belts. At all testing temperatures, the Al-8 wt.% Mg alloy matrix composite exhibited lower wear rate than that of pure Al matrix composite. Wear rates of the composites were almost independent from test temperatures up to 200 °C. At temperatures higher than 200 °C, the wear rate of both composites increased dramatically with increasing test temperature.

1. INTRODUCTION

Aluminium matrix composites are attractive engineering materials for manufacturing of aircraft and automotive components involving high speed rotation or reciprocating motion, such as motor blocks, pistons, cylinder liners, connecting rods, drive shafts, brake drums and discs [1-5]. Today, 10-20 vol.% Al_2O_3 or SiC reinforced composites are commercially available for such applications [1,3,5].

Recently, metal matrix composites containing a high volume fraction of reinforcement (>50 vol.%) have been developed for automobile brake applications [2,4] and for electronic appliances, such as microwave housings and heat sinks [6].

High temperature wear behaviour of Al matrix composites reinforced with 10-20 vol.% Al_2O_3 or SiC particles have been studied previously [7-13]. However, reports on the high temperature wear behaviour of the composites containing high volume fraction reinforcements (>50 %) are scarce. Therefore, this work deals with the effect of temperature on the abrasive wear behavior of Al/SiC_p composites containing about 60 vol.% reinforcement.

2. EXPERIMENTAL DETAILS

In this study, 7 mm diameter rod shaped SiC reinforced composites produced by pressure infiltration technique were examined. A schematic of the pressure infiltration apparatus used in this study is illustrated in Figure 1. A fused silica tube was used as a compact holder. The mean diameter of the SiC particles compacted in the silica tube was 23 μm. Commercial pure Al and Al–8 wt.% Mg alloy melts were infiltrated into the compacts under argon pressure of 765 kPa at a melt temperature of 750 °C. Details of the infiltration procedure were explained in [14].

Metallographic samples were prepared from longitudinal cross sections of the composites by gentle grinding and polishing. Microstructures of the composites were examined by using optical and scanning electron microscopes. The volume

fractions of the microstructural components were quantified by an image analyzer.

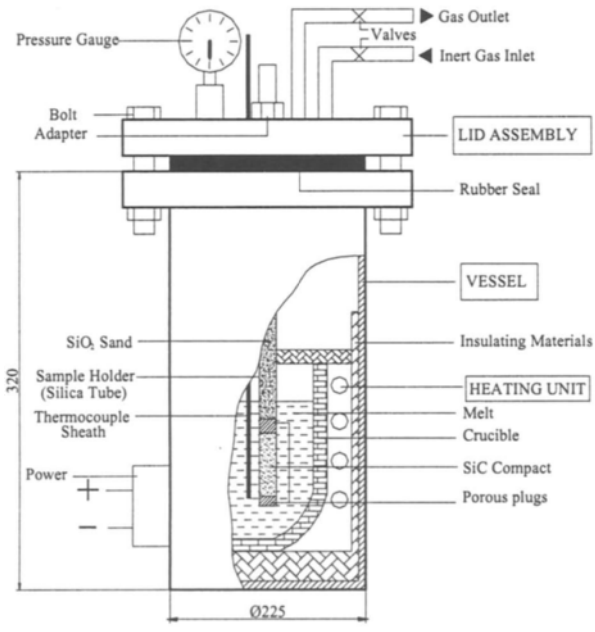


Figure 1. Schematic view of the pressure infiltration apparatus used in this study.

Mechanical properties of the composites were determined by uniaxial compression and three point bending tests. The lengths of the composite samples were 14 mm for the compressive tests, while they were 30 mm for the bending tests. Both tests were conducted on an Instron 1185 Universal testing machine with a crosshead speed of 0.5 mm/min at room temperature.

Wear tests were carried out by using the metal-abrasive configuration, shown in Figure 2, at various temperatures up to 500 °C. Samples were heated by means of an electrical resistance heater. The composites were worn under the normal stress of 0.7 N/mm^2 by rubbing on 125 μm Al_2O_3 abrasive grains bonded to belts at a sliding speed of 0.2 m/s. During the wear tests, samples that were rubbing on the abrasive belts were moving perpendicular to the sliding direction so that they always passed over fresh abrasive grains. The weight losses of the samples due to abrasion was measured for sliding distance intervals of 5.5 m with an electronic balance having a resolution of 0.1 mg. Total sliding distance was 22 m. After the wear tests, worn surfaces of the composites were examined by a scanning electron microscope.

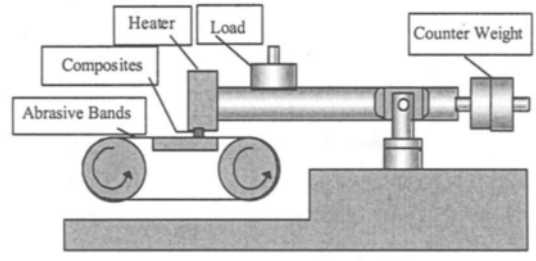


Figure 2. Schematic view of the abrasive wear tester used in this study.

3. RESULTS

Low and high magnification micrographs of the pure Al and the Al-8 wt.% Mg alloy matrix composites are shown in Figure 3 and Figure 4, respectively. The composition, in terms of volume fraction of reinforcement, matrix and pores are listed in Table1. As shown in Figure 3, pores are mainly located near the intersecting points of the SiC particles. High magnification microstructural examinations (Figure 4) depicts the presence of intermetalics in the matrix. A scanning electron microscope micrograph and EDS analysis of the intermetalics are given in Figure 5 for Al-8 wt.% Mg alloy matrix.

Table 1. Image analysis results of the investigated composites.

Materials	Volume fraction (%)		
	SiC	Matrix	Porosity
Pure Aluminum	60.2±0.7	33.7±2.1	6.1±2.7
Al-wt. % 8 Mg	59.0±4.1	41.0±4.1	0.01±0.006

The results of mechanical (compression and bending) and abrasive wear tests are presented in Figure 6 and Figure 7, respectively. The weight loss of the composites due to abrasion increased linearly with increasing sliding distance. General views of the worn surfaces are shown in Figure 8 and Figure 9.

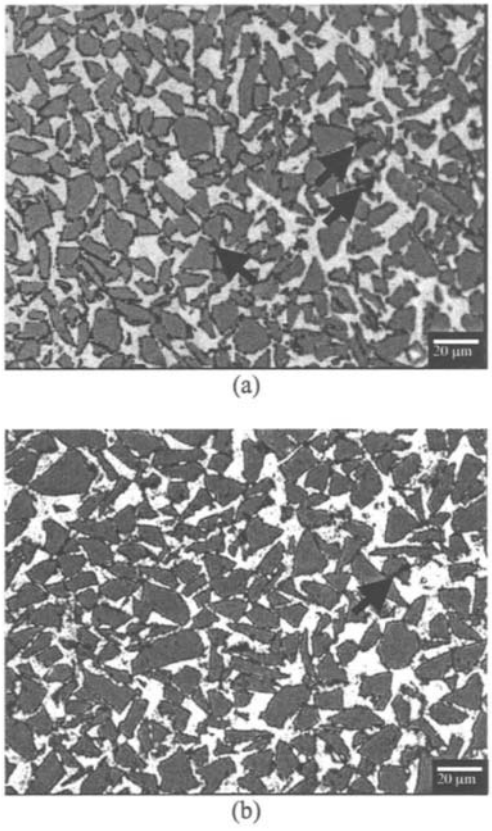


Figure 3. Low mignification optical micrographs of (a) pure Al matrix and (b) Al-8 wt.% Mg alloy matrix composites. The arrows show some of the porosities present in microstructures.

4. DISCUSSION

Microstructural examinations of both composites showed that, reinforcing SiC particles were uniformly distributed in the matrixes. Although the pure Al matrix and the Al-8 wt.% Mg alloy matrix composites had similar amounts of SiC particles (60 vol. %), the resulting porosity contents were quiet different (Table 1). Addition of 8 wt.% Mg to the matrix decreased the amount of porosity significantly as compared to pure Al matrix composite, which can be attributed to improved wetting between matrix and reinforcing SiC particles [15,16] as well as the decrease in threshold pressure [17,18]. The threshold pressure corresponds to a minimum external pressure required for beginning of penetration of liquid metal into the SiC compacts during infiltration process.

High magnification optical micrographs of the composites (Figure 4) showed that the matrix of the composites were composed of gray and dark coloured intermetalics. The results of EDS analysis given in Figure 5 revealed that the gray coloured intermetalics are rich in Si content whereas dark coloured intermetalics are rich in both Mg and Si contents. In previous studies, these types of gray colored precipitates were described mainly as eutectic Si [19,20], while the dark coloured precipitates were identified as Mg_2Si [19-22]. It should be mentioned that Mg_2Si precipitates were not observed in the matrix of the pure Al matrix composite.

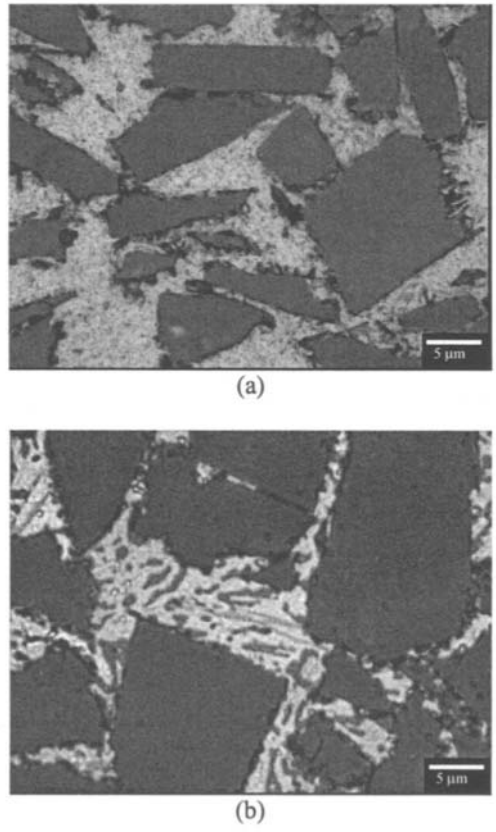
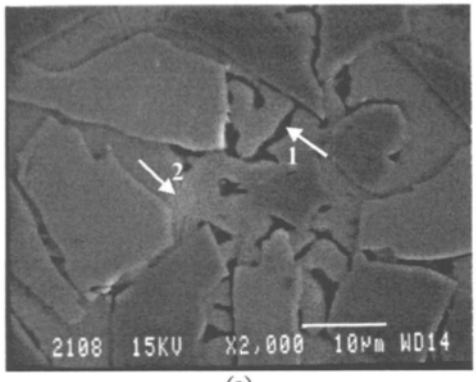


Figure 4. High mignification optical micrographs of (a) pure Al matrix and (b) Al-8 wt.% Mg alloy matrix composites.



(a)

Element (wt. %)	Intermetalic	
	1	2
Mg	17.1	-
Si	22.4	92.9
Al	60.5	7.1

(b)

Figure 5. (a) Scanning electron micrograph of Al-8 wt.% Mg alloy matrix composite and (b) EDS analysis results of phases marked as 1 and 2, which represent the dark and gray colored intermetalic present in the matrix, respectively.

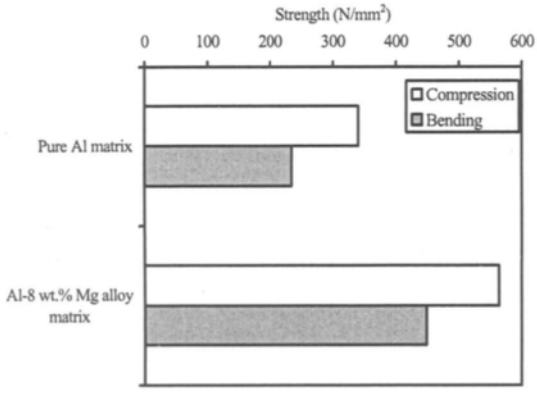
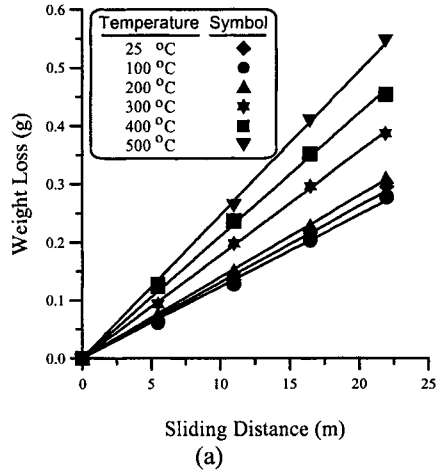


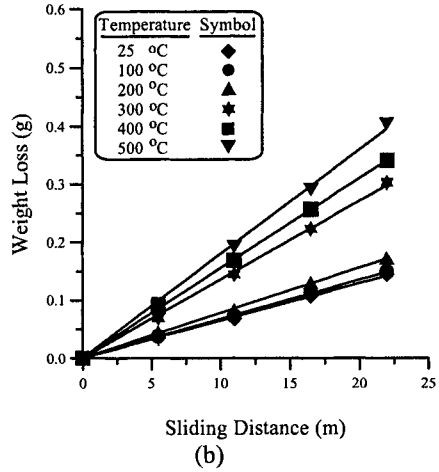
Figure 6. Room temperature strength values of the investigated composites obtained by compression and bending tests.

As illustrated in Figure 6, alloying of the matrix with 8 wt.% Mg increased the compression and bending strengths about 60 % and 90 %, respectively. The increase in room temperature strength with addition of Mg was believed to be due

solid solution hardening of the matrix as well as :cipitation of intermetalics [21,23-26].



(a)



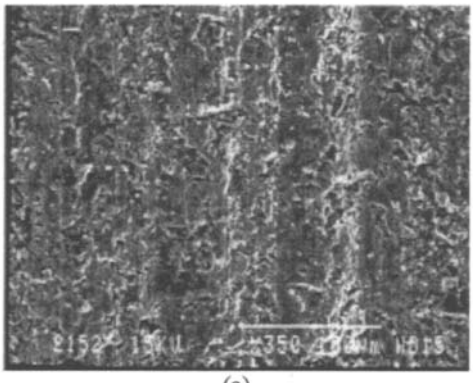
(b)

ɟure 7. The effect of sliding distance on the weight s of (a) pure Al matrix and (b) Al-8 wt.% Mg ɔy matrix composites at various temperature.

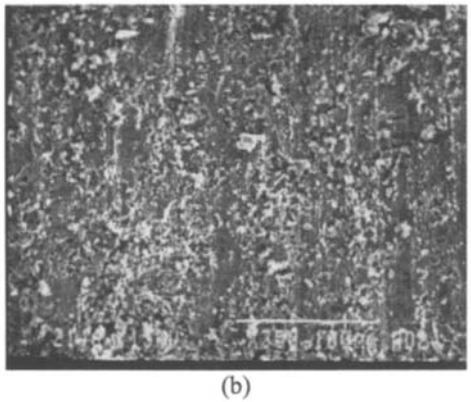
The variation in the wear rates, which were calculated from the slope of Figure 7 in the unit of g/m, are presented in Figure 10 as a function of test temperature. The addition of Mg to Al matrix resulted with an improvement of the wear resistance of the composites. Thus, the high strength composite exhibited high resistance to abrasion at ambient and elevated temperature.

Ambient temperature wear resistance of the composites was maintained up to a critical temperature of 200 °C. A sharp transition from mild to severe wear was observed above this critical

temperature (Figure 10). In agreement with present work, Singh and Alpas [8,9] have also reported results of high temperature sliding wear tests of Al 6061 alloy matrix composite reinforced with 20 vol.% Al_2O_3 particle and attributed transition from mild to severe wear to massive decrease of compression strength above 200 °C.



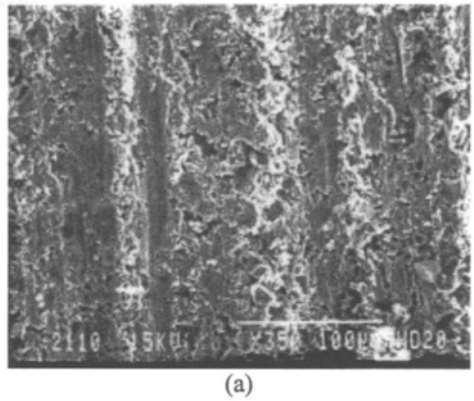
(a)



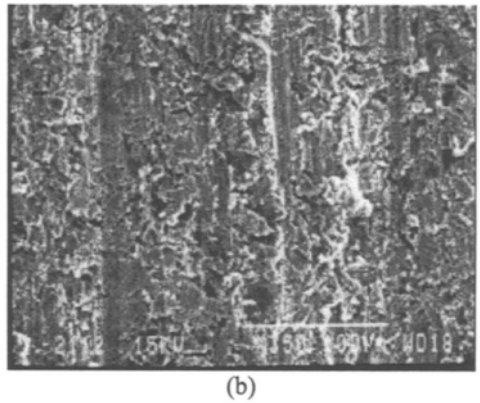
(b)

Figure 8. Worn surface scanning electron micrographs of (a) pure Al matrix and (b) Al-8 wt.% Mg alloy matrix composites abraded on 125 μm Al_2O_3 abrasive grains at room temperature.

Figure 8 and Figure 9 show the abraded surfaces of the composites tested at room temperature and 500 °C, respectively. Worn surface examinations revealed the wear mechanism as grooving. The width and depth of the grooves were smaller for the high strength Al-8 wt.% Mg alloy matrix composite than that of the low strength pure Al matrix composite. The size of the grooves increased with increasing test temperature for both composites which is due to the softening of the matrix at elevated temperature.



(a)



(b)

Figure 9. Worn surface scanning electron micrographs of (a) pure Al matrix and (b) Al-8 wt.% Mg alloy matrix composites abraded on 125 μm Al_2O_3 abrasive grains at 500 °C.

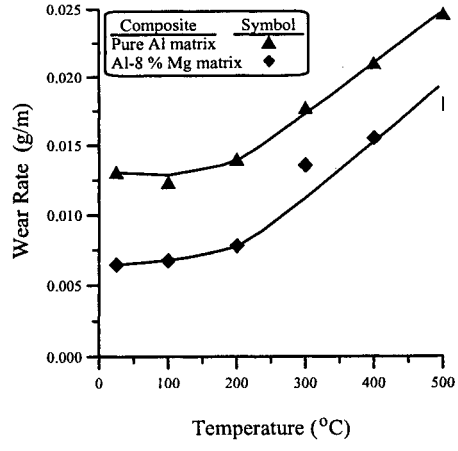


Figure 10. Temperature dependence of abrasive wear rates of pure Al matrix and Al-8 wt.% Mg alloy matrix composites.

4. CONCLUSION

The experimental results of this study have highlighted the effect of 8 wt.% Mg addition on the compression strength, bending strength and abrasion resistance of Al matrix composite containing 60 vol.% SiC.

Al-8 % Mg alloy matrix composite exhibited higher compression strength, bending strength and abrasion resistance than pure Al matrix composite. At ambient temperature, wear rates were maintained up to a critical temperature of 200 °C, above which wear resistance decreased dramatically with increasing temperature.

REFERENCES

1. P.K. Rohatgi, R. Asthana and S. Das, International Metals Reviews, v 31 (1986) 115.
2. P. Rohadgi, JOM, (1991) 11.
3. J.E. Allison, and G.S. Cole, JOM, v 45 (1993) 19.
4. J. Eliasson, and R. Sandstrom, Key Engineering Materials, v 104-107 (1995) 3.
5. V.M. Kevorkijan, JOM, November (1999) 54.
6. B. Maruyama and W.H. Hunt Jr., JOM, November (1999) 56.
7. M.A. Martinez, A. Martin and J. Llorca, Scripta Metallurjica et Materialia, v 28 (1993) 207.
8. J. Singh and A.T. Alpas, Scripta Metallurgica et Materialia, v 32 (1995) 1099.
9. J. Singh and A.T. Alpas, Metallurgical and Materials Transaction, 27 A (1996) 3135.
10. S. Wilson and A.T. Alpas, Wear, 196 (1996) 270.
11. S.Y. Yu, H. Ishii, K. Tohgo, Y.T Cho and D. Diao, Wear, 213 (1997) 21.
12. A. Martin, J. Rodriguez and J. Llorca, Wear, 225-229 (1999) 615.
13. M. Muratoðlu and M. Aksoy, Materials Science and Engineering, 282 A (2000) 91.
14. E. Candan, H. Ahlatci and H. Çimenoðlu, Wear, 247 (2001) 133.
15. T.M. Liu, and C.G. Chao, Materials Science and Engineering, 169 A (1993) 79.
16. E. Candan, H.V. Atkinson and H. Jones, Key Engineering Materials, v 127 (1997) 463.
17. J. Narciso, A. Alonso, A. Pamies, C. Garcia-Cordovilla and E. Louis, Scripta. Metallurgica et Materialia, v 31 (1994) 1495.
18. E. Candan, Phd Thesis, University of Sheffield, July 1998.
19. Y. Liu, S.B. Kang and H.W. Kim, Materials Letters, 41 (1999) 267.
20. M.S. Karaka°, M.S. Thesis, Istanbul Technical University, 2001.
21. E. E. Schmid, K.V. Oldenburg and G. Frommeyer, Z. Metalkunde, 81 (1990) 809.
22. Y.E. Lü, Q.D. Wang, X.D. Zeng, Y.P. Zhu and W.J. Ding, Materials Science and Engineering, 301 A (2001) 255.
23. D. Lüdeoke, Z. Metallkunde, 77 (1986) 278.
24. J. Paul and H.E. Exner, Z. Metalkunde, 81 (1990) 816.
25. L. Salvo, G.L. Esperance, M. Suery and J.G. Legoux, Materials Science and Engineering, 177 A (1994) 173.
26. J. Zhang, Z. Fan, Y. Wang and B. Zhou, Materials Design, 21 (2000) 149-153.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Friction and wear of spheroidal graphite cast iron under severe sliding conditions

K. Hirasata, K. Hayashi and H. Matsunami

Department of Mechanical Engineering, Osaka Sangyo University
3-1-1 Nakagaito, Daito-shi, Osaka, Japan

The dry friction and wear characteristics of spheroidal graphite cast iron under the severe conditions of high sliding speed and high contact pressure were experimentally investigated by using the pin-on-disk type test rig. The variations of the friction force, the wear amount and the temperature rise with sliding distance were experimentally measured and the relations among them were clarified. The friction and wear characteristics of spheroidal graphite cast iron pin are strongly influenced with the friction-induced temperature rise. Its wear rate changes in three steps as the increment of sliding distance because of the changes of its hardness by temperature rise (the first wear region, the second wear region and the third wear region). The wear rate of cast iron pin depends on the sliding speed and the contact pressure, and its empirical formula in each region was derived. Meanwhile, the coefficient of friction is large at the start of sliding and rapidly decreases with the increment of sliding distance in the first and the second wear regions. Finally, it converges to some constant value in the third wear region. The coefficient of friction depends on the sliding speed and decreases with its increment, but it is almost independent of contact pressure. Some empirical formulae were derived.

1. INTRODUCTION

A cast iron is widely used as brake shoe material and many investigations have been carried out on its friction and wear characteristics[1-8]. Many of them, however, have treated the phenomena under relatively mild sliding conditions, and the investigations on the friction and wear characteristics of cast iron under the severe conditions of high sliding speed and high contact pressure as seen in some kinds of emergency brakes have been not so many reported, and many things to be clarified are still remained. Under the conditions of high sliding speed and high contact pressure, the temperature rises of sliding materials induced by friction are large and the changes of properties of sliding materials by the temperature rise have strong influences on the friction and wear characteristics[9-10].

Some of the authors previously carried out the investigations in which the flake graphite cast iron was used as a pin and experimentally clarified its friction and wear characteristics under the severe sliding conditions[11-12].

Then, the dry friction and wear characteristics of spheroidal graphite cast iron under the

conditions of high sliding speed and high contact pressure at room temperature, especially their changes by the temperature rise are experimentally investigated here. For this purpose, the pin-on-disk type test rig with high rigidity has been built, and the wear tests using a spheroidal graphite cast iron pin and a mild steel disk are carried out. The friction force, the linear wear, the wear rate and the temperature rise of cast iron pin are measured. Then, the relations between them are discussed and some empirical formulae of friction coefficients and wear rates are derived.

2. APPARATUS AND PROCEDURES OF EXPERIMENTS

The pin-on-disk type test rig used in this investigation is shown in Fig.1. A cast iron pin with 6mm × 3.25mm rectangular section is fixed on the load cell and is pressed against the surface of the rotating disk by means of air cyinder. The flywheel with sufficiently large of inertia and the electric motor with sufficiently large torque are used in order to keep the rotating speed of disk constant during a test. The contact force and the friction force can be simultaneously measured by the load cell fixing the cast iron pin. The wear length of cast iron pin is measured by the eddy current type gap sensor. The temperature of the cast iron pin was measured by the fiber-type radiation thermometer on its side surface near the sliding surface shown in Fig.3. The measured temperature is the average one over the focussed circle with 1mm diameter.So,the temperature of the slidingsurface of the cast iron pin is higer than this measured one[5-6]. The dimensions and the chemical components of test pieces used here are shown in Fig.2 and Table.1.

All the tests were carried out under the conditions of room temperature and dry friction. Before starting the test, the cast iron pin and the mild steel disk were cleaned up with acetone. Then, the disk was rotated with a certain con-

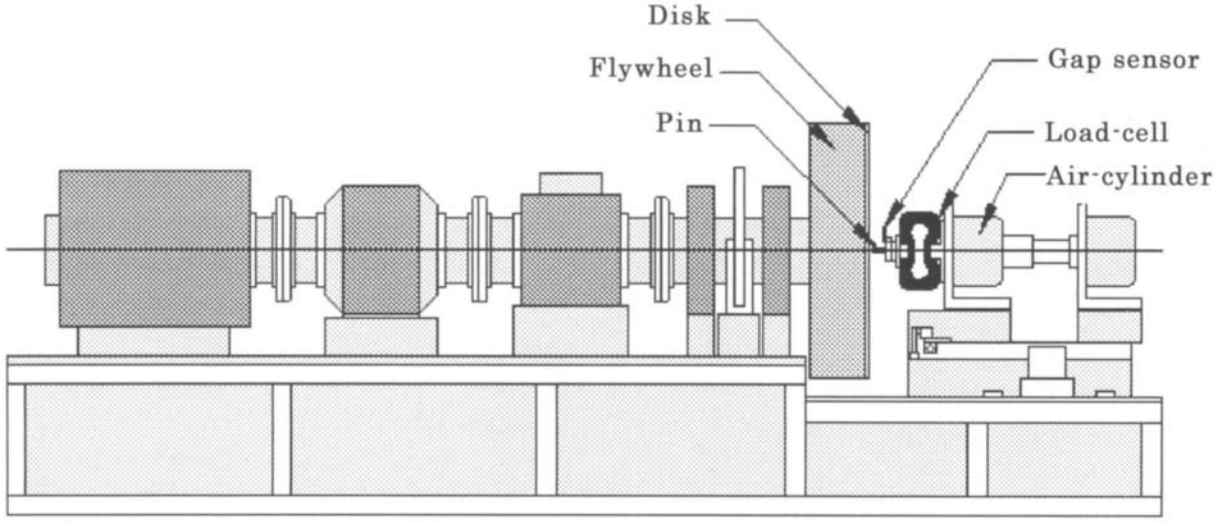


Fig.1 Experimental apparatus

stant speed and the cast iron pin was pressed against the rotating disk with some contact force by means of air cylinder. When the cast iron pin slid on the mild steel disk for some distance and its wear amount (wear length) reached to acertain value (2mm in this test), the pin was pulled apart from the rotating disk and the test was finished. The experiments were carried out in the range of contact pressure P=29.4~98.0MPa and sliding speed V=2~15m/s

Moreover, the crystalline structure in the section of the cast iron pin after the sliding test was observed by microscope in order to investigate the relation to the friction and wear characteristics.

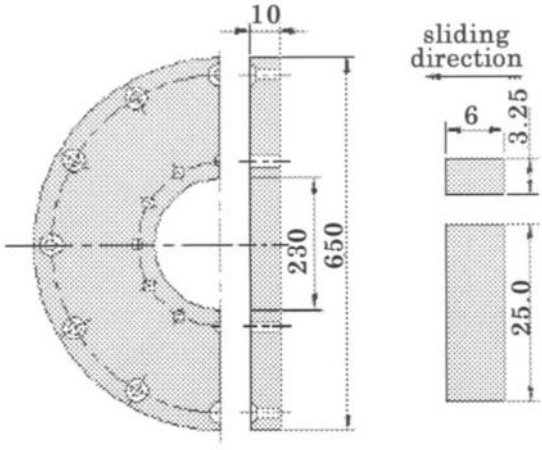


Fig.2 Dimensitions of pin and disk (Unit:mm)

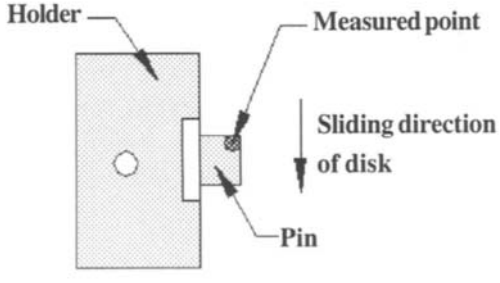


Fig.3 Mesurement of temperature

Table.1 Chemical components of pin and disk

	Pin	Disk
Material	Spheroidal graphite cast iron (FCD800)	mild steel (SS400)
Chemical components(wt%)		
C	3.880	0.140
Si	2.800	0.230
Mn	0.110	0.610
P	0.020	0.011
Si	0.007	0.005
Mg	0.039	

3. EXPERIMENTAL RESULTS

3.1. Variations of friction, wear and temperature with sliding distance

An example of the experimental results showing how the friction force, the wear amount (wear length) and the temperature of cast iron pin varied with sliding distance is shown in Fig.4.

Fig.4 shows the following things. The wear of cast iron pin was relatively moderate for a while after the start of sliding but the wear rate remarkably increased after sliding some distance, and then it decreased again and became stable. In this manner, the wear rate of spheroidal graphite cast iron pin changed in three steps with sliding distance. Then, these wear conditions are named the first wear region, the second wear region and the third wear region in order, and the transition points from the first wear region to the second one and from the second one to the third one are named the first transition point and the second one, respectively (see Fig.4).

The friction force was large at the start of sliding but decreased with the sliding distance in

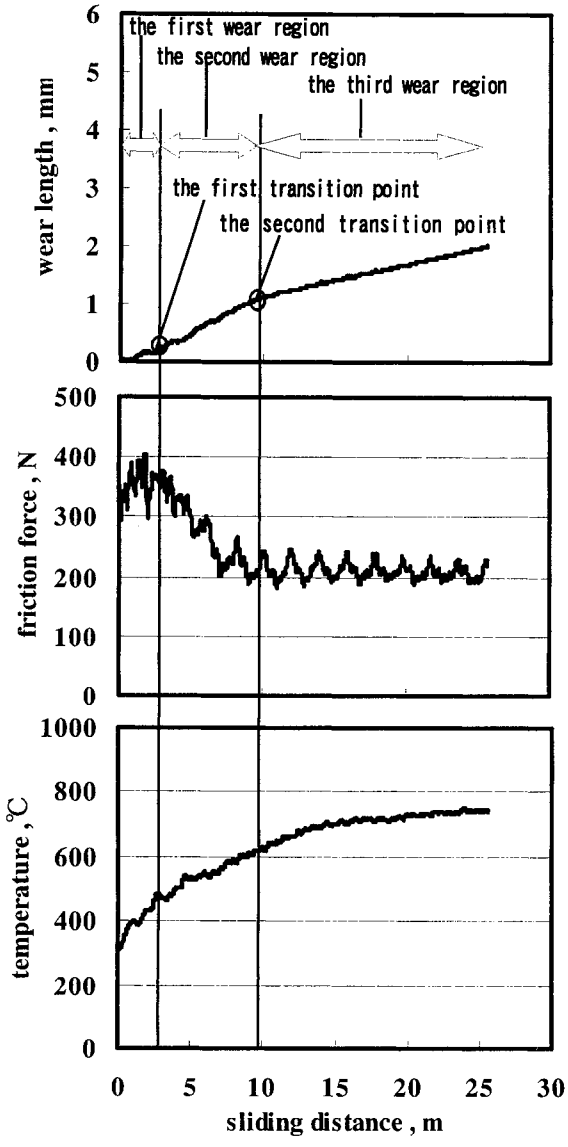


Fig.4 Variations of friction force, wear amount and temperature with sliding distance (sliding condition : V=10m/s,P=49.0MPa)

the first , and , the second wear regions, and it was settled in some constant value in the third wear region. Meanwhile, the temperature of cast iron pin rapidly rose by the frictional heat for a while after the start of sliding but the rate of temperature rise decreased with sliding diatance, and it became much moderate and settled in some constant value in the third wear region. On the other hand, the temperature rise of rotating mild steel disk was less than 10°C.

The hardness test was carried out for the purpose of clari fying the reduction of hardness of castiron pin by the temperature rise. The test result is shown in Fig.5. The hardness of the spheroidal graphite cast iron falls with the temperature rise. The rate of the fall of hardness is relatively moderal below 500°C, but it rapidly increases beyond 500°C and increases much more beyond 600°C. The hardness change of mild steel disk is also shown here.

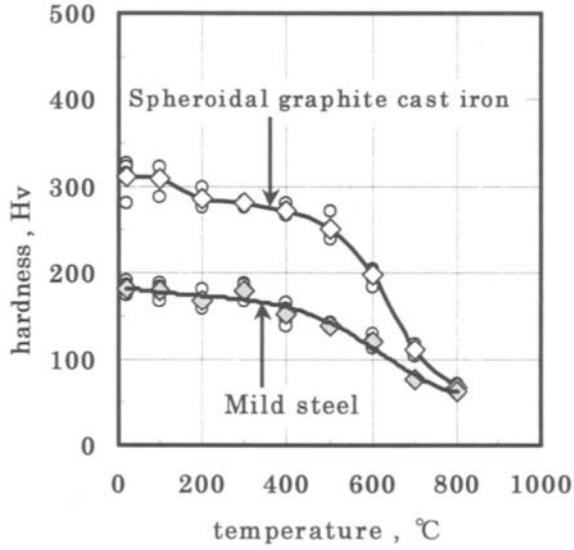


Fig.5 Variations of hardness of cast iron and mild steel with temperature

3.2. Coefficient of friction

The friction characteristics will be discussed with the coefficient of friction. The coefficient of friction was calculated as the ratio of the friction force to the contact force measured in the sliding tests. The relation between the friction force and the sliding distance is shown in Fig.4 as an example. So, the coefficient of friction also varies with the sliding distance in the same manner. Namely, the coefficient of friction decreases with the increment of sliding distance in the first and the second wear regions, and

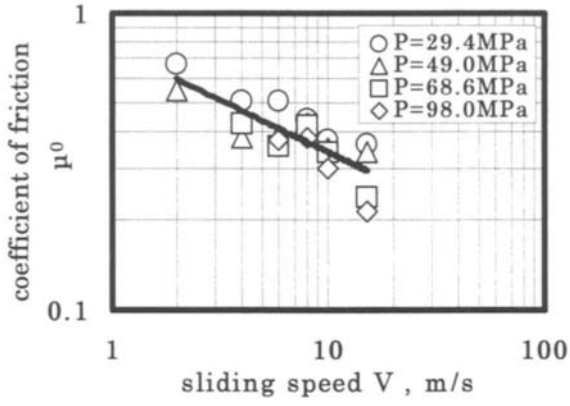


Fig.6(a) Variations of coefficient of friction (at the start point of sliding)

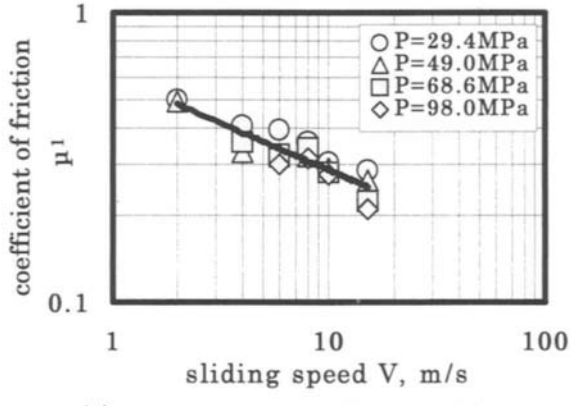


Fig.6(b) Variations of coefficient of friction (at the first transition point)

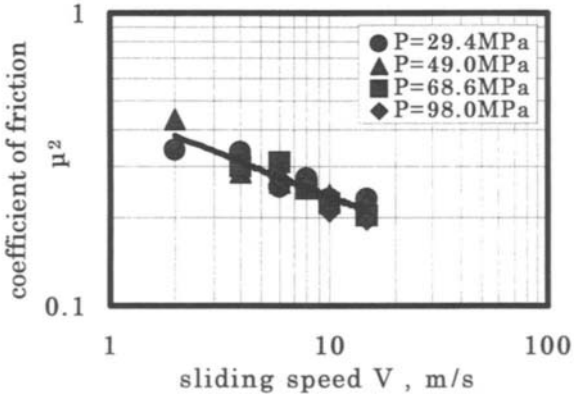


Fig.6(c) Variations of coefficient of friction (at the second transition point)

then it converges to some constant value in the third wear region. The variations of coefficients of friction at the start point of sliding, μ_0, at the first transition point, μ_1, and the second transition point, μ_2 with the sliding speed are shown in Figs.6(a), 6(b) and 6(c).

The experimental data plotted in these figures are the mean values of the plural measured data for each sliding condition. As seen in these figures, the coefficients of friction, μ_0, μ_1 and μ_2 decrease with the increment of sliding speed but are not influenced with the contact pressure. The solid lines in the figures were drawn by applying the method of least squares to the experimental data, and the following empirical formulae are derived.

$$\mu_0 = 0.76V^{-0.35} \quad (1)$$

$$\mu_1 = 0.60V^{-0.32} \quad (2)$$

$$\mu_2 = 0.46V^{-0.29} \quad (3)$$

3.3. Wear rate

Just after the start of sliding, the powdery and / or flaky wears were observed. Then, the part of the cast iron pin near the frictional surface was softened by the temperature rise with the frictional heat and it flew and attached itself to the circumference of the pin. The external appearance of the cast iron pin after sliding test is shown in Fig.7. Meanwhile, the examples of the crystalline structures in the section of the cast iron pin after the sliding tests observed by microscope are shown in Fig.8(a) and Fig.8(b). Fig.8(a) is the crystalline structure in the section of the cast iron pin for the case finishing the sliding test in the second wear region and Fig.8(b) is the one for the case finishing the sliding test in the third wear region. The flow of the crystalline structure near the sliding surface of the cast

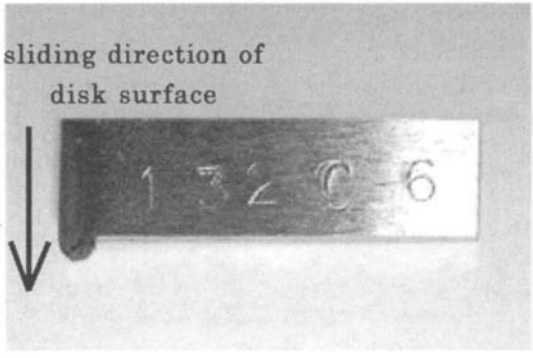


Fig.7 Cast iron pin after sliding test

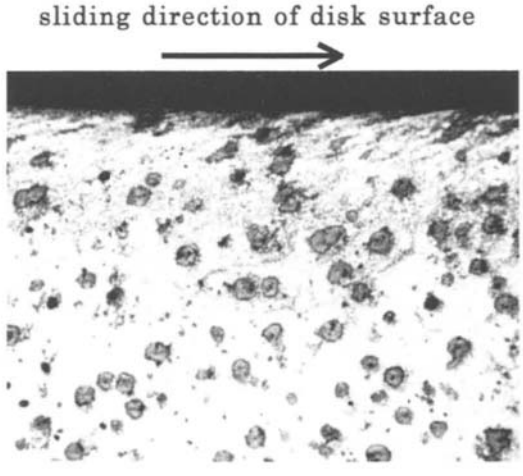


Fig.8(a) Crystalline structure of cast iron pin (in the second wear region) (V=10m/s,P=49.0MPa)

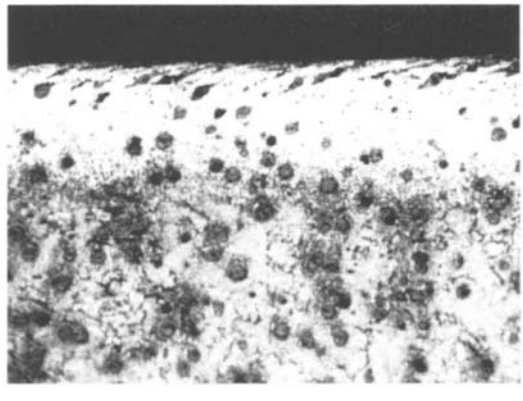


Fig.8(b) Crystalline structure of cast iron pin (in the third wear region) (V=10m/s,P=49.0MPa)

iron pin (the fluidity layer) in the sliding direction of the disk are observed , and it is seen that the thickness of the fluidity layer in the second wear region is thicker than that in the third wear region.

The specific wear amount, W, is generally used in order to discuss the wear characteristics and it is defined as

$$W = v / (F \cdot L) \tag{4}$$

where v is the wear amount in volume, F is the contact force and L is the sliding distance. It is not suitable, however, to apply this conventional specific wear amount for the purpose of estimations and discussions of wear characteristics under the severe sliding conditions such as accompanied with the thermal wear treated here. Then, the new physical value defined by Eq.(5), W_m, is introduced here[9-10];

$$W_m = dv / (F \cdot dL) \tag{5}$$

where dv is a increase of wear amount induced by the increment of sliding distance dL. This physical value, W_m, has the same dimension as the aforementioned conventional specific wear amount and is named as the modified specific wear amount. Eq.(5) can be rewritten into the following expression.

$$W_m = A(dl / dt) / [F(dL / dt)] = \dot{l} / (P \cdot V) \tag{6}$$

where t is a time and $\dot{l}=dl/dt$ is a wear rate of cast iron pin in length. It is seen from Eq.(6) that the modified specific wear amount introduced here is a ratio of wear rate in length to PV value.

An example of experimental results arranged in the relation of modified specific wear amount versus sliding distance is shown in Fig.9.As

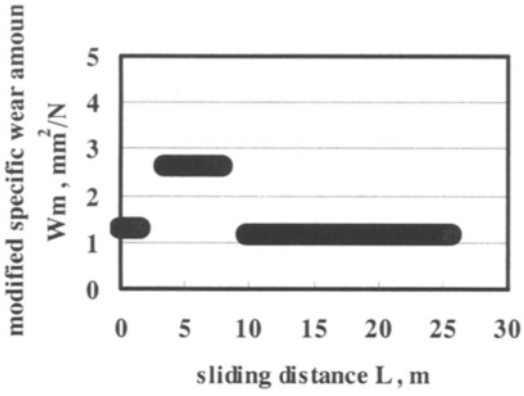


Fig.9 Variation of modified specific wear amountwith sliding distance (sliding condition V=10m/s,P=49.0MPa)

seen from this figure, the modified specific wear amount in the first, the second and the third wear regions are respectively constant in each region irrespective of sliding distance. The experimental results for other sliding conditions also showed the same tendencies.Therefore, the usage of the modified specific wear amount is convenient and useful to discuss the wear characteristics under the severe sliding conditions as treated here.

Figs.10(a), 10(b) and 10(c) show how the modified specific wear amounts in the first, the second and the third wear regions varied with the sliding conditions.The experimental data shown in these figures are the mean values of the plural test results. It is seen from these figures that the modified specific wear amount in each region, W_{mi} (i = 1,2,3), depends on the sliding speed, V, but it is little influenced with the contact pressure, P. The solid lines in these figures were drawn by applying the method of least squares to the experimental data and the empirical formula for them is as follows;

$$W_{mi} = ai \cdot V^{bi}, (i = 1,2,3), (mm^2/N) \qquad (7)$$

where $a_1 = 2.00$, $b_1 = -0.17$
$a_2 = 17.1$, $b_2 = -0.95$
$a_3 = 23.4$, $b_3 = -1.24$

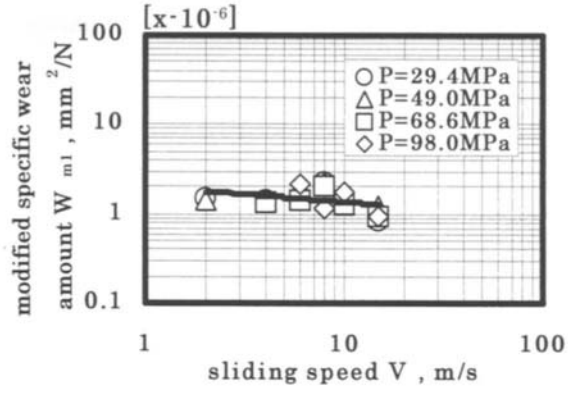


Fig.10(a) Modified specific wear amount in the first wear region

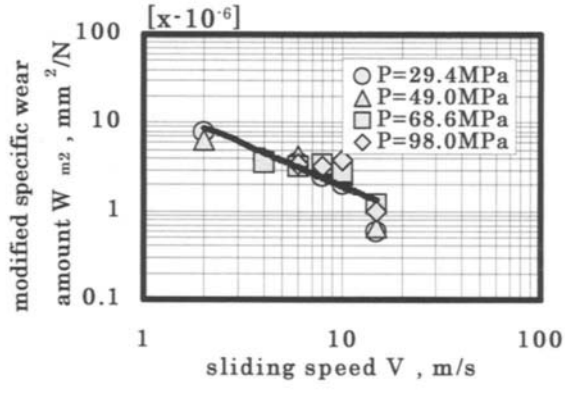


Fig.10(b) Modified specific wear amount in the second wear region

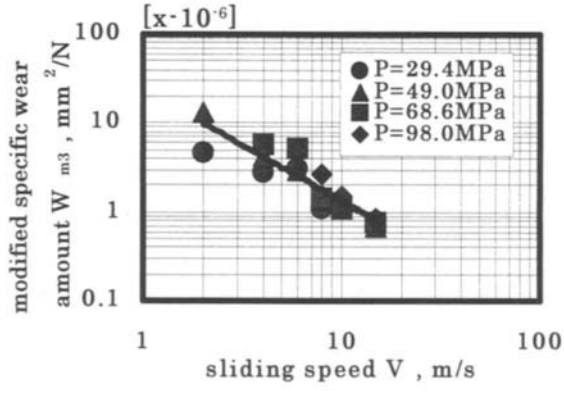


Fig.10(c) Modified specific wear amount in the third wear region

and the suffixes 1, 2 and 3 are corresponding to the first, the second and the third wear regions, respectively. From Eqs.(6) and (7), the wear rate of cast iron pin in length in each region, $\dot{l}_i$ (i= 1,2,3), is expressed as

$$\dot{l}_i = W_{mi} \cdot P \cdot V = a_i \cdot P \cdot V^{1+b}, \ (i = 1,2,3), \ (mm/s) \qquad (8)$$

where $\dot{l}_1$, $\dot{l}_2$ and $\dot{l}_3$ are the wear rates of cast iron pin in the first, the second and the third wear regions, respectively. Then, the following results are clarified from Eq.(8). The wear rates in the aforementioned three regions are proportionally increased by the increment of contact pressure, P. On the other hand, they vary with sliding speed, V, in the different manners, respectively. The wear rate in the first wear region increases with the increment of sliding speed and that in the second wear region almost independent of sliding speed, but it conversely decreases with the increment of sliding speed in the third wear region.

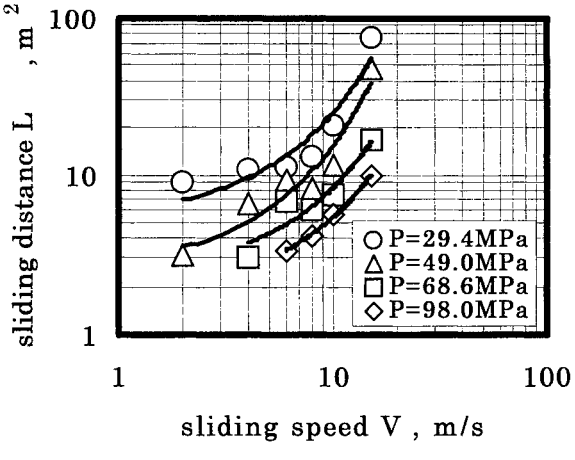


Fig.11 Variations of sliding distance with sliding speed

3.4. Sliding distance

It is interested to know how far the pin slides until the coefficient of friction and the wear rate reach steady state. Fig.11 shows how the sliding distance from the start point to the second transition point, L_2, varied with sliding conditions. From this figure, it is seen that the sliding distance, L_2, increases as the sliding speed increases and the contact pressure decreases.

4. DISCUSSIONS

The friction-induced temperature rise of cast iron pin gives the strong effects on the friction and wear characteristics under such severe sliding conditions. Just after the start of sliding, he temperature of sliding surface of cast iron pin rapidly rises by the generations of enormous frictional heat because of large coefficient of friction, but the temperature gradient of cast iron pin at right angle to its sliding surface is large and the part of pin with high temperature is limited in the one near the sliding surface because the thermal conductivity of spheroidal graphite cast iron is fairly small.So, the softened layer of cast iron pin by temperature rise is not so thick and the wear rate is relatively small (the first wear region). With the increment of sliding distance and the lapse of time, however, the enormous frictional heat, which have not sufficiently been removed by the wear in the first wear region, makes the rapid temperature rise of cast iron pin by the heat conduction and generates the thick softened layer with high temperature. As the results, the wear rate remarkably increases (the second wear region). The wear of cast iron pin in the first stage of sliding were removed as powdery and/or flaky. With the progress of softening the cast iron pin by temperature rise, the cast iron pin near the sliding surface was fluidified, and was removed from sliding surface and remained

around the pin. Meanwhile, the coefficients of friction in the first and the second wear regions decrease as the sliding distance increases because the plastic flow resistance of cast iron pin near the sliding surface is reduced by the violent temperature rise. After the thermal balance has been achieved bythe removal of enormous heat with high wear rate in the second wear region, the decrease of friction-induced heat generation by decreasing of coefficient of friction and the heat conduction, the wear becomes rather moderate and stable, and the coefficient of friction also converges to some constant value (the third wear region). As seen in Fig.4, the settlement of temperature of cast iron pin was behind them because of the influences of the measured point and the heat conduction. Moreover, the following things are found out for the first and the second transition points shown in Fig.4. It is seen from Fig.5 that the hardness of speroidal graphite cast iron gradually falls with the temperature rise under 500 °C, and it remarkably falls beyond 500°C and much more beyond 600°C. In the meantime, the measured temperature of the cast iron pin corresponding to the first and the second transition points were about 500°C and 600°C respectively as seen from Fig.4. Therefore, both are in good correspondences. Then, the hardness of the disk surface becomes higher than that of the cast iron pin when the temperature of the cast iron pin is beyond about 600 °C because the temperature of the disk surface rose littile and its hardness did not change during the test. By the experiments in which the wear length of cast iron pin was set in 4mm, it was confirmed that the aforementioned third wear region was the region where the friction and wear characteristics became stable.

5. CONCLUSIONS

The dry friction and wear characteristics of spheroidal graphite cast iron under severe conditions of high sliding speed and high contact pressure were experimentally investigated. From the experimental results obtained here, the following conclusions are derived.

1) The reductions of hardness of cast iron pin according to the temperature rise caused by friction are strongly influenced with the friction and wear characteristics.

2) The wear of spheroidal graphite cast iron pin changes with sliding distance in three steps. For a while after the start of sliding, the temperature of cast iron pin near the sliding surface rapidly rises but its wear rate is not so large because the cast iron pin has not sufficiently softened (the first wear region). After some short sliding distance, the wear rate remarkably increases because the temperature rise becomes high enough to soften the pin (the second wear region). Then, after this violent wear, the wear rate decreases again and becomes stable since the thermal balance has been achieved (the third wear region).

3) The wear rate of the spheroidal graphite cast iron pin depends on the contact pressure and the sliding speed. It is proportional to the contact pressure in all wear regions. On the other hand, it increases with the increment of sliding speed in the first and the second wear regions but it goes by contraries in the third wear region.

4) The coefficient of friction is large at the start of sliding and decreases more and more with the increment of sliding distance in the first and the second wear regions. In the third wear region, it becomes stable and converges to some constant value.

5) The coefficient of friction decreases with the increment of sliding speed but it is almost independent of contact pressure.

6. ACKNOWLEDGEMENT

The authors wish to special thanks to the members of Elevator Mechanical Development Section of Inazawa Works, Mitsubishi Electric Corporation for their helps and supports.

REFERENCES

(1) E. Takeuchi; "The mechanism of wear of spheroidal graphite cast iron in dry sliding", Wear, 19, 3 (1972) 267

(2) V. Ivusic; "Transition wear behavior of an abrasiton-resistant cast iron", Wear, 74, 2 (1981) 307

(3) E. Takeuchi, T. Nakagawa, F. S Dai, M.Matsunaga and H. Y. Ra; "Friction and wear of sintered cast iron products", Wear 75, 2 (1982) 303

(4) P .W. Leach, "Prow formation in the wear of cast iron", Wear 96, 2 (1984) 135

(5) K. Okada, T. Yoshida and A. Hijikata, "Friction and Wear Characteristics of Gray Cast Iron under High Sliding Speed and High Contact Pressure", J. of Japan Society of Lubrication Engineers, 24,1(1979), 27,(in Japanese)

(6) K. Okada, "Thermal Wear of Cast Iron under High Bearing Pressure", J. of Japan Society of Lubrication Engineers, 30,10 (1985), 733, (in Japanese)

(7) K. Takasawa, T. Tsujimura and S. Yuri, "Frictionand Wear Characteristics of Several Cast Irons", J. of Japan Foundrymen's Society, 57,8(1985), 516, (in Japanese)

(8) T. Tsujimura and S. Manabe, "High Speed Friction Performance of Several Cast Irons and Cast Claddings", J. of Japan Foundrymen's Society, 59,7(1987), 421, (in Japanese)

(9)S .W .E . Earls; "Wear characteristics of some metals in relation to surface temperature", Wear, 20, 1 (1972) 51

(10)F .E. Kennedy; "Thermal and thermomechanical effects in dry sliding", Wear,100 (1984) 453

(11) K. Hayashi, K. Hirasata and K. Sugita, "Effects of thin layer on friction and wear of cast iron under severe sliding conditions", Proc. of 22nd Leeds-Lyon Symp. on Tribology,(1995), 665

(12)K.Hayashi,K.Hirasata and K.Sugita, "Thermaleffects on friction and wear of cast iron under severe sliding conditions", Proc. of lst Asia Int. Conf. on Tribology (1998), 455

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Wear particles produced in repeated contact of two asperities

J. Sugimura[a], J. W. Kang[b], S. Tanaka[a] and Y. Yamamoto[a]

[a] Department of Mechanical Engineering, Kyushu University, Fukuoka 812-8581, Japan
[b] Tsukuba Technical Laboratory, NOK Corporation, Ibaraki 300-4247, Japan

Wear particles produced in repeated contact of aluminium with two pyramidal asperities are studied with image analysis. Four types of wear particles which reflect four different wear mechanisms can be identified with morphological parameters. Type 1 wear occurs only in two-asperity contact, and the particles had thin fibrous shape. Type 2 wear particles are small fragments produced at side ridges. Type 3 wear particles are from a wedge formed in front of the contact, and are of relatively thick, bulky shape. Type 4 wear is cutting caused by asperities with smaller tip angles, yielding chips of characteristic shapes. The criterion of identification based on the morphological parameters is applied to wear particles produced in a conventional ball-on-disk test, in which most particles are identified as Types 2 and 3 though Type 1 particles also exist. It is suggested that several additional processes, including stacking of loose particles within small gaps between sliding surfaces and adhesion of the particles on the surfaces, work to make wear more complex than those with two asperities.

1. INTRODUCTION

Analysis of wear particles suspended in oils is widely accepted as a powerful tool for monitoring tribological conditions in machinery [1]. Introduction of computers in processing images and in analysing particle shapes seems to have gained availability of wear particle analysis in industrial applications [2-8]. However, complex relationships between particle shapes and the processes involved in their formation have yet to be solved, with which identification of particle shape could be more reliable and systematic. A great problem in seeking any of these relationships is that the geometrical and mechanical conditions particles face during their birth differ from particle to particle. It is necessary, and is the aim of this work, to first study particles formed under simple conditions, which may then be extended to analyse particles from actual, more complex, sliding situations.

The simplest form of real contact between solids in sliding may consist of an asperity and a flat surface. Most experimental and theoretical works on fundamentals mechanisms of wear have used this configuration [9,10]. However, this is not the case with actual sliding systems in which a number of asperities make repeated contact. Recently, the authors have found that the repeated sliding of two asperities along close tracks gives rise to a characteristic type of wear that never occurs with only one asperity [11,12]. This suggests that it is necessary to take account of multiple asperities.

This paper describes morphological features of wear particles produced through different wear mechanisms occurring in misaligned dual-pin/disk experiments. Combination of these parameters obtained with image analysis enables identification of wear particles produced through different wear mechanisms. This is then applied to wear particles produced in a conventional ball-on-disk experiment.

2. EXPERIMENTAL

2.1 Dual-pin/disk experiments

As the simplest test to simulate multiple asperity contacts, misaligned dual-pin/disk experiments were conducted [12]. Two diamond pyramidal indenters were slid repeatedly along close tracks on pure aluminium, as shown schematically in Fig. 1. Figure 2 shows the main part of the dual-pin/disk rig. A disk specimen is mounted on a table driven at slow speeds by an AC servomotor via harmonic gears. Two pins were supported by loading arms, and pressed against the disk. The arms were pivoted on position-

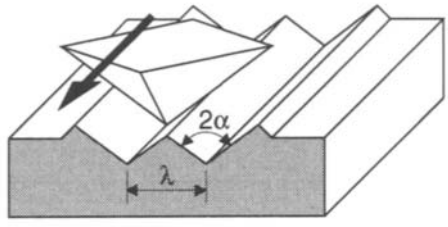


Figure 1. Repeated contact with two asperities

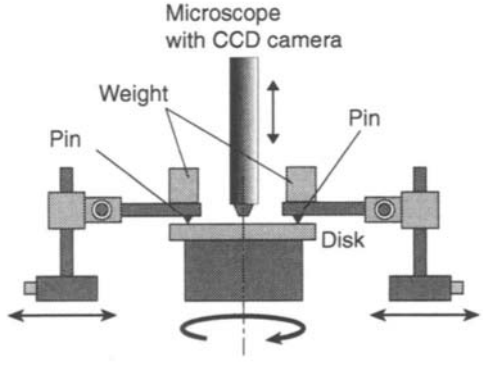


Figure 2. Misaligned dual-pin/disk test rig

controllable tables so that the pins could be located to make contact with the disk on arbitrary tracks. The surface of the disk, when it was at rest, was observed with a CCD microscope, whose images were captured in a computer.

At the tips of the two pins were mounted diamond asperities of pyramidal shapes. Three pairs of pyramids were used with the tip angles 2α of 128°, 148° and 168°. The disk material used was JIS A1050 aluminium containing more than 99.5 % of aluminium, and having a hardness of 46HV. The disks were finished by lapping to 0.04 μmRa.

Experiments were made both with one asperity and two asperities. Each asperity was slid against a disk with one of its edges facing forward, as shown in Fig. 1. With one or both of the asperities loaded against a disk, the disk was rotated for 50 revolutions. Sliding speeds of 2.8 mm/s and 7.0 mm/s, and loads of 0.48 N and 0.57 N were employed. The range of the distance λ between two tracks in tests with the two pins was chosen such that ridges between them joined together to form a central ridge within first five revolutions.

All the experiments were conducted under lubrication with a mineral oil H150, whose kinematic

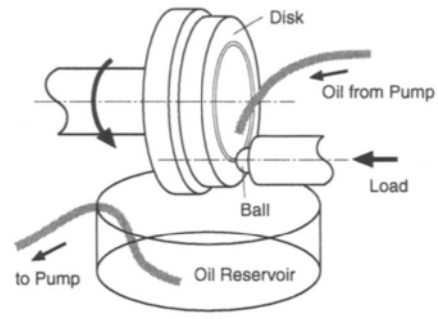


Figure 3. Ball-on-disk experiment

viscosity was 30.0 mm^2/s at 40°C. The oil was put on the disk before each test, and no further oil was supplied during the test. Wear particles produced were collected membrane filters of a pore size of 0.45 μm.

2.2 Ball-on-disk experiment

In order to obtain wear particles formed through ordinary multiple asperity contact, a conventional ball-on-disk test was conducted. Figure 3 shows the test configuration. A disk specimen was made of aluminium A1050. A ball specimen was 19.05 mm in diameter and was made of bearing steel SUJ2, whose surface was shot-blasted to 1.9 μmRrms. The average tip angle of surface asperities is 159°. Sliding was made for six revolution of the disk, *i.e.* for a distance of 0.65 m, at a speed of 7.3 mm/s and a load of 3.5 N, with H150 as a lubricant.

2.3 Analysis of wear particles

The following parameters were determined by analysing microscopic views of wear particles on membrane filters, where shadow images were used for morphological parameters while reflection images used for optical parameters [4-6]. Figure 4 illustrates the basic information for calculating the parameters.

Representative diameter D is particle's size, and defined as square root of projected area A of a particle.

Elongation ϕ is the aspect ratio L/W, where particle width W is the minimum Feret diameter and the length L is measured perpendicular to W.

Modified roundness ψ' represents complexity of contour shape, and is defined as a square of the ratio of the perimeter P to a perimeter P_{el} of an ellipse having the same area and the same aspect ratio. Formally,

$$\psi' = P^2 / P^2_{el} \qquad (1)$$

where

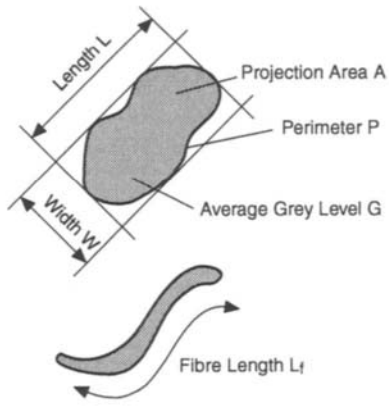


Figure 4. Dimensions of wear particle

$$P_{el} = 4\sqrt{\frac{A\phi}{\pi}}E\left(\sqrt{1-\frac{1}{\phi^2}}\right) \quad (2)$$

and a function E(k) is the elliptic integral of the second kind given by

$$E(k) = \int_0^{\pi/2} \sqrt{1-k^2\sin^2\theta}\,d\theta \quad (3)$$

The modified roundness ψ' is an extension of the well-known roundness parameter, which uses a perimeter of a circle so that it contains information regarding shape elongation as well as further complexity of the shape. The modified roundness ψ' has been introduced in order to describe these information with the two parameters f and ψ' separately [5].

Fibre ratio F is defined by

$$F = L_f^2 / A \quad (4)$$

where L_f is the fibre length measured along the fibre. This is for describing thin and fibrous particle. A fibre length of a particle is determined through thinning operation of a binarised shadow image.

Reflectivity R of a particle is defined here by

$$R = (G/G_0)^{1/\gamma} R_0 \quad (5)$$

where G is the grey level averaged over a particle's reflection image, G_0 and R_0 is the grey level and reflectivity of a reference surface, and γ is an exponent that describes the non-linearity in output of the CCD camera. In this work, a white ceramics plate is used as a reference surface with R_0 of unity under the light source adjusted to give G_0 of 140.

Contrast C used in this work is defined as a root mean square of the average of square of grey level difference between neighbouring pixels, where the average is taken over all vertical and horizontal neighbours, *i.e.* based on 4-connectivity.

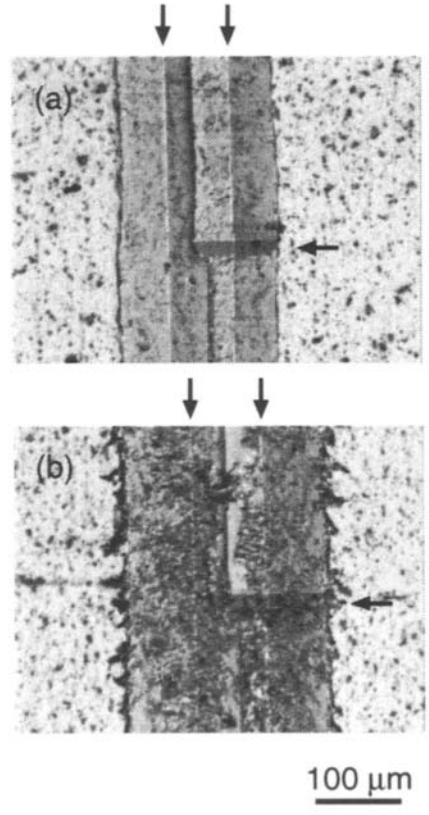


Figure 5. Wear tracks on aluminium disk; 2α=168°, λ=90µm, 2.8 mm/s, 0.48 N; (a) after 10 revolutions, (b) after 50 revolutions; vertical arrows indicate bottom of grooves and the sliding direction

Determination of these feature parameters was made with an image analysis system updated for Microsoft Windows. Microscopic images of wear particles on membrane filters were taken with transmitted light and reflected light separately, and stored as 24-bit colour images of 640x480 pixels. Shadow images by the transmitted light were binarised to black and white, and processed by a thinning-thickening and thickening-thinning operations such that any disconnected points and holes were eliminated.

Particles having representative diameters D larger than 3 µm were selected randomly on the membrane filters. From five to ten views for each of the filter samples were analysed to yield data of about 800 and 100 particles for the dual-pin/disk tests and the ball-on-disk test, respectively.

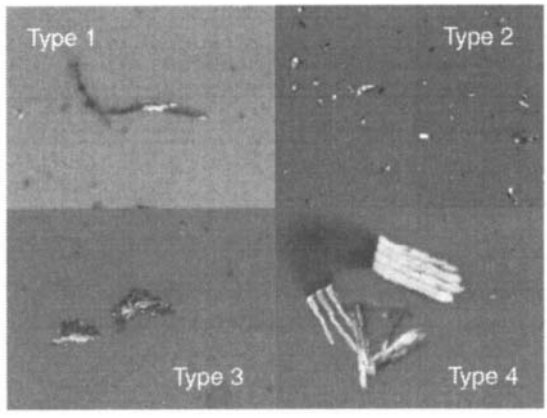


Figure 6. Four types of wear particles

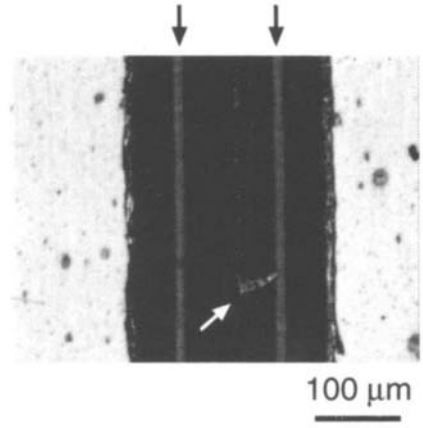


Figure 7. Wear tracks on aluminium disk; 2α=128°, λ=115μm, 2.8 mm/s, 0.48 N; after 6 revolutions

3. RESULTS

3.1 Dual-pin/disk experiments

Typical wear tracks on the aluminium disk are shown in Fig. 5, and four different types of aluminium wear particles generated are shown in Fig. 6.

Wear tracks consisting of grooves and side ridges are formed during initial cycle of sliding, and the grooves grow wider and deeper as the asperities repeat sliding. Sliding of the asperities causes plastic deformation in aluminium without substantial wear during earlier cycles, as shown in Fig. 5(a). Figures 5(a) and 5(b) are microscopic views of the disk surface where the right asperity has just stopped sliding after 10 and 50 cycles; tip angle of the asperities is 2α = 168°. After some cycles, wear particle generation begins to occur at side ridges as shown in Fig. 5(b), and possibly from plastic wedges formed in front of the contacts, as shown by arrows. Particles from the ridges may be small fragments, which are shown as Type 2 particles in Fig. 6. Particles from the wedges may be larger and thicker than Type 2, and may correspond to Type 3 particles shown in Fig. 6.

Although wear of Types 2 and 3 may occur also at the central ridge between the two grooves, many elongated strip-like particles form, which are named Type 1 particles as shown in Fig. 6. Figure 7 demonstrates that this type of particle is about to detach from the central ridge. This type is distinctive of repeated contact of two asperities, and does not occur with only one asperity. Detailed observation has suggested that particles of this type are generated through generation, deformation and fracture of extrusions on the ridge between the two grooves [11]. This is different from the ratchetting wear [13] in that deformation of extrusion is caused by reciprocal loadings. It has been found that Type 1 predominates wear of soft metals such as aluminium and copper when subject to repeated contact with two asperities having larger tip angles [12].

Type 4 wear particles shown in Fig. 6 are cutting chips produced only with the smallest asperity tip angle 2α = 128°. This angle may not normally cause cutting with only one traverse. However, changes in contact geometry after the first traverse, and also the existence of the neighbouring groove [14], must have enabled cutting. Type 4 particles are generally large, with their length or width comparable with the width of the grooves, and have characteristic shape with striations at constant intervals on their face like bellows. They often are fragmented into straight bars, which are called Type 4f.

Figure 8 summarises where and how the four types of wear particles are produced. Types 2, 3 and 4 correspond to three modes of abrasive wear processes, *i.e.* ploughing, wedge formation, and cutting [9]. Type 1 is distinctive of repeated contact of two asperities. Because practical sliding systems inherently suffer from multi-asperity contact, mechanisms involved in this wear mode may be very important.

In Figs. 9(a) to 9(e) are plotted elongation ϕ, modified roundness ψ', reflectivity R, contrast C and fibre ratio F against representative diameter D, in which data for particles from a number of different sliding

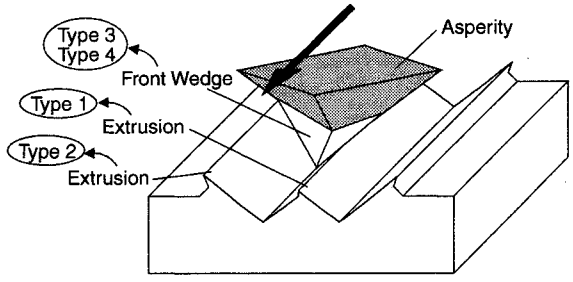
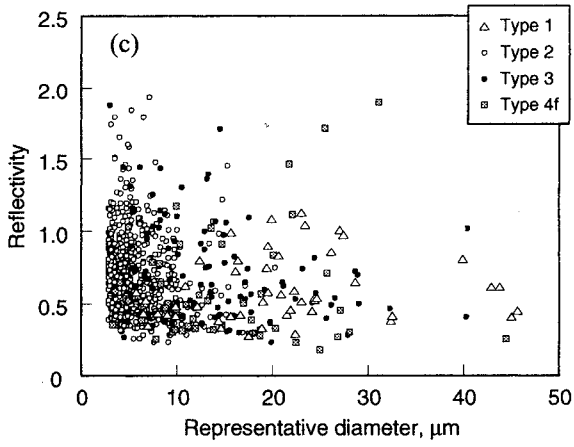
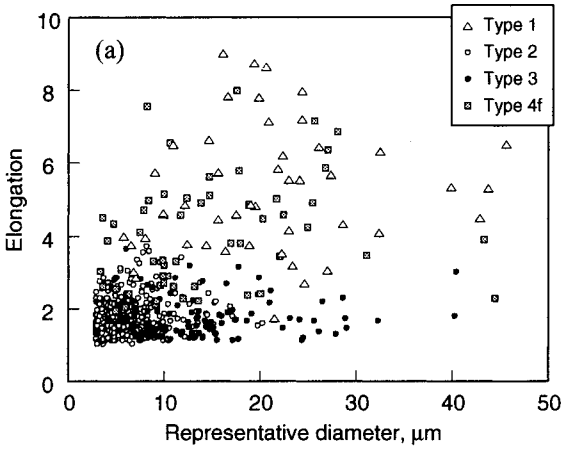
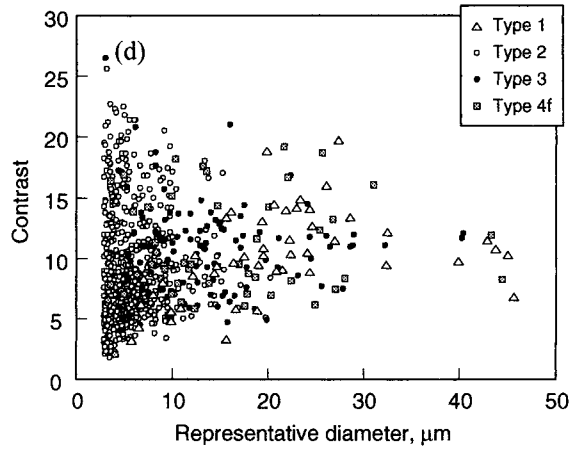
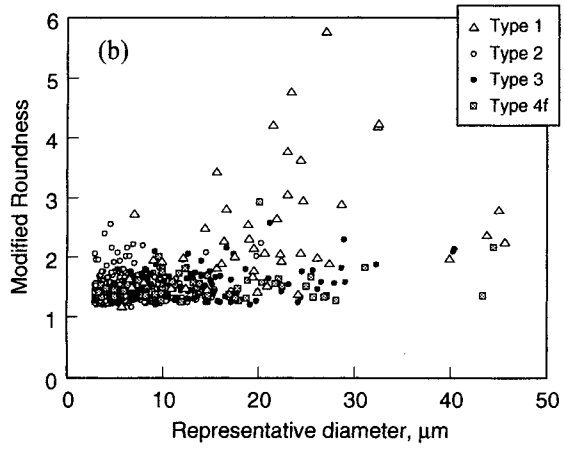


Figure 8. Four types of wear by repeated contact of two asperities









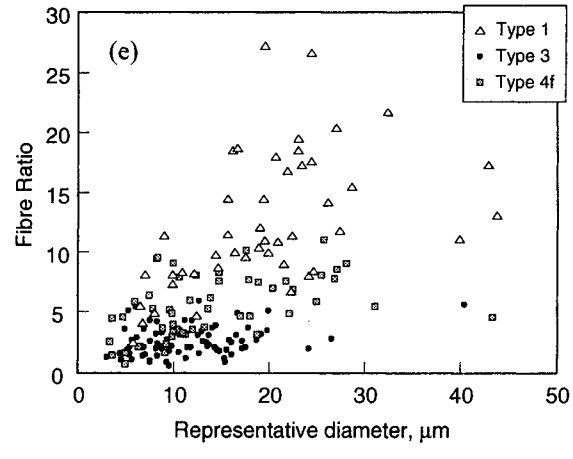


Figure 9. Feature parameters for aluminiium wear particles produced in dual-pin/disk experiments; (a) elongation, (b) modified roundness, (c) reflectivity, (d) contrast and (e) fibre ratio plotted against representative diameter

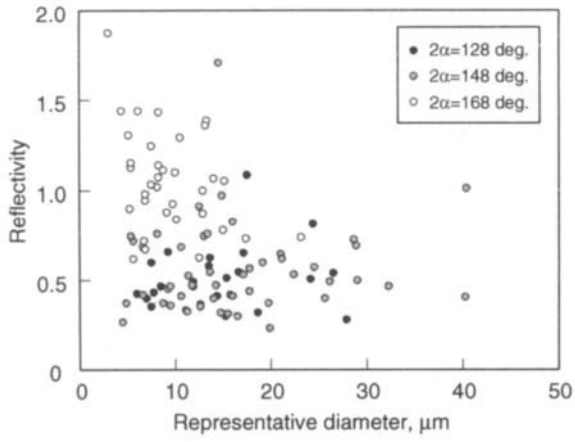


Figure 10. Reflectivity vs. representative diameter for Type 3 particles; 2.8 mm/s, 0.48 N; after 50 revolutions

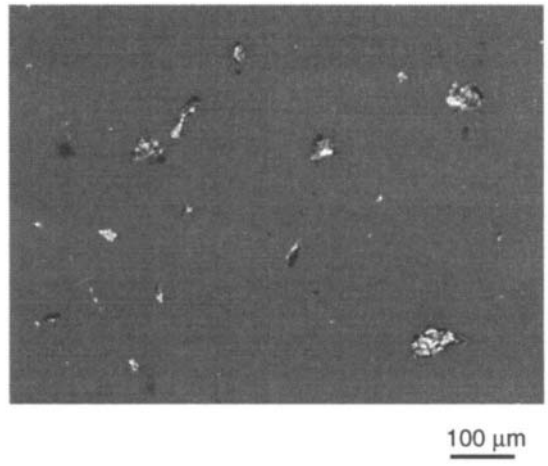


Figure 11. Wear particles generated in ball-on-disk experiment

conditions are plotted altogether. The marks in the figures represent the types of wear particles determined by the authors' eyes before the image analysis was made. For Type 4 particles, only fragmented ones are included in the plots, because of a problem in analysing the connected and partly-connected shapes of large chips. Although the scatter is large, the ranges of parameter values differ from type to type as follows.

Firstly, elongated shapes of Type 1 and Type 4f particles present elongation values larger than 2, which is larger than most of the other two types of particles. Modified roundness, on the other hand, is larger than 2 for slightly curved or wavy shape of Type 1, while it is less than 2 for Type 4f particles whose contours are relatively straight. Difference between these two types is definitely seen in fibre ratio. Thus Type 1 particles and Type 2 particles can be identified with D and two of the three parameters ϕ, ψ' and F, although with some exceptions.

Secondly, rounded shapes of Types 2 and 3 can be distinguished from the above two types using the shape parameters, although there appears to be no parameter to distinguish between these two. In general, Type 3 particles have larger representative diameters D than Type 2. However, both types share the range between 5 and 20 μm. More detailed analysis of shapes or three-dimensional description may be necessary to identify the two types.

Thirdly, the parameter ranges are the same for the same type of particles irrespective of the tip angles of the asperities and other sliding conditions. This is partly because the number of particle data for each of the test conditions is too small to make statistical examination. One of a few exceptions in the present data is that of reflectivity. Figure 10 shows reflectivity of Type 3 particles for three tip angles of the asperities. Particles formed by asperities of $2\alpha = 168°$ are more shiny suggesting that they have less deformed surface than those generated by smaller tip angles.

3.2 Ball-on-disk experiment

On the basis of the above results, an attempt is made to identify wear particles produced in more realistic situation of multiple contacts in the ball-on-disk wear experiment.

Figure 11 shows wear particles collected on a filter. Some of them are very large compared to the average size of the particles analysed in the dual-pin/disk experiments. Most of them look Type 2 or 3, though some elongated particles can also be seen.

Figures 12(a) to 12(c) show elongation ϕ, modified roundness ψ' and reflectivity R plotted against representative diameter D for about 100 particles collected after 90 seconds of run. Comparing these plots with Figs. 9(a) to 9(c), most particles are identified as Types 2 and 3. There is a possibility that ten to twenty particles having elongation larger than 2 are Types 1 or 4f. Photographs of these particles demonstrate that they are Type 1 particles though not highly elongated.

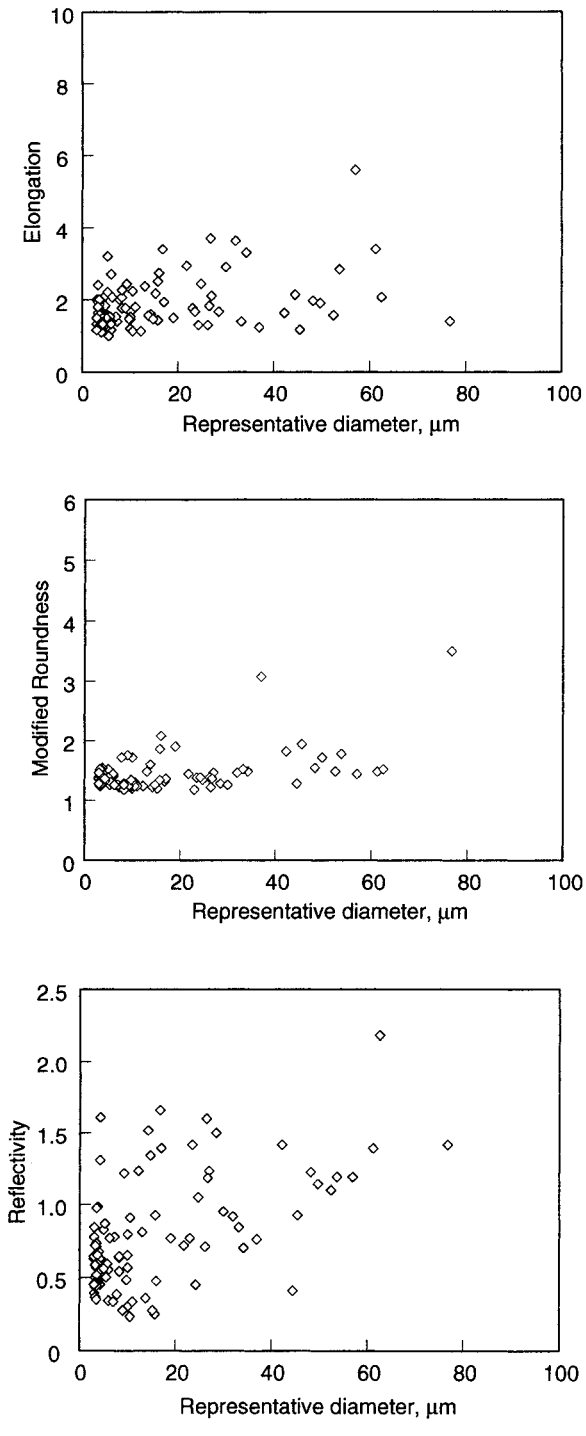


Figure 12. Feature parameters for aluminium wear particles produced in ball-on-disk experiments; (a) elongation, (b) modified roundness, (c) reflectivity plotted against representative diameter

4. DISCUSSION

As the roughened ball would provide multiple asperity contacts, it was expected that parallel grooves would be formed on the aluminium surface and Type 1 wear would occur at ridges between the grooves. However, the particle analysis suggests that Type 1 is not the dominant wear mechanism.

The disk surface after the run had parallel grooves, while aluminium adhered on the ball was found, which also formed parallel grooves. This may have been brought about by simultaneous contact of parallel grooves with parallel grooves, when lateral deformation for Type 1 generation could hardly occur. Also, the average distance between the original asperities on the ball are estimated from surface profiles to be less than 20 μm, which implies that valleys in between the asperities are too shallow to provide space to accommodate particles detached from the disk. Loose particles might be stacked in small gap between the surfaces, which is different from the situation in the dual-pin/disk experiments where there is ample space for particles to escape after removal from the sliding surface. Another factor that might have caused the difference in wear processes is the specific pyramidal shape in the dual-pin/disk experiments.

Because of the changes in the geometry and the small gap between the surfaces, the only plastic deformation allowed for the aluminium surface is that in front of the ball. This is supported by the fact that many Type 3 particles larger than 50 μm are observed, which are larger than the initial asperity distance.

It is still questionable whether the criteria to classify wear particles could be applied to different wear situations. Relations between processes of particle formation and particle shapes in the ball-on-disk experiment, and in most of normal wear situations, are not so simple and clear as those in the dual-pin/disk configuration. It is necessary to clarify any deviations between them, and to find how to incorporate the deviations in the classification of wear particles. Any progress in this respect will be useful in condition monitoring as well as in exploring mechanisms of wear.

For this purpose, the particle analysis should be improved. The parameters used in the present analysis provide only rough pictures of particles, but obviously lack capability to describe detailed features of particles; Types 2 and 3 particles could have features more than just rounded shape. Three-dimensional shape

description and surface texture description [6,8] would be useful, if they are improved to afford information regarding features specific to wear processes.

5. CONCLUSIONS

Four different wear mechanisms were observed in repeated contact of aluminium with two pyramidal asperities. Type 1 wear occurred only in two-asperity contact, and the particles had thin fibrous shape. Type 2 wear particles were small fragments produced at ploughed ridges. Type 3 wear particles were from a wedge formed in front of the contact, and were of a relatively thick, bulky shape. Type 4 wear was cutting wear caused by asperities with smaller tip angles, yielding chips of characteristic shapes. The four types of wear particles could be identified, to some extent, with morphological parameters obtained by image analysis.

The criterion to identify the four types of wear particles was applied to wear particles produced in a conventional ball-on-disk test. Most particles were identified as Types 2 and 3, although there seemed to be some Type 1 particles. It was suggested that several additional processes, including stacking of loose particles within the small gaps between sliding surfaces and adhesion of the particles on the surfaces, contributing to wear more complex than those with two asperities.

REFERENCES

1. T. M. Hunt, Handbook of Wear Debris Analysis and Particle Detection in Liquids, Elsevier, 1993.
2. B. J. Roylance, I. A. Albidewi, A. R. Luxmoore and A. L. Price, The Development of a Computer-Aided Systematic Particle Analysis Procedure - CAPSA, Lubr. Engrg., 48, 12 (1992) 940-946.
3. T. B. Kirk, D. Panzera, R. V. Anamalay and Z. L. Xu, Conputer Image Analysis of Wear Debris for Machine Condition Monitoring and Fault Diagnosis, Wear, 181-183 (1995) 717-722.
4. J. Sugimura, A. Umeda and Y. Yamamoto, Wear Debris Identification with Neural Networks, J. JSME, Pt.C, 61, 590 (1995) 4055-4060 (in Japanese).
5. J. Sugimura, M. Takesue and Y. Yamamoto, Wear debris analysis for condition monitoring of gears, Tribology for Energy Conservation, D. Dowson et al. eds., Elsevier (1998) 431-442.
6. A. Umeda, J. Sugimura and Y. Yamamoto, Characterization of wear particles and their relations with sliding conditions, Wear, 216 (1998) 220-228.
7. G. W. Stachowiak, Numerical characterization of wear particle morphology and angularity of particles and surfaces, in "New Directions in Tribology", I. M. Hutchings ed., Mechanical Engineering Publications, 1997, 371-389.
8. G. W. Stachowiak and P. Podsiadlo, Characterization and classification of wear particles and surfaces, Wear, 249 (2001) 194-200.
9. K. Kato, Wear Mechanisms, in "New Directions in Tribology", I. M. Hutchings ed., Mechanical Engineering Publications, 1997, 39-56.
10. J. A. Williams, Wear Modelling: Analytical, Computational and Mapping: A Continuum Mechanics Approach, Wear, 255-259 (1999) 1-17.
11. J. Sugimura, K. Sakai and Y. Yamamoto, Wear of Ductile Materials by Repeated Contact of Multiple Asperity Tips, Proc. 12th Intern. Colloquium Tribology (2000) 1657-1662.
12. J. W. Kang, J. Sugimura and Y. Yamamoto, Effects of Sliding Conditions on Wear by Repeated Contact of Two Asperities, Proc. International Tribology Conference, Nagasaki, 2000 (2001) 351-356.
13. A. Kapoor and K. L. Johnson, Plastic ratchetting as a mechanism of metallic wear, Proc. Roy. Soc. Lond., A445 (1994) 367-381.
14. J. A. Williams and Y. Xie, The generation of wear surface by the interaction of parallel grooves, Wear, 155 (1992) 363-379.

SESSION XVIII
ELASTOHYDRODYNAMIC LUBRICATION (3) - EXPERIMENTAL AND THEORETICAL

Chairman : Dr C J Hooke

Paper XVIII (i)	**In Situ Pressure Measurements in an Elastohydrodynamically Lubricated Point Contact Using Raman Microspectrometry. Comparison with Numerical Calculations**
Paper XVIII (ii)	**Numerical Modelling of Thermal Effects in Elastohydrodynamic Lubrication Solvers**
Paper XVIII (iii)	**Analysis of the Oil Distribution in Starved Elliptical E.H.D. Contacts**
Paper XVIII (iv)	**Ordinary Shear Thinning Behaviour and its Effect Upon EHL Film Thickness**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

In situ pressure measurements in an elastohydrodynamically lubricated point contact using Raman microspectrometry. Comparison with numerical calculations.

I. Jubault[a], J.L. Mansot[a], P. Vergne[b], A.A. Lubrecht[b] and J. Molimard[c]

[a]Groupe de Technologie des Surfaces et Interfaces
Université des Antilles et de la Guyane 97159 Pointe à Pitre, Guadeloupe

[b]Laboratoire de Mécanique des Contacts, UMR CNRS/INSA 5514, INSA de Lyon
20 avenue Albert Einstein 69621 Villeurbanne cedex, France

[c]Centre SMS, Ecole des Mines de Saint Etienne
158 Cours Fauriel 42023 Saint-Etienne cedex 02, France.

Raman microspectrometry has been applied to in situ pressure measurements in an elastohydrodynamically lubricated (EHL) point contact. An experimental set-up coupling a Raman spectrometer and an elastohydrodynamic ball on disc test rig has been developed. Pressure profiles and maps have been recorded under pure rolling conditions for a wide range of applied loads, rolling speeds and lubricant temperatures.

The measured pressure profiles are compared quantitatively with numerical solutions of the isothermal EHL point contact problem. The good agreement obtained allows a validation of both the measurement method and the experimental set-up. Discrepancies between the experimental and the numerical profiles observed at high rolling speeds are attributed to inlet shear heating of the lubricant. The corresponding temperature rise is estimated. The evolution of the film thickness with speed is studied and the transition from isothermal to thermal regime is pointed out.

1. INTRODUCTION

Elastohydrodynamic lubrication occurs in lubricated contacts that are formed between heavily loaded machine components such as gear teeth and rolling element bearings. In this lubrication regime high pressures are generated within the contact and induce both large elastic deformations of the contacting surfaces and a huge viscosity increase in the lubricant (piezo-viscous effect).

Detailed understanding of the performance of elastohydrodynamically lubricated contacts can be gained through theoretical studies. The first theoretical studies concerning elastohydrodynamic lubrication took place in the 1940s [1-2]. More recently, multigrid techniques have successfully been introduced as a fast and efficient numerical method to simulate EHL line and point contact problems [3-4]. Nowadays, three-dimensional representations of pressure and film thickness are common outputs of numerical simulations [5]. Solutions have been proposed for many complex problems including thermal effects [6-7], surface roughness or starved lubrication [8]. In order to establish and to extend the limits of validity of the models used in numerical simulations, the predictions of these simulations have to be compared to experimental results.

A great variety of experimental techniques has been developed to measure film thickness in EHL contacts. Optical interferometry is the most commonly used technique and has significantly contributed to a better understanding of the way an EHL contact works [9]. The recently developed colorimetric interferometry techniques that combine optical interferometry and computer image analysis allow the mapping of film thickness distributions

with high accuracy and spatial resolution [10-11]. Film thickness measurements confirmed the predictions of the EHL theory with respect to the film shape in the contact zone and its dependence upon operating conditions and lubricant properties.

On the contrary, only a few techniques can provide pressure measurements in lubricated contacts. The earliest techniques used microtransducers made of evaporated manganin deposited onto one of the contacting surfaces [12-14]. These intrusive techniques have several drawbacks: for example, measurements are dependent upon the temperature and the strains at the solid surface [15].

Raman microspectrometry is a powerful technique well adapted to local measurements in contacts owing to its spatial resolution of a few microns [16-17]. In a previous study, Raman microspectrometry was applied to in situ measurements of pressure in a rolling elastohydrodynamic point contact [18]. Pressure profiles and 3D-pressure maps were recorded for two model lubricants. The obtained results were in good qualitative agreement with the predictions of EHL theory. In particular, the presence of the Petrusevich pressure spike and the evolution of its position and intensity with load and speed were clearly evidenced. This study has demonstrated the ability of Raman microspectrometry to measure pressure in dynamic lubricated contacts. However, the experimental pressures have not been compared quantitatively with theoretical predictions. Furthermore, the experimental conditions (load, speed and temperature of the lubricant) were limited by the capacities of the EHL test rig.

In the present study, a new experimental set-up has been developed coupling a Raman spectrometer to an EHL test rig with extended performances. The pressure measurement resolution, the mechanical stability of the test rig and the range of operating conditions are improved. Pressure maps and longitudinal pressure profiles measured for various operating conditions are shown. The numerical solutions of the isothermal EHL point contact problem are compared to experimental results. The aim is to validate the developed experimental method and to show its possibilities, not only for pressure measurement but also for the study of the lubricant behaviour. Evidence of thermal effects in the inlet of the contact at high speed is inferred from the comparison between experimental and numerical pressure profiles. The lubricant's temperature rise in the inlet and the film thickness reduction in the contact are then evaluated.

2. PRESSURE MEASUREMENT

2.1. Principle

The principle of pressure measurement using Raman spectroscopy has been previously described [18]. Under the effect of an applied pressure, the equilibrium distances between neighbouring atoms of the lubricant molecules decrease, inducing an increase of their vibration frequencies and therefore a shift of the corresponding Raman bands on the Raman spectrum. In the pressure range of interest for the present work (0 to 2 GPa), it was shown in [19] and experiments confirm [20-21] that the frequency (expressed in wave number $\bar{\nu}$) varies with the applied pressure p according to the relation:

$$\Delta\bar{\nu} = K \times p \qquad (1)$$

The K coefficient of the lubricant Raman vibrational mode (usually the most intense one) is first calibrated in a high pressure hydrostatic cell. The pressure in the contact is determined by measuring the frequency shift of the corresponding Raman band. The time-scale over which a vibrational mode adapts its frequency is several orders of magnitude shorter than both the lubricant transit time in EHD contacts and the time required to calibrate the frequency shift at a given hydrostatic pressure.

The resolution of the pressure measurements depends on both the pressure sensitivity of the lubricant vibrational mode and the precision of the frequency measurement given by the spectrometer.

2.2. Experimental set-up

The experimental set-up combines a Raman microspectrometer with an EHL test rig (Figure 1). The pressure measurements were performed in a contact formed between a sapphire disc and a steel ball. The characteristics of the ball and the disc are given in Table 1. The ball and the disk are driven by independent motors to produce pure rolling conditions. The velocities are controlled with high precision and stability.

The bottom of the ball dips in a test reservoir containing the lubricant, ensuring fully flooded conditions in the contact. The contact and the

lubricant are thermally isolated from the outside and heated/cooled by an external thermal controlling system. A platinum temperature probe monitors the lubricant temperature in the test reservoir.

	Steel ball	Sapphire disc
Elastic modulus	E_1=210 GPa	E_2=360 GPa
Poisson ratio	ν_1=0.3	ν_2=0.34
Reduced elastic modulus	E'=294 GPa	

Table 1. Characteristics of the contacting materials.

Raman spectra were recorded with a U1000 Jobin-Yvon spectrometer coupled to a microscope. The argon-ion laser 514.5 nm line was used at a power of approximately 16 mW on the sample. The recording time varies from 120 s to 360 s for each spectrum depending on the thickness of the lubricating film. Measurements were performed with a x50 long working distance objective giving an analytical probe diameter of about 10 μm within the plane of the contact.

A monochromator with two 600 gr/mm gratings was used to disperse the Raman spectrum on a CCD detector, giving a spatial dispersion of 0.4 cm^{-1} per pixel. The central frequency of a Raman band was determined by fitting the data to a Gaussian curve, providing a precision on its position of $\Delta\bar{\nu} = \pm 0.1\ cm^{-1}$.

In order to perform measurements at any point within the contact area, the Raman microscope can be placed precisely above the EHL test rig by means of a motorised micropositioning XY system. The exact position of the laser beam in the contact is located from the interference Newton's rings observed in white light through a video camera mounted on the microscope. The Z displacement is used to focus the beam in the lubricant film.

2.3. Model lubricant and operating conditions

The used model lubricant is a polyphenyl ether oil commonly named 5P4E or Santovac 5. The choice of this oil was motivated by both its vibrational and rheological properties.

The 5P4E is a very good Raman scatterer, whose Raman spectrum is dominated by an intense band at 1001 cm^{-1} related to the trigonal mode of vibration of the aromatic rings within the 5P4E molecules. The pressure sensitivity of this mode has been calibrated in diamond anvil cells by several authors [22-24]. The value of K=3.3 cm^{-1}/GPa given by Laplant et al. [24] was used in the present study. The pressure resolution of our measurements is about 30 MPa for this lubricant and the Raman spectrometer used.

The 5P4E presents both a high viscosity and an important pressure viscosity coefficient so it can form sufficiently thick films in the contact to obtain spectra with a good signal to noise ratio. Its rheological properties at 25°C and 50°C are given in Table 2.

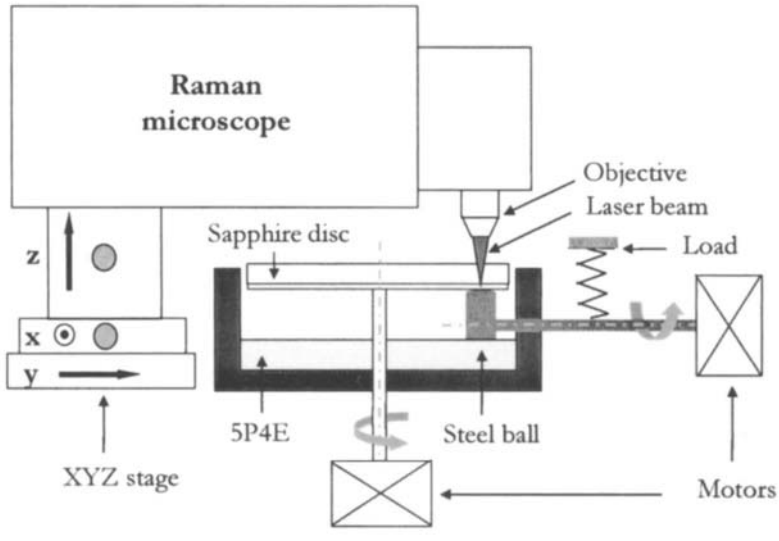


Figure 1. Schematic representation of the experimental set-up.

	T=25°C	T=50°C
Viscosity η_0	2.10 Pa.s	0.16 Pa.s
Pressure-viscosity coefficient α*	45.9 GPa^{-1}	28.4 GPa^{-1}

Table 2. Rheological properties of 5P4E.

Load	17-237 N
Hertzian pressure	0.75-1.84 GPa
Hertzian contact radius	104-248 μm
Speed u_e	0.05-2 m/s
Central film thickness	0.3-2.2 μm
Dimensionless load parameter M	7-438
Dimensionless material parameter L	17-52

Table 3. Operating conditions.

Pressure measurements have been performed for various operating conditions covering a wide range of load, speed and temperature. These conditions are gathered in Table 3 together with the maximum pressure p_0, the contact radius a calculated by the Hertz theory and the central film thickness h_c. The film thickness in the contact at 50°C and for low speeds has been measured using a differential colorimetric interferometry technique. The principle of this technique has been already described elsewhere [11]. When thickness measurements could not be done (at 25°C and for high speeds), the central film thickness has been estimated by the elastohydrodynamic (EHD) models of Hamrock and Dowson [25] or Venner [5], according to the assumptions made in these models (isothermal flow, Newtonian behaviour...).

Longitudinal pressure profiles were made along the rolling direction (x direction) from the inlet to the outlet, performing measurements every 10-30 μm in the regions of low pressure gradients and every 5-10 μm in the regions of high pressure gradients. The spacing between two consecutive measurements was also fitted to the contact diameter, which changes with the load. For some conditions, complete mappings of the contact were achieved. Several longitudinal profiles were made, separated from each other by 10-30 μm in the central region of the contact and by 5-10 μm on the sides.

3. NUMERICAL SOLUTION

The solution of the EHL problem requires the simultaneous solution of the Reynolds and the elasticity equations to obtain the pressure and the film thickness distributions, taking into account that the fluid's properties (viscosity and density) are functions of pressure. The governing equations describing the steady-state EHL point contact under isothermal and fully flooded conditions are given bellow, equations (2-6). The lubricant is assumed to be Newtonian and the surfaces are assumed to be perfectly smooth.

The domain of study is rectangular and the co-ordinates are defined such that the x-axis has the same direction as the rolling speed. The origin is located at the centre of the contact. The stationary Reynolds equation is:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial y}\right)=u_e\frac{\partial(\rho h)}{\partial x} \qquad (2)$$

The elasticity equation is given by:

$$h(x,y) = h_0 + \frac{x^2+y^2}{2R}+\frac{2}{\pi E'}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{p(x',y')dx'dy'}{\sqrt{(x-x')^2+(y-y')^2}} \qquad (3)$$

The lubricant is assumed to be compressible according to the Dowson-Higginson equation:

$$\rho(p)=\rho_0\left(1+\frac{0.6.10^{-9}p}{1+1.7.10^{-9}p}\right) \qquad (4)$$

The viscosity is assumed to depend on the pressure according to the Roelands equation:

$$\eta(p)=\eta_0\exp\left[\left(\alpha\frac{p_R}{z}\right)\left(-1+\left(1+\frac{p}{p_R}\right)^z\right)\right] \qquad (5)$$

$$\text{with}\quad \left(\alpha\frac{p_R}{z}\right)=\ln\eta_0+9.67$$

and z=0.6 and p_R=1.98x10^8 Pa.

The applied load must be balanced by the pressure generated in the contact. This condition is written as:

$$w = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy \qquad (6)$$

A fast multilevel solver has been applied to the solution of the EHL problem [4, 8]. The operating conditions are expressed in terms of the Moes dimensionless point contact parameters M and L, and dimensionless variables are used:

$$X=\frac{x}{a}, Y=\frac{y}{a}, P=\frac{p}{p_0}$$

where a is the contact radius and p_0 the maximum pressure given by the Hertz theory.

The solutions have been calculated on a domain $\{(X,Y)\in\Re^2, -4.5 \le X \le 1.5, -3 \le Y \le 3\}$, with pressure taken as zero along the edges of the computation zone and the application of the Reynolds boundary condition at the outlet.

4. RESULTS AND DISCUSSION

Pressure mappings of the contact measured for two different operating conditions are shown in Figure 2 and Figure 3 together with the numerical pressure distributions calculated for the corresponding M and L parameters.

Measurements have been done on the whole contact area so that the symmetry of the pressure distribution with respect to the central longitudinal line Y=0 can be verified. The integration of the experimental pressure distributions gives the applied load within an error less than 10% at low load (Figure 2) and less than 3% at higher load (Figure 3).

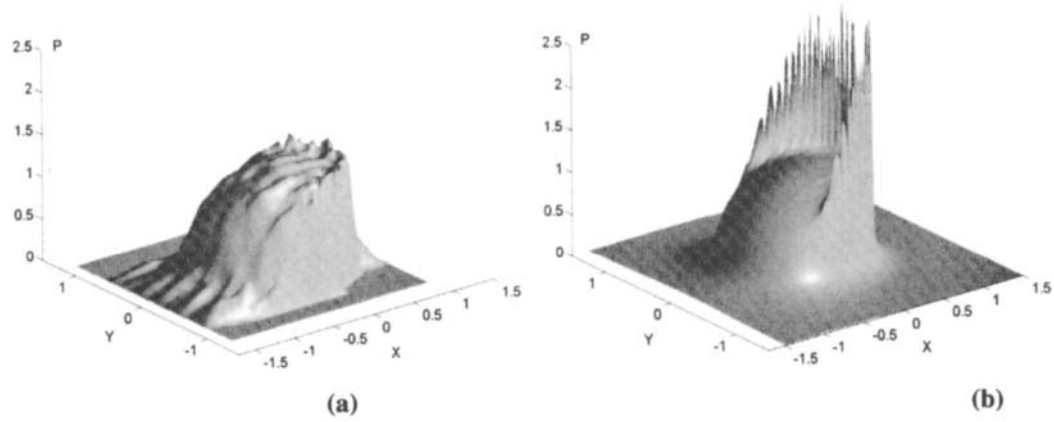


Figure 2. Three-dimensional representations of experimental (a) and numerical (b) dimensionless pressure distribution for w=20 N, u_e=0.3 m/s, T=50°C (M=37, L=19).

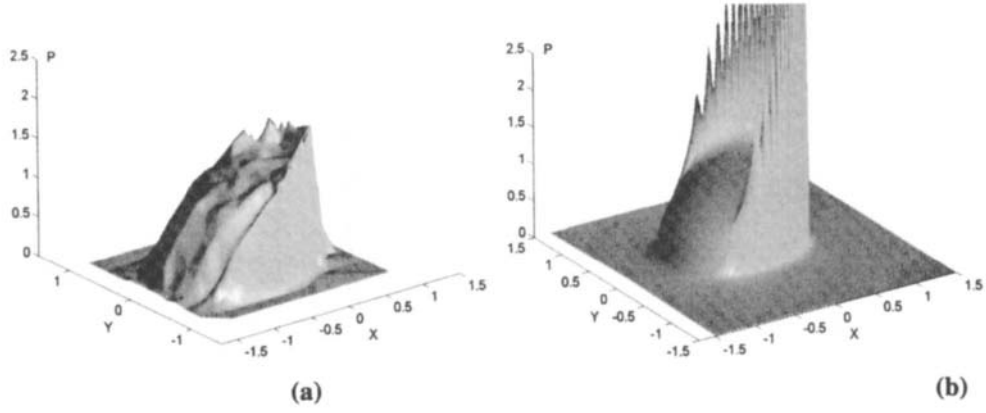


Figure 3. Three-dimensional representations of experimental (a) and numerical (b) dimensionless pressure distribution for w=43 N, u_e=0.094 m/s, T=25°C (M=28, L=43).

The zone of highest pressure is located at the rear of the contact. It has the horseshoe shape usually observed in film thickness maps, which corresponds to a thickness reduction. On the longitudinal pressure profiles, this local overpressure is generally referred as the pressure spike. At 50°C (Figure 2), the pressure distribution shows a central zone of constant high pressure surrounded by a horseshoe shaped zone of low intensity compared to the distribution obtained at 25°C (Figure 3). At 25°C, the pressure distribution deviates significantly from the Hertzian one, the central region of constant pressure is nearly non-existent and large pressure gradients occur. Since 5P4E is more viscous at 25°C than at 50°C (see Table 2), it generates a thicker film in the contact ($h_c \approx 1.4$ μm in the case of Figure 3 and $h_c \approx 0.4$ μm in the case of Figure 2) and so exhibits broader and higher pressure spikes.

The main differences between the experimental and the numerical pressure distributions concern the pressure spikes, which are much more intense on the numerical distributions. The pressure measured by Raman microspectrometry is averaged over the surface of the laser probe, whose diameter is estimated at 10 μm in our experiments. Consequently, the intensity of the experimental pressure spike is lower than its real value. On the other hand the intensity of the calculated spike strongly depends on the numerical solution, notably on the level of discretization. For these reasons, it is difficult to know accurately the real intensity of the pressure spike. Furthermore, the high L values cause the solution to converge slowly, and a small error may persist in some solutions.

On the basis of longitudinal profiles, quantitative comparisons between experimental and numerical results are proposed. In Figure 4 and Figure 5 the experimental pressure profiles measured for various loads and speeds at temperature of 50°C and 25°C are compared to the numerical profiles calculated for the corresponding M and L parameters. The dimensionless pressure P is shown as a function of the dimensionless position X along the centre line of the contact (Y=0) in the longitudinal direction. To make comparison easier, all the profiles are at the same scale.

The experimental and the numerical profiles show the same evolution with load and speed: as the load is increased or as the speed is decreased, the pressure profile approaches the Hertzian shape. The pressure spike becomes narrower, its intensity decreases and its position moves toward the exit of the contact. In all cases, the intensity of the pressure spike is much higher for the numerical profiles than for the experimental ones, for the reasons previously explained.

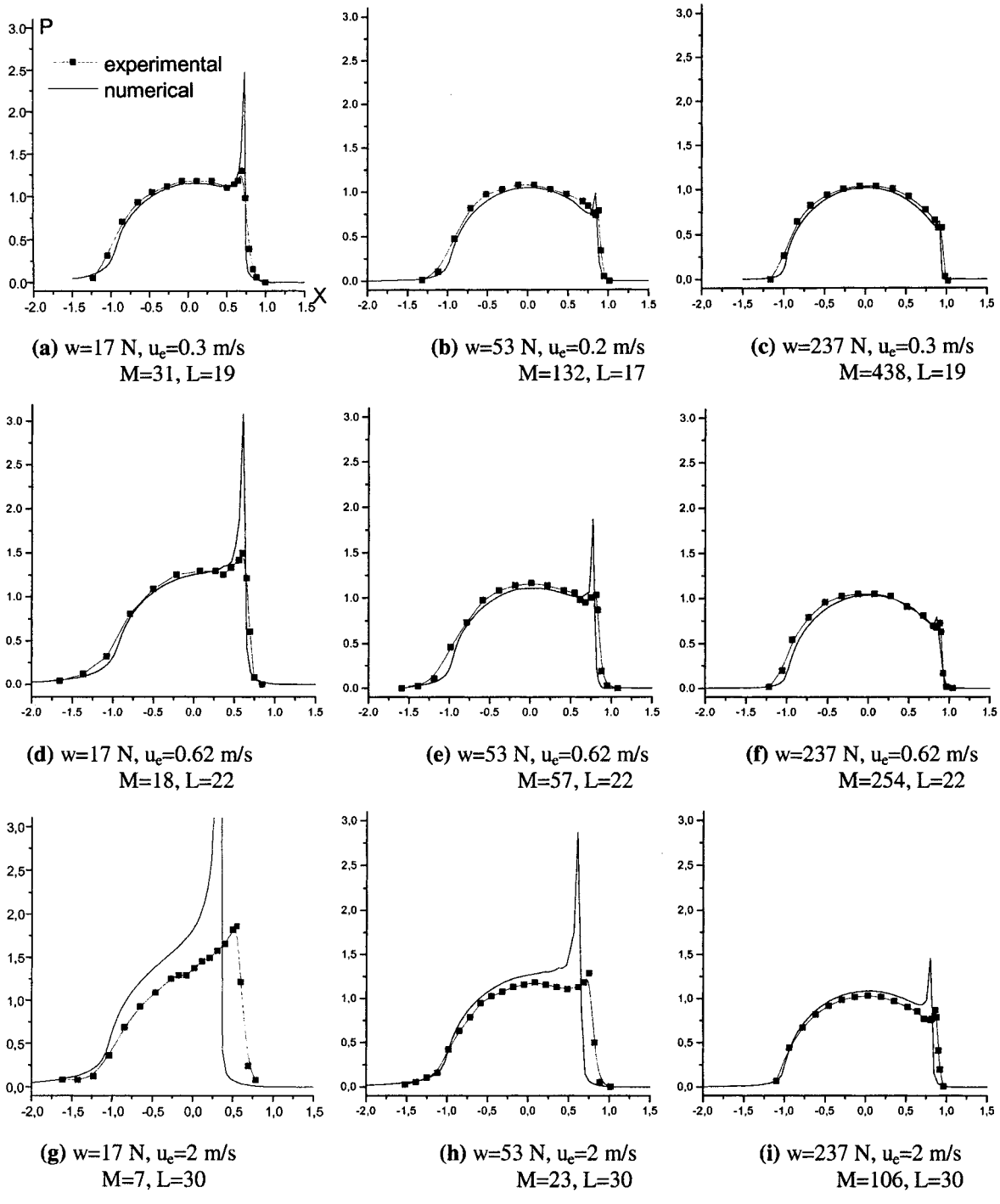
- For low and moderate speeds ($u_e \leq 0.62$ m/s at 50°C, Figures 4a to 4f and u_e=0.05 m/s at 25°C, Figure 5a), there is a good agreement between the experimental and the numerical profiles, in particular with regard to the shape of the profiles, the value of the pressure at the centre of the contact and the position of the pressure spike.
  For high speeds (u_e=2 m/s at 50°C, Figures 4g to 4i and u_e=0.2 m/s at 25°C, Figure 5c), significant discrepancies are observed between the experimental and the numerical profiles. The measured pressures are lower than the calculated ones and the pressure spike position is shifted toward the exit of the contact.

The actual rheological behaviour of the lubricant in the contact is not exactly known and the rheological laws used in the calculations are only approximations. Some studies show that the viscosity of 5P4E varies less at low pressure and more at high pressure than predicted by the Roelands equation (5) [26]. Therefore, the agreement obtained between experimental and numerical profiles for moderate speeds under various load and temperature conditions is very good and allows to validate the measurement method and the experimental set-up.

The problem is then to explain the discrepancies observed for high speeds. The shear stresses in the contact inlet increase with speed and can induce thermal dissipation in the film. The viscosity of 5P4E is very sensitive to temperature and can be modelled by:

$$\eta(T)=\eta_0 \exp\left[\beta\left(\frac{1}{T}-\frac{1}{T_0}\right)\right] \qquad (7)$$

where the temperature-viscosity coefficient β is equal to 8500 K at 50°C and to 13000 K at 25°C. Thus, the viscosity of 5P4E decreases of 13% per degree around 25°C. Therefore a temperature rise of a few degrees induces an important viscosity decrease which significantly changes the film thickness and pressure distributions.



(a) w=17 N, u_e=0.3 m/s M=31, L=19

(b) w=53 N, u_e=0.2 m/s M=132, L=17

(c) w=237 N, u_e=0.3 m/s M=438, L=19

(d) w=17 N, u_e=0.62 m/s M=18, L=22

(e) w=53 N, u_e=0.62 m/s M=57, L=22

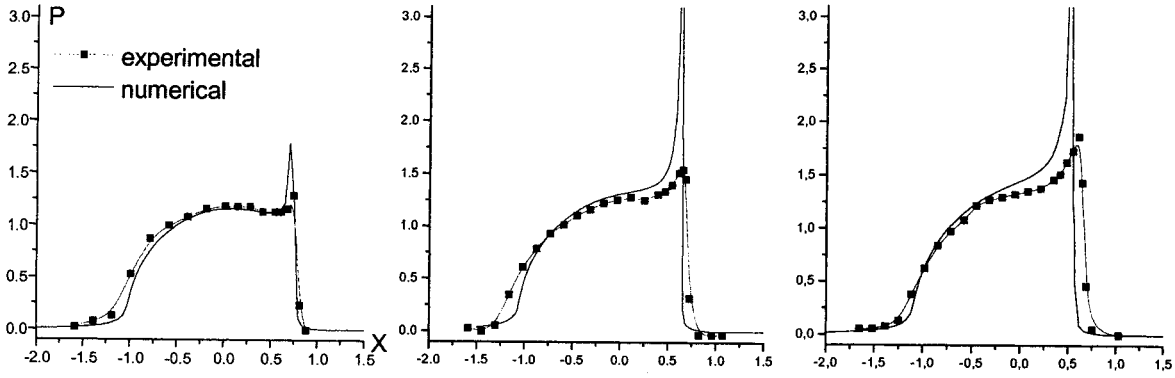
(f) w=237 N, u_e=0.62 m/s M=254, L=22

(g) w=17 N, u_e=2 m/s M=7, L=30

(h) w=53 N, u_e=2 m/s M=23, L=30

(i) w=237 N, u_e=2 m/s M=106, L=30

Figure 4. Comparison between experimental and numerical dimensionless central line pressures for various loads and speeds at T=50°C.



(a) w=53 N, u_e=0.05 m/s M=54, L=37

(b) w=43 N, u_e=0.094 m/s M=28, L=43

(c) w=53 N, u_e=0.2 m/s M=19, L=52

Figure 5. Comparison between experimental and numerical dimensionless central line pressures for various loads and speeds at T=25°C.

In order to verify if the discrepancies between experimental and numerical pressure profiles are really due to inlet shear heating, new isothermal solutions have been computed for the same loads and speeds but for higher inlet temperatures of the lubricant. The numerical profiles calculated for a load of 17 N, a speed of 2 m/s and for four different temperatures of 5P4E (50°C, 55°C, 60°C and 65°C) are compared in Figure 6 to the corresponding experimental profile measured for a temperature in the reservoir of 50°C. The profile calculated for a temperature of 55°C is above the experimental profile, with a pressure spike shifted toward the centre of the contact. On the contrary, the pressures calculated for a temperature of 65°C are lower than the experimental ones and the pressure spike is shifted toward the exit of the contact. The profile calculated for a temperature of 60°C agrees very well with the experimental one. We infer from these results that owing to shear heating, the lubricant's temperature in the inlet is 10°C above its temperature in the reservoir, within an estimated uncertainty of ± 2°C.

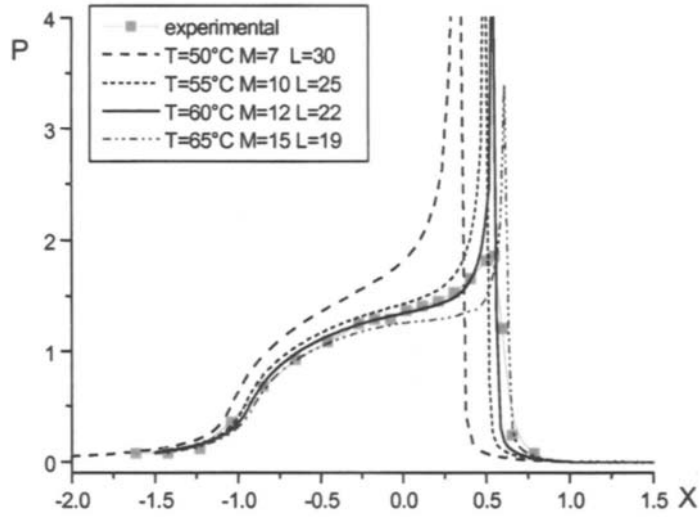
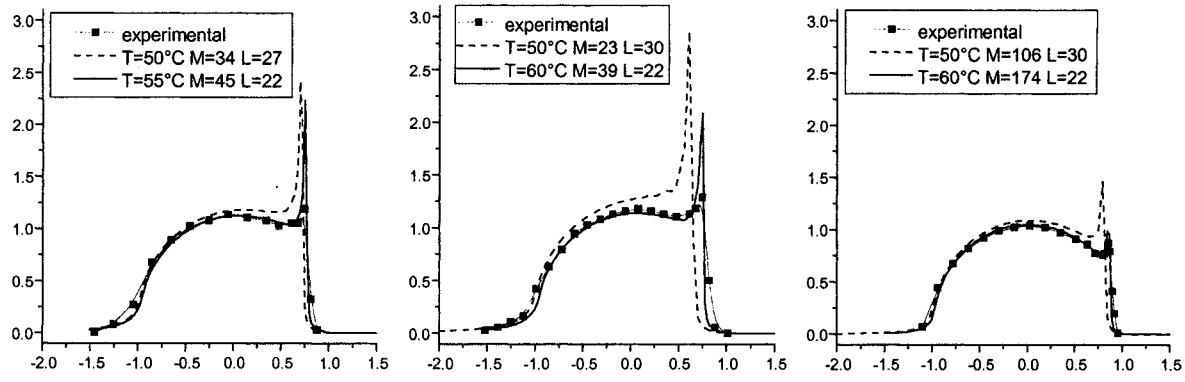


Figure 6. Comparison of the experimental central line pressures measured for w=17N and u_e=2m/s with the numerical profiles calculated for various inlet temperatures (T=50°C in the reservoir).

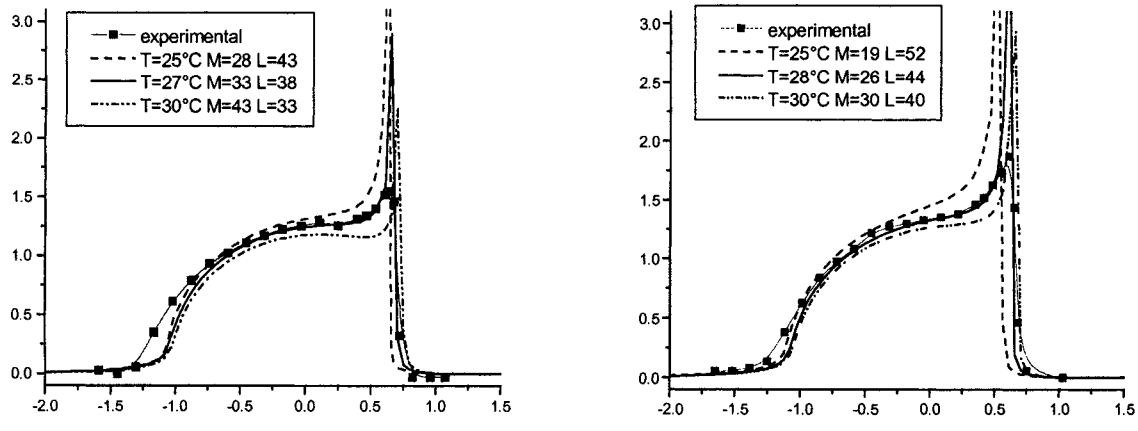
The same process has been repeated for all the conditions for which deviations have been observed between experimental and numerical profiles. In each case, the calculations have been carried out for several temperatures of the lubricant. The estimated inlet temperature is the one giving the best agreement between experimental and numerical profiles. At 50°C in the reservoir, the temperature rise of the lubricant is estimated to be equal to 5 ± 2°C at 1.2 m/s (Figure 7a) and to 10 ± 2°C at 2 m/s whatever the load (Figure 7b and c). At 25°C, a good agreement is obtained for a temperature rise of 2 ± 1°C at 0.094 m/s (Figure 8a) and 3 ± 1°C at 0.2 m/s (Figure 8b). Only direct measurements of the inlet temperature would permit verification of these results.

As the film thickness in the high pressure zone is determined by the properties of the lubricant in the inlet zone, a temperature increase in the inlet, which results in a viscosity decrease, will lead to a reduction in film thickness as shown by Cheng [27]. This reduction can have severe practical consequences, for example on the durability of machine elements. Therefore, it is useful to predict the experimental conditions for which thermal effects occur as well as the real film thickness in the contact.



(a) w=53 N, u_e=1.2 m/s **(b)** w=53 N, u_e=2 m/s **(c)** w=237 N, u_e=2 m/s

Figure 7. Comparison of the experimental central line pressures with the numerical profiles calculated accounting for lubricant inlet shear heating (T=50°C in the reservoir).



(a) w=43 N, u_e=0.094 m/s **(b)** w=53 N, u_e=0.2 m/s

Figure 8. Comparison of the experimental central line pressures with the numerical profiles calculated accounting for lubricant inlet shear heating (T=25°C in the reservoir).

Film thickness measurements using white light optical interferometry being limited to thickness lower than 700 nm, measurements at high speeds (u_e>0.62 m/s) could not be performed.

The numerical solution of the EHL equations provides at once pressure and film thickness distributions. Figure 9 shows the evolution of the central film thickness with the speed for a reservoir temperature of 50°C and a load of 53 N. The film thickness is calculated with and without lubricant inlet shear heating, considering an inlet temperature respectively of 55°C at 1.2 m/s and 60°C at 2 m/s and equal to 50°C. Film thickness measurements performed by optical interferometry are also represented in Figure 9. Two regions can be observed: a low speed region where the film thickness increases linearly with speed (in logarithmic coordinates) and a high-speed region where the film thickness remains almost constant. The first region corresponds to a regime for which thermal effects in the contact inlet can be neglected; the second region corresponds to a regime for which these effects become dominant. The transition between these two regimes occurs for a speed of 0.7 m/s at 50°C (case of Figure 9) and 0.05-0.1 m/s at 25°C. Of course, these values and the associated temperature rises depend on the experimental set-up, on the properties and the geometry of the contacting solids, etc.

In the thermal regime, the isothermal EHD models overestimate the real film thickness. Nevertheless, these models can estimate the film thickness if the real inlet temperature of the lubricant is known.

These results show that the experimental and the numerical approaches are complementary and that it is useful to use them jointly so as to obtain information that neither of them can provide alone. They also point out the necessity to develop both finer models describing the real rheological and thermal behaviour of the lubricant, and in situ investigations performed at a local scale within lubricated contacts.

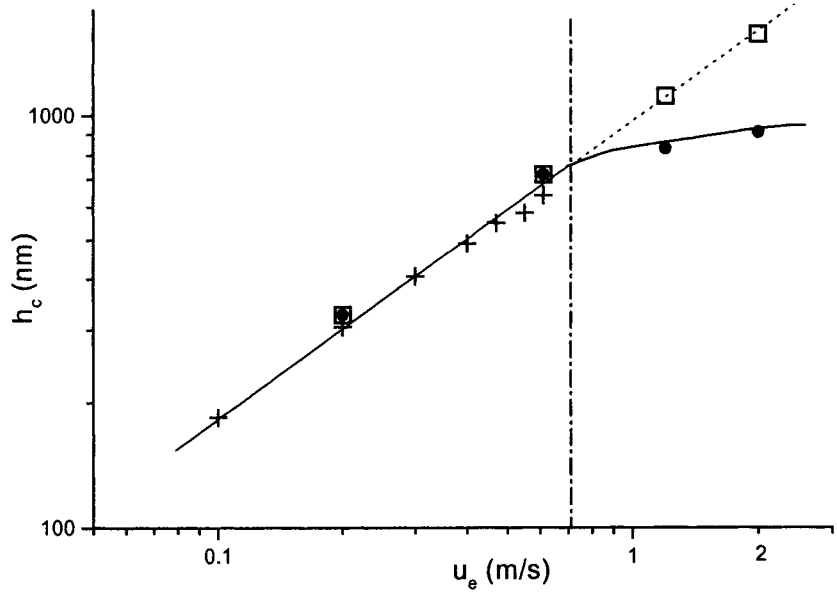


Figure 9. Evolution of the central film thickness h_c with the speed u_e for w=53 N (T=50°C in the reservoir). Values of h_c:

- measured by optical interferometry (+),
- calculated with (●) (T≥50°C in the inlet) and without (□) (T=50°C in the inlet) inlet shear heating.

5. CONCLUSION

An in situ pressure measurement method based on Raman microspectrometry has been developed. Pressure profiles and maps have been measured within a rolling EHL point contact for a wide range of operating conditions. The experimental pressure profiles were compared quantitatively with the numerical solutions of the isothermal EHL problem. These comparisons allowed a validation of the experimental method developed and showed the occurrence of thermal effects in the inlet of the contact at high speeds. Furthermore, it is possible to estimate the inlet temperature rise of the lubricant and the film thickness in the contact.

The obtained results show that the experimental method developed in this work is a very useful tool for the study of dynamic lubricated contacts. This method can be applied to the study of more complex and less known situations than the EHL contact under pure rolling conditions. Thus, pressure measurements have been recently performed in an EHL contact under rolling-sliding conditions [19] and in a dented lubricated contact under pure sliding conditions [28].

ACKNOWLEDGMENTS

The authors wish to thank G. Roche and M. Querry (from LMC) for the development of the experimental set-up, and for their precious support and comments during the experiments.

REFERENCES

[1] A.M. Ertel, Akad. Nauk SSSR Prikadnaya Mathematica i Mekhanika 3 n°2 (1939) 41.
[2] A.N. Grubin, Central Scientific Research Institute for Technology and Mechanical Engineering, Moscow, D.S.I.R. translations (1949) 115.
[3] A. Brandt, Mathematics Comput. 31 (1977) 333.
[4] A.A. Lubrecht, Ph.D. Thesis, University of Twente, Enschede, The Netherlands, 1987.
[5] C.H. Venner, W.E. Ten Napel, Wear 152 (1992) 368.
[6] F. Sadeghi, P.C. Sui, J. of Tribology 112 (1990) 189.
[7] B. Piccigallo, Wear 193 (1996) 56.
[8] C.H. Venner, A.A. Lubrecht, Proc. Instn. Mech. Engrs. 214 Part J (2000) 43.
[9] H.A. Spikes, Proc. Instn. Mech. Engrs. 213 Part J (1999) 335.
[10] M. Hartl, I. Krupka, M. Liska, Trib. Trans. 42 (1999) 36.
[11] J. Molimard, Ph.D. Thesis, INSA de Lyon, France, n°99ISAL0121, 1999.
[12] W.J. Bartz, J. Ehlert, Trans. ASME J. Lub. Tech. 98 (1976) 501.
[13] M.M.A. Safa, J.C. Anderson, J.A. Leather, Sensors and Actuators 3 (1982) 119.
[14] P.M. Johns-Rahnejat, R. Gohar, Tribotest Journal 1 (1994) 33.
[15] H.J. Marginson, R.S. Sayles, A.V. Olver, Trib. Int. 28 8 (1995) 517.
[16] J.L. Mansot, Ph.D. Thesis, University of Lyon I, France, n°86-50, 1986.
[17] D.J. Gardiner, E. Baird, C. Craggs, M.P. Dare-Edwards, J.C. Bell, Lub. Science 1 (1989) 301.
[18] I. Jubault, J.L. Mansot, P. Vergne, D. Mazuyer, J. of Tribology 124 (2002) 114.
[19] I. Jubault, I., Ph D. Thesis, Université des Antilles et de la Guyane, France, n°02AGUY0077, 2002.
[20] R.D. Etters, A.A. Helmy, Physics of Solids Under High Pressure, J.S. Schilling and R.N. Shelton (eds), North-Holland, Amsterdam, 1981.
[21] J.R. Ferraro, Vibrational Spectroscopy at High External Pressures, Academic Press, 1984.
[22] D.J. Gardiner, E. Baird, E., A.C. Gorvin, W.E. Marshall, M.P. Dare-Edwards, Wear 91 (1983) 111.
[23] A. Haiahem, Ph.D. Thesis, INSA de Lyon, France, n°89ISAL0018, 1989.
[24] F. Laplant, E.J. Hutchinson, D. Ben-Amotz, J. of Tribology 119 (1997) 817.
[25] B.J. Hamrock, D. Dowson, Trans. ASME J. Lub. Tech. 99 (1977) 264.
[26] S. Yasutomi, S. Bair, W.O. Winer, J. of Tribology 106 (1984) 291.
[27] H.S. Cheng, ASLE Transactions 8 (1965) 397.
[28] S. Coulon, I. Jubault, A.A. Lubrecht, F. Ville, P. Vergne, accepted for the 6th Int. Trib. Conf. Austrib 2002.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Numerical Modelling of Thermal Effects in Elastohydrodynamic Lubrication Solvers

R Fairlie[a*] C.E.Goodyer,[a] M.Berzins,[a] and L.E.Scales,[b]

[a]Computational PDEs Unit, School of Computing, University of Leeds, Leeds, UK

[b]Shell Global Solutions, Chester, UK

The numerical modelling of thermal effects in elastohydrodynamic lubrication solvers is addressed by extending the code of the authors. This code is used to investigate the important effect of shear upon the temperature profiles for both a steady state and a transient line reversal example and also for a steady state point contact problem. Additional equations for the modelling of mean and surface temperatures within an elastohydrodynamic lubrication contact are described and are added into the existing multigrid driven solver. Temperature profiles in steady line and point contact cases are presented and used to demonstrate that even small amounts of shear produce temperatures in excess of those in the pure rolling case.

1. Introduction

Within an elastohydrodynamically lubricated (EHL) contact the temperature of both the entrained fluid and the contact surfaces can vary dramatically. In the pure rolling case temperature rises of only a matter of degrees may be expected, however when shear occurs the temperature may rise in severe cases by in excess of a hundred degrees [1]. Fluid temperature plays a direct role in governing the density and viscosity of a fluid, which in turn affects the pressure profile and surface deformation across the contact. Clearly, then, with temperature fluctuations of this magnitude occurring it is important to include thermal effects in EHL models.

These thermal effects have been extensively studied for many years since the initial theoretical work by Crook [2] in 1961, with numerical methods for the thermal solutions in line contacts developed in the 1960's by Sternlicht [3], Cheng and Sternlicht [4], Cheng [5] and by Dowson and Whitaker [6]. Recent papers, for example, those by Lee *et al [7], [8], by Kim* et al [9,10] and by Kazama et al [11] have utilised the multigrid techniques applied to EHL problems by Venner and Lubrecht [12] to solve both line and point contact cases.

In this paper the transient EHL model developed in the CPDE Unit at Leeds, [13,14] using the multigrid techniques applied to EHL problems by Venner and Lubrecht [12] is extended to include thermal effects [16]. This involves extensions to the non-Newtonian fluid model [16] governing the density and viscosity and introducing three additional equations to the code, one for mean temperature and one for each of the two contact surfaces. Inclusion of these into the equation set for the line contact case will be discussed along with the solution method. An extension to the point contact case is also described.

Results are presented and discussed for an essentially steady state line contact showing the thermal effects at both pure rolling and at several shear rates, followed by thermal profiles in a line contact reversal problem, again for both pure rolling and varying shear rates. A thermal point contact solution will then be presented.

2. Notation

Dimensional Values

C_p	spec heat capacity $(Jkg^{-1}K^{-1})$
$C_p : a$	spec heat capacity of roller A $(Jkg^{-1}K^{-1})$

*The authors would like to thank Shell Global Solutions and EPSRC for funding this work

$C_p : b$	spec heat capacity of roller B $(Jkg^{-1}K^{-1})$
p	pressure (Pa)
h	film thickness (m)
ρ	density (Nm^{-1})
ρ_a, ρ_b	density of roller A/B (Nm^{-1})
u_a, u_b	velocity of roller A/B (ms^{-1})
u_e	entrainment velocity (ms^{-1})
η	viscosity (Pas)
θ	temperature (K)
x	distance (m)
t	time (s)
μ	Dowson-Higginson mu parameter (Pa^{-1})
ν	Dowson-Higginson nu parameter (Pa^{-1})
$\beta_{e;0}$	thermal expansion coefficient (K^{-1})
κ_e	pressure coeff of thermal expansion (Pa^{-1})
θ_0	reference temperature (K)
b	Hertzian radius (m)
p_h	maximum Hertzian pressure (Pa)
R	reduced radius (m)
ρ_0	standard density (Nm^{-1})
η_0	standard viscosity (Pas)
α_0	pressure coefficient of viscosity (Pa^{-1})
z_0	Roelands z0 parameter
z_1	Roelands z1 parameter
θ_r	ref temp. in Roelands law (K)
θ_0	ref temp. for density and viscosity (K)
k	thermal conductivity coeff $(Wm^{-1}K^{-1})$
k_a, k_b	thermal con. coeff of A/B $(Wm^{-1}K^{-1})$

Reference Values

u_0	reference velocity (ms^{-1})
η_x	ambient viscosity (Nm^{-1})
ρ_x	ambient density (Nm^{-1})
θ_x	ambient temperature (K)

Non-Dimensional Values

B	Hertzian radius	$(B = 2b/R)$
H	film thickness	$(H = Rh/b^2)$
P	pressure	$(P = p/p_h)$
T	time	$(T = u_0 t/2b)$
X	distance	$(X = x/b)$
U_a	velocity of roller A	$(U_a = u_a/u_0)$
U_b	velocity of roller B	$(U_b = u_b/u_0)$
U_e	entrainment velocity	$(U_e = u_e/u_0)$
U_m	mean velocity	$(U_m = u_m/u_0)$
$\overline{\rho}$	density	$(\overline{\rho} = \rho/\rho_x)$
$\overline{\eta}$	viscosity	$(\overline{\eta} = \eta/\eta_x)$
$\overline{\theta}$	temperature	$(\overline{\theta} = \theta/\theta_x)$
Γ_m	mean shear rate	$(\Gamma_m = \lvert U_b - U_a \rvert / H)$

3. Governing Equations

There are five governing equations for the standard non-thermal line contact EHL problem, those of pressure, film thickness, force balance, density and viscosity. In addition there are three equations linked with the thermal model, these being the energy equation and a surface temperature equation for each surface.

The pressure distribution is modelled by the Reynolds Equation [17]. In non-dimensional form this is:

$$\frac{\partial(\overline{\rho}H)}{\partial T} = \frac{\partial}{\partial X}\left(\frac{\overline{\rho}H^3}{\overline{\eta}\lambda}\frac{\partial P}{\partial X}\right) - U_e(T)\frac{\partial(\overline{\rho}H)}{\partial X}, \quad (1)$$

where

$$U_e(T) = U_a(T) + U_b(T), \quad (2)$$

and

$$\lambda = \frac{6\eta_x u_e R^2}{b^3 p_h}. \quad (3)$$

The Film Thickness Equation represents the geometry of the roller and is given, in non-dimensional form, by:

$$H(X) = H_0 + \frac{X^2}{2} - \frac{1}{\pi}\int_{X_a}^{X_b} \ln|X - X'|P(X')\mathrm{d}X', \quad (4)$$

where H_{00} is the central offset film thickness, which defines the relative positions of the surfaces if no deformation was to occur. The parabolic term represents the undeformed shape of the rollers, and the integral defines the deformation of the surface due to the pressure distribution across the domain.

The Force Balance Equation relates the pressure distribution across the domain to the applied load. It is given by:

$$\int_{-\infty}^{\infty} P(X)\mathrm{d}X = \frac{\pi}{2}. \quad (5)$$

The density of the fluid is normally modelled as a function of pressure only using the Dowson and Higginson [18] model. In this work we consider the density as being a function of both pressure and temperature and use an extended form of the Dowson and Higginson model in dimensional form:

$$\rho(p,\theta) = \rho_0 \left(1 + \frac{\mu p}{1+\nu p}\right) \{1 - \beta_e(p)(\theta - \theta_0)\}, \quad (6)$$

where

$$\beta_e(p) = \beta_{e;0} \exp(-\kappa_e p). \quad (7)$$

Similarly to the density, in this work the viscosity is considered to be a function of both pressure and of temperature, instead of merely pressure. As a result the viscosity model used is a modified version of the Roelands law [19], this is:

$$\eta(p,\theta,\dot{\gamma}) = \eta_0 \exp(y), \quad (8)$$

where

$$y = \frac{\alpha_0 p_0}{z_0} \left\{ \left(1 + \frac{p}{p_0}\right)^{z(\theta)} \left(\frac{\theta - \theta_r}{\theta_0 - \theta_r}\right)^{-s} - 1 \right\}, \quad (9)$$

$$z(\theta) = z_0 - z_1 \ln\left(\frac{\theta - \theta_r}{\theta_0 - \theta_r}\right), \quad (10)$$

and

$$s = \frac{\beta_0 z_0}{\alpha_0 p_0}(\theta_0 - \theta_r). \quad (11)$$

Temperature varies not only in the direction of flow through the contact which is solved for but also in the direction across the flow, between the rollers which is not solved for. Consequently an approximation must be made to take this into account. We assume a parabolic temperature approximation [15].

The appropriate Energy Equation for EHL problems [16] assuming the parabolic temperature approximation is given in non-dimensional form as:

$$\overline{\rho}\left\{\frac{1}{2}\frac{\partial \overline{\theta}}{\partial T} + U_m \frac{\partial \overline{\theta}}{\partial X} + \frac{(\overline{\theta} - \overline{\theta_b})(U_m - U_b)}{H}\frac{\partial H}{\partial X}\right\}$$
$$= \frac{3\overline{k}}{2B^2H^2}(\overline{\theta_a} + \overline{\theta_b} - 2\overline{\theta}) + \overline{\beta_e}\theta\left(\frac{1}{2}\frac{\partial P}{\partial T} + U_m\frac{\partial P}{\partial X}\right)$$
$$-2B\overline{\mu}\left(U_m - \frac{U_e}{2}\right)\frac{\partial P}{\partial X} + \frac{B\overline{\mu}\overline{\kappa}}{3}\overline{\eta}\Gamma_m^2, \quad (12)$$

where

$$U_m = -\frac{H^2}{2\overline{\kappa}\overline{\eta}}\frac{\partial P}{\partial X} + \frac{U_e}{2}, \quad (13)$$

$$\overline{\mu} = \frac{p_h R}{C_p \rho_x \theta_x b}, \quad (14)$$

$$\overline{\kappa} = \frac{6\eta_x u_r R^2}{b^3 p_h}, \quad (15)$$

$$\overline{\beta_e} = \frac{p_h \beta_e}{C_p \rho_x}, \quad (16)$$

and

$$\overline{k} = \frac{k}{C_p \rho_x u_0 b}. \quad (17)$$

The term on the left is heat convection, while the terms on the right of the equation model thermal conductivity, compression heating, viscous heat generation due to Pouseille flow and viscous heat generation due to Couette flow.

There are two equations for surface temperature, one for each surface. In the pure rolling case, these will be identical. The equations are given in non-dimensional terms as:

$$\overline{\theta_a(X)} = 1 + 2\frac{\overline{\kappa\chi_a}}{\sqrt{U_a}}\int_{-\infty}^{X} \frac{(3\overline{\theta} - 2\overline{\theta_a} - \overline{\theta_b})}{(X - \bar{\zeta})^{1/2} H(X)} \mathrm{d}\bar{\zeta}, \quad (18)$$

and

$$\overline{\theta_b(X)} = 1 + 2\frac{\overline{\kappa\chi_b}}{\sqrt{U_b}}\int_{-\infty}^{X} \frac{(3\overline{\theta} - \overline{\theta_a} - 2\overline{\theta_b})}{(X - \bar{\zeta})^{1/2} H(X)} \mathrm{d}\bar{\zeta}, \quad (19)$$

where

$$\overline{v} = \frac{b^3 p_h}{12 R^2 \eta_x \nu}, \quad (20)$$

$$\overline{\kappa} = \frac{k}{\rho_0 C_p \nu b}, \quad (21)$$

$$\overline{\nu} = \frac{6R^2}{b^2}, \quad (22)$$

$$\overline{\chi_a} = \frac{\rho_0 C_p \nu R}{\sqrt{\pi \rho_a C_{p:a} k_a u_0 b}}, \quad (23)$$

$$\overline{\chi_b} = \frac{\rho_0 C_p \nu R}{\sqrt{\pi \rho_b C_{p:b} k_b u_0 b}}. \quad (24)$$

Extension for Point Contact

The equations for the isothermal point contact model are well documented. After non-dimensionalisation, additional y-directional terms occur in the Reynolds equations, the force balance equation is altered and there are changes to the integral in the film thickness equation. The density and viscosity equations remain unchanged.

The Reynolds equation (1) becomes,

$$\frac{\partial(\overline{\rho}H)}{\partial T} = \frac{\partial}{\partial X}\left(\frac{\overline{\rho}H^3}{\overline{\eta}\lambda}\frac{\partial P}{\partial X}\right) + \frac{\partial}{\partial Y}\left(\frac{\overline{\rho}H^3}{\overline{\eta}\lambda}\frac{\partial P}{\partial Y}\right) - \frac{U_e(T)}{U_e(0)}\frac{\partial(\overline{\rho}H)}{\partial X}, \quad (25)$$

The Force Balance equation in the point contact case is given by:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(X,Y)\mathrm{d}X\mathrm{d}Y = \frac{2\pi}{3}. \quad (26)$$

The film thickness equation is given by,

$$H(X,Y) = H_{00} + \frac{X^2}{2} + \frac{Y^2}{2} - \frac{2}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{P(X',Y')\mathrm{d}X'\mathrm{d}Y'}{\sqrt{(X-X')^2+(Y-Y')^2}}, \quad (27)$$

The mean energy equation has extra terms in the two dimensional case. The two-dimensional form of this equation is given by,

$$\overline{\rho}\left\{\frac{1}{2}\frac{\partial\overline{\theta}}{\partial T} + U_m\frac{\partial\overline{\theta}}{\partial X} + V_m\frac{\partial\overline{\theta}}{\partial Y} + \frac{\overline{\theta}-\overline{\theta_b}}{H}\left((U_m-U_b)\frac{\partial H}{\partial X} + V_m\frac{\partial H}{\partial Y}\right)\right\} = \frac{3k}{2B^2H^2}(\overline{\theta_a}+\overline{\theta_b}-2\overline{\theta}) + \frac{B\overline{\mu}\kappa}{3}\overline{\eta}\Gamma_m^2 + \overline{\beta_e\theta}\left(\frac{1}{2}\frac{\partial P}{\partial T} + U_m\frac{\partial P}{\partial X} + V_m\frac{\partial P}{\partial Y}\right) - 2B\overline{\mu}\left\{\left(U_m - \frac{U_e}{2}\right)\frac{\partial P}{\partial X} - V_m\frac{\partial P}{\partial Y}\right\}, \quad (28)$$

where

$$V_m = -\frac{H^2}{2\overline{\kappa\eta}}\frac{\partial P}{\partial Y}. \quad (29)$$

4. Discretisation

A uniform mesh is used with N elements where $N = 2^k + 1$. The level of refinement on a grid can then be described as grid level k refinement. The equations are discretised using first order upwind differencing, or second order central differencing where appropriate. This discretisation has been documented elsewhere [20] for the non thermal equations, so here we shall only consider the energy and two surface temperature equations.

The Energy Equation (12) is discretised as:

$$\overline{\rho_i}\left\{\frac{(\overline{\theta_i^{t_n}} - \overline{\theta_i^{t_{n-1}}})}{2\Delta T} + U_m\frac{(\overline{\theta_i}-\overline{\theta_{i-1}})}{\Delta X} + \frac{(\overline{\theta_i}-\overline{\theta_b})}{H}\left((U_m-U_b)\frac{H_i-H_{i-1}}{\Delta X}\right)\right\} = \frac{3k}{2B^2H_i^2}(\overline{\theta_a}+\overline{\theta_b}-2\overline{\theta_i}) + \overline{\beta_e\theta_i}\left(\frac{(P_i^{t_n}-P_{i-1}^{t_{n-1}})}{2\Delta T} + U_m\frac{(P_i-P_{i-1})}{\Delta X}\right) - 2B\overline{\mu}\left(U_m - \frac{U_e}{2}\right)\frac{(P_i-P_{i-1})}{\Delta X} + \frac{B\overline{\mu}\kappa}{3}\overline{\eta}\Gamma_m^2, \quad (30)$$

where,

$$U_m = -\frac{H_i^2}{2\overline{\kappa\eta}}\frac{(P_i-P_{i-1})}{\Delta X} + \frac{U_e}{2}. \quad (31)$$

The Surface Temperature Equations (18) and (19) become:

$$\overline{\theta_{a;i}} = 1 + 2\overline{\kappa\chi_a}\sum_{j=1}^{i}\left\{\frac{\overline{\theta_{A;j-1}}+\overline{\theta_{A;j}}}{H_{j-1}+H_j}2\{(X_j-X_{j-1})^{1/2} - (X_i-X_j)^{1/2}\}\right\}, \quad (32)$$

and

$$\overline{\theta_{b;i}} = 1 + 2\overline{\kappa\chi_b}\sum_{j=1}^{i}\left\{\frac{\overline{\theta_{B;j-1}}+\overline{\theta_{B;j}}}{H_{j-1}+H_j}2\{(X_j-X_{j-1})^{1/2} - (X_i-X_j)^{1/2}\}\right\}, \quad (33)$$

where,

$$\overline{\theta_A} = 3\overline{\theta} - 2\overline{\theta_a} - \overline{\theta_b}, \quad (34)$$

and

$$\overline{\theta_B} = 3\overline{\theta} - \overline{\theta_a} - 2\overline{\theta_b}. \quad (35)$$

5. Solution Method

The equations, excluding the thermal equations, are solved using multigrid techniques in in the standard manner as described in [21,20,22] and utilising the multilevel multi-integration algorithm of Brandt and Lubrecht [23].

The temperature, both at the surface and in the fluid is calculated only on the finest mesh. First the surface temperatures are calculated using equations (32) and (33). This can be an expensive process as at each point a summation along the line up to that point is required. Furthermore the calculation at each point is expensive when repeated. Following the surface calculation, the Energy Equation (30) is calculated at each point using a Newton method. The relaxation factor is initially small but can be allowed to increase once the solution is converging.

In the multigrid method the temperature is also required on the coarser grids. This is obtained by coarsening the temperature on the fine mesh, no temperature solve being carried out on these meshes.

6. Results

Moes Parameter M	27.68
Moes Parameter L	9.53
Speed Parameter U	1.6×10^{-11}
Material Parameter G	4000
Load Parameter W	1.57×10^{-4}
Pressure-viscosity index α	$2.0\times10^{-8}\ Pa^{-1}$
Maximum Hertzian Pressure p_h	$1.0\times10^{9}\ Pa$
Minimum Domain Bound	$-4.6\times10^{-3}\ m$
Maximum Domain Bound	$4.6\times10^{-3}\ m$
Number of Points	2049

Table 1
Operating Parameters for Line Contact Case

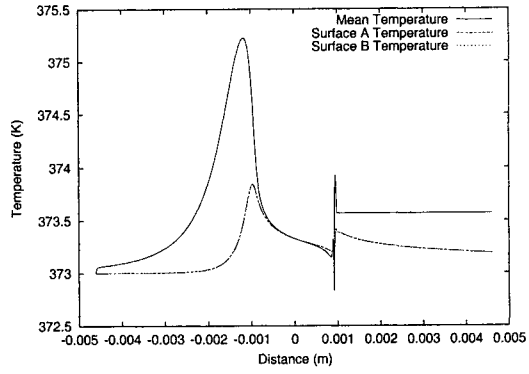


Figure 1. Temperature Profile in Pure Rolling Case

6.1. Steady State Line Contact

Figure 1 shows the mean fluid and surface temperature profiles for a pure rolling case, as described in Table 1, with both rollers always travelling with an identical velocity of $3.0ms^{-1}$. The mean temperature can be seen to smoothly rise to a peak upstream of the contact region, before rapidly decreasing as the contact region is entered. The temperature then slowly decreases across the contact region until a spike coinciding with the exit of the contact region. The surface temperature follows the same profile but at a lower temperature and translated slightly further downstream than the mean temperature profile.

If small amounts of shear occur then very different solution profiles are obtained. Figure 2 shows three mean fluid temperature solutions, the first solution is for pure rolling and the subsequent two are with slight amounts of shear. The structure of the pure rolling profile can barely be made out as the rise in temperature in the middle of the contact region is so large. The shear profile does exhibit these features but they remain of the same size as in the rolling case.

6.2. Transient Line Contact

The behaviour of the model in Table 1 under transient conditions is investigated using a reversal case. Each roller at $t = 0$ has a velocity and over a time period this is gradually changed until

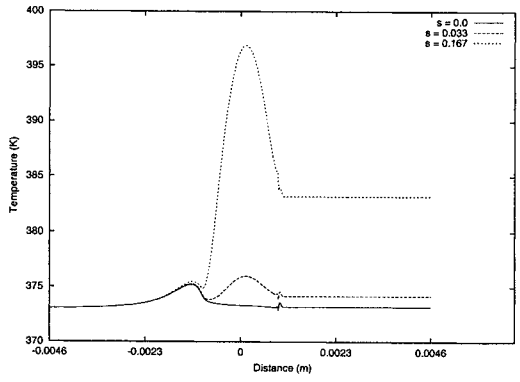


Figure 2. Temperature Profile in Shear Cases

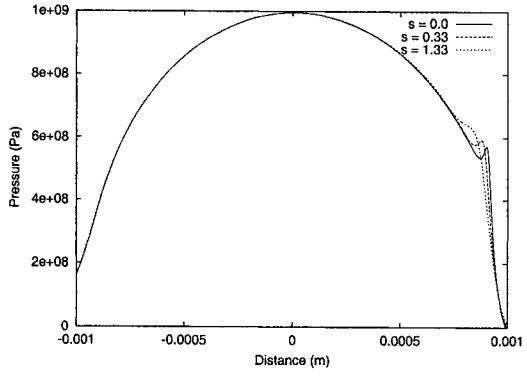


Figure 3. Pressure Profile in Shear Cases

finally it is travelling at the opposite velocity.

In the first case pure rolling is considered with both rollers always travelling at identical velocities to each other. Initially they are travelling at $3.0ms^{-1}$ and decrease linearly over a period of $0.2s$ to a final velocity of $-3.0ms^{-1}$. There are 500 timesteps of $0.0004s$ taken.

In the second and third cases shear is introduced, at a low level and at a high level. The second model has the rollers initially with velocities of $2.5ms^{-1}$ for Roller A and $3.5ms^{-1}$ for Roller B, the third with velocities of $1.0ms^{-1}$ for Roller A and $5.0ms^{-1}$ for Roller B. In both cases these velocities vary linearly with time over a period of $0.2s$ to final opposite velocities. There are 500 timesteps of $0.0004s$ taken.

Figures 4 show the temperature profiles for all three cases at 5 points during the reversal run. As can be seen the temperature of the model with the largest shear dominates the profiles displaying a maximum rise in excess of 100 degrees. This decreases as the velocity is reduced and at the point of reversal the temperature has reduced to ambient throughout the domain.

A similar effect can be seen with the temperatures at the surface of the roller. Figures 5 show the surface temperatures on the two rollers. The temperature profile on each roller is clearly the same as for the other roller, but at different values. Roller A, the slower of the two rollers clearly has lower temperature profiles than the faster roller.

7. Steady Point Contact

Figure 6 shows two mean fluid temperature solutions in the point contact case with operating parameters detailed in Table 2. The lower of the two solutions is for the pure rolling case, the higher solution is for a case with a small amount of shear. The pure rolling model has both rollers travelling with a velocity of $4.0ms^{-1}$, the shear case has roller velocities of $4.01ms^{-1}$ and $3.99ms^{-1}$ giving a shear rate of $s = 0.005$. In the pure rolling solution the temperature can be seen to rise upstream of the contact region, falling to ambient within this region, before a spike at the exit of the contact region. The case with the shear

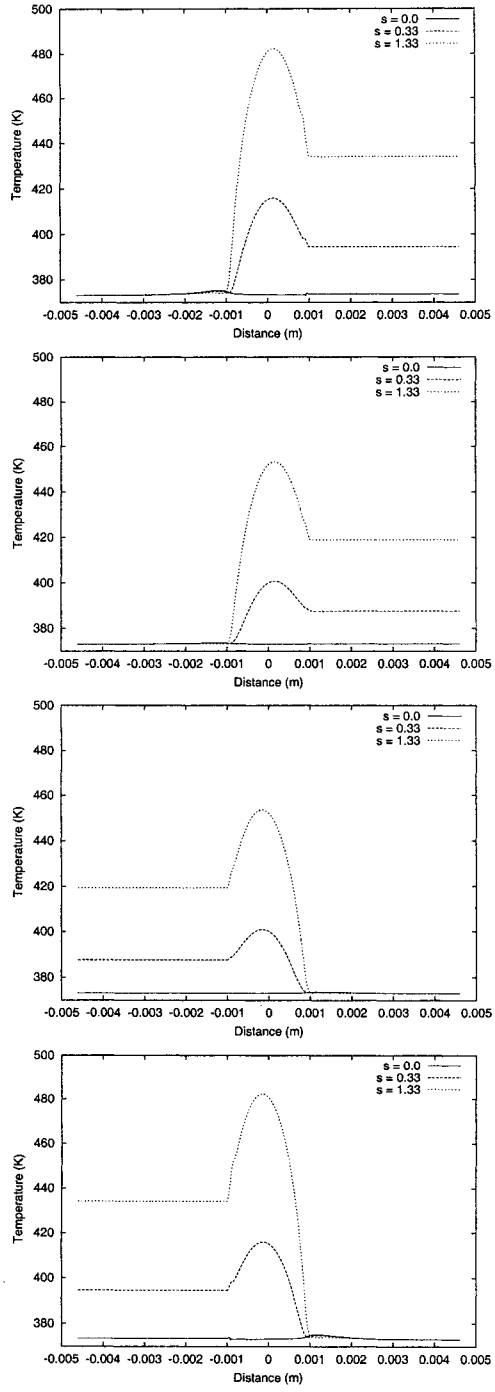


Figure 4. Mean Fluid Temperature Profiles at t=0.0s,0.05s,0.15s,0.20s

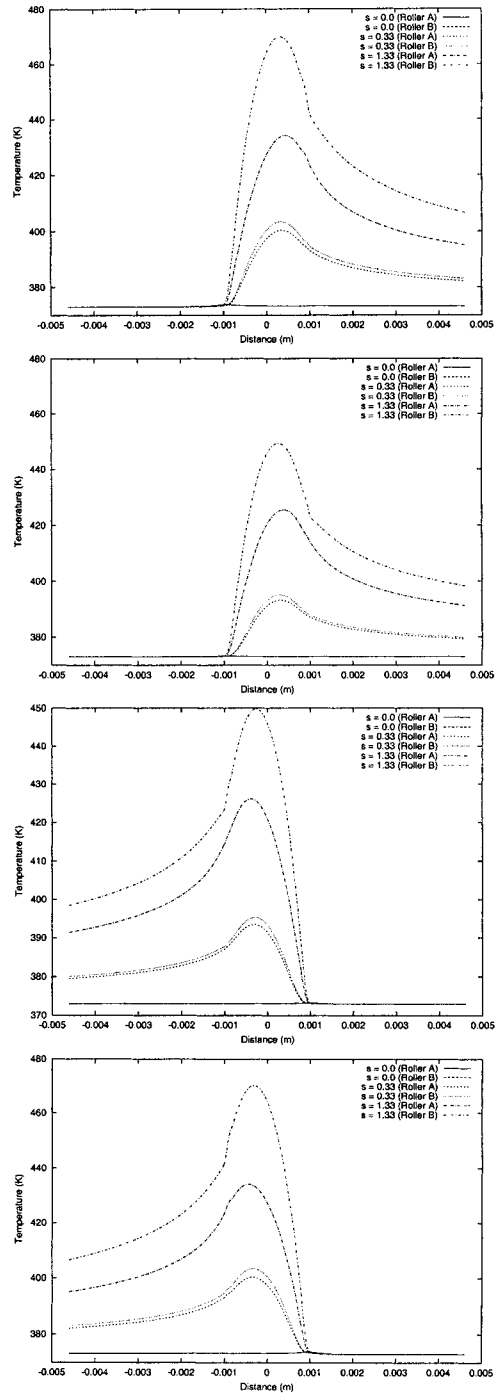


Figure 5. Surface Temperature Profiles at t=0.0s,0.05s,0.15s,0.20s

Moes Parameter M	33.38
Moes Parameter L	17.90
Speed Parameter U	1.6×10^{-10}
Material Parameter G	4200
Load Parameter W	2.58×10^{-6}
Pressure-viscosity index α	$2.1\times10^{-8}Pa^{-1}$
Maximum Hertzian Pressure p_h	$1.0\times10^{9}Pa$
Minimum Domain Bound	$-4.86\times10^{-3}m$
Maximum Domain Bound	$2.50\times10^{-3}m$
Number of Points	257×257

Table 2
Operating Parameters for Point Contact Case

displays the same features except that there is a large temperature rise in the centre of the contact region relative to the rolling case, the ambient temperature being $373K$ and the peak temperature $384K$.

8. Conclusions

It has been shown that an energy equation can be successfully incorporated into the EHL solver of [14] to obtain mean and surface temperature solutions for both steady state and transient cases. The importance of shear in an EHL contact has also been demonstrated with even small amounts of shear producing temperatures far in excess of those found within the contact under pure rolling conditions. This increase in temperature can also be seen to influence the spike in the pressure profile.

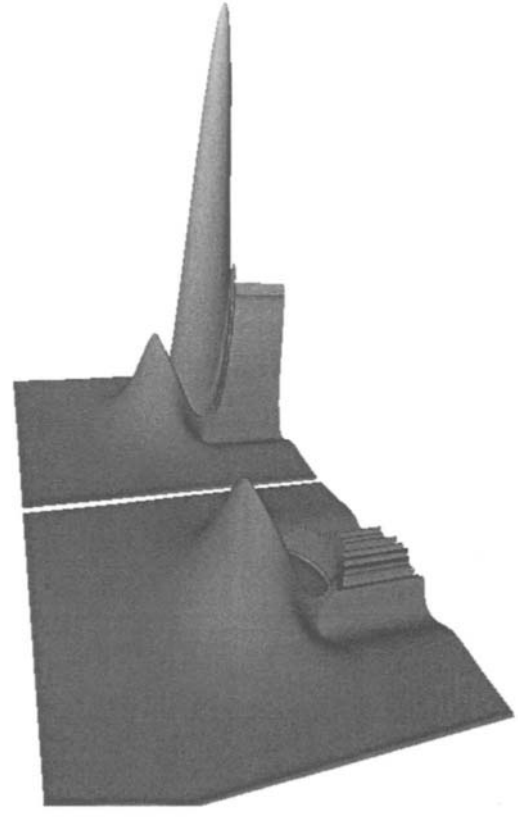


Figure 6. Mean Fluid Temperature Profiles

REFERENCES

1. Ausherman, V.K., Nagaraj, H.S., Sanborn, D.M. and Winer, W.O., Infrared temperature mapping in elastohydrodynamic lubrication, *Trans. ASME, Journal of Tribology*, 1993, 115, 102-110.
2. Crook, A.W., The lubrication of rollers. Part III., *Phil. Trans. R. Soc. Lond. A*, 1961, 83, 213-226.

3. Sternlicht, B., Lewis, P. and Flynn, P., Theory of lubrication and failure of rolling contacts, *ASME J.Basic Eng*, 1961, 83, 213-226.
4. Cheng, H.S and Sternlicht, B., A numerical solution for the pressure, temperature and film thickness between two infinitely long, lubricated rolling and sliding cylinders, under heavy loads, *Trans. ASME, J.Basic Eng*, 1965, 103, 695-707.
5. Cheng, H.S., A refined solution to the thermal-elastohydrodynamic lubrication of rolling and sliding cylinders, *ASLE Trans.*, 1965, 8, 397-410.
6. Dowson, D. and Whitaker, A.V., A numerical procedure for the solution of the elastohydrodynamic problem of rolling and sliding contacts lubricated by a Newtonian fluid, *Proc. Inst. Mech. Eng.*, 1965, 180, 57-71.
7. Lee, R.T. and Hsu, C.H., Advanced multilevel solution for thermal elastohydrodynamic lubrication of simple sliding line contacts, *Wear*, 1994, 171, 227-237.
8. Lee, R.T. and Hsu, C.H., Multilevel solution for thermal elastohydrodynamic lubrication of rolling/sliding circular contacts, *Tribology International*, 1995, 28(8), 541-552.
9. Kim, H.J., Ehret, P., Dowson, D. and Taylor, C.M., Thermal elastohydrodynamic analysis of circular contacts. Part 1: Newtonian model, *Proc Instn Mech Engrs, Part J, J. of Engineering Tribology*, 2001, 215, 339-352.
10. Kim, H.J., Ehret, P., Dowson, D. and Taylor, C.M., Thermal elastohydrodynamic analysis of circular contacts. Part 2: non-Newtonian model, *Proc Instn Mech Engrs, Part J, J. of Engineering Tribology*, 2001, 215, 353-362.
11. Kazama, T., Ehret, P. and Taylor, C.M., On the effects of the temperature profile approximation in thermal Newtonian solutions of elastoydrodynamic lubrication line contacts, *Proc Instn Mech Engrs, Part J, J. of Engineering Tribology*, 2001, 215(J1), 109-120.
12. Venner, C.H. and Lubrecht, A.A., Multilevel methods in lubrication, *Tribology Series, 37*, ed D.Dowson, Elsevier, 2000.
13. Nurgat, E., Berzins, M. and Scales, L., Solving EHL problems using iterative, multigrid, and homotopy methods, *Trans. ASME, J. Tribology*, 1999, 121(1), 28-34.
14. Goodyer, C.E., Fairlie, R., Berzins, M. and Scales, L.E., Adaptive techniques for elastohydrodynamic lubrication solvers, *Tribology Research : From Model Experiment to Industrial Problem, Proceedings of the 27th Leeds-Lyon Symposium on Tribology*, ed G.Dalmaz et al, Elsevier, 2001.
15. Salehizadeh, H. and Saka, N., Thermal non-Newtonian elastohydrodynamic lubrication of rolling line contacts, *Trans. ASME, J. Tribology*, 1991, 113, 481-491.
16. Scales, L.E., Quantifying the Rheological Basis of Traction Fluid Performance, *SAE Technical Paper Series*, No.1999-01-3610, 1999.
17. Reynolds, O., On the theory of lubrication and its application to Mr Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil, *Philosophical Transactions of the Royal Society of London*, 1886, 177, 157-234.
18. Dowson, D. and Higginson, G.R., Elastohydrodynamic Lubrication, The Fundamentals of Roller and Gear Lubrication., Permagon Press, Oxford, Great Britain, 1966.
19. Roelands, C.J.A., Correlational Aspects if the viscosity-temperature-pressure relationship of lubricating oils., *PhD thesis*, Technische Hogeschool Delft, The Netherlands, 1966.
20. Goodyer, C.E., Elastohydrodynamic Lubrication, *PhD thesis*, University of Leeds, Leeds, England, 2002.
21. Venner, C.H., Multilevel Solution of the EHL Line and Point Contact Problems, *PhD thesis*, University of Twente, The Netherlands, 1991.
22. Goodyer, C.E., Fairlie, R., Berzins, M. and Scales, L.E., An in-depth investigation of the multigrid approach for steady and transient EHL problems, *Thinning films and Tribological Interfaces: Proceedings of the 26th Leeds-Lyon Symposium on Tribology*, ed D.Dowson et al, Elsevier, 2000.
23. Brandt, A. and Lubrecht, A.A., Multilevel matrix multiplication and fast solution of integral equations, *Journal of Computational Physics*, 1990, 90(2), 348-370.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

ANALYSIS OF THE OIL DISTRIBUTION IN STARVED ELLIPTICAL E.H.D. CONTACTS

B. Damiens[a], P.M.E. Cann[b], C.H. Venner[c], A.A. Lubrecht[a].
[a] Laboratoire de Mécanique des Contacts, UMR CNRS 5514, INSA de Lyon, France
[b] Department of Mechanical Engineering, Tribology Section, Imperial College, London SW7 2BX, UK.
[c] University of Twente, Tribology group, Enschede, The Netherlands

1. INTRODUCTION

In a recent paper, the authors analyzed starved lubrication of elliptical contacts [1]. The aim was to predict quantitatively starved film decay as a function of contact operating conditions. In this respect the work was an extension of earlier analysis of circular contacts by Chevalier et al. [2].

Although the elliptical contact analysis allowed quantitative prediction of the starvation level, it failed to provide a single variable capable of describing starvation as a function of ellipticity. However, the elliptical analysis did show that the degree of starvation can vary substantially with position across the contact. This variation was not found for circular contacts and hence a simple transformation from circular to elliptical is not possible.

The current paper therefore explores some of the issues raised in the elliptical contact analysis. The aims were: (a) to provide additional insight into the local character of starvation in elliptical contacts. (b) to quantify the amount of oil lost from the track in the inlet and outlet regions. (c) to analyze the local oil film redistribution in the inlet region

The numerical calculations were performed using a Multi-Level algorithm [3].

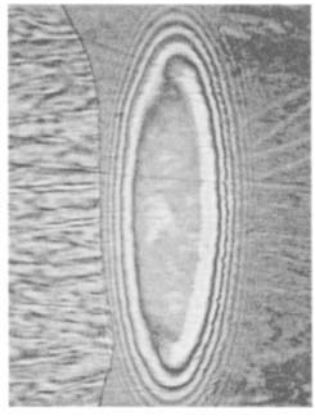


Figure 1. *Starved elliptical contact.*

NOMENCLATURE

a	ellipse axis in rolling direction $a = (3\mathrm{w}R/E')^{1/3}(2\kappa\mathcal{E}/\pi)^{1/3}$
b	ellipse axis perpendicular to rolling direction $b = a/\kappa = (3\mathrm{w}R/E')^{1/3}(2\kappa\mathcal{E}/\pi)^{1/3}/\kappa$
$\mathcal{D}$	domain $X_0 \leq X \leq X_e, -Y_0 \leq Y \leq Y_0$
E'	reduced modulus of elasticity $2/E' = (1-\nu_1^2)/E_1 + (1-\nu_2^2)/E_2$
G	material parameter, $G = \alpha E'$
h	film thickness
h_c	central film thickness
H	dimensionless film thickness $H = h(2R\mathcal{E})/(a^2\mathcal{K})$
H_c	dimensionless central film thickness
H_{cff}	dimensionless fully flooded central film thickness
H_0	dimensionless rigid body displacement
H_{oil}	dimensionless inlet oil film thickness
H_{out}	dimensionless outlet oil film thickness
H_{outff}	dimensionless fully flooded outlet oil film thickness
L	dimensionless material parameter (Moes)

$L = G(2U)^{1/4}$

m parameter of the elliptic integrals $\mathcal{K}$ and $\mathcal{E}$
$m = (1-\kappa^2)^{1/2}$

M dimensionless load parameter (Moes)
$M = W(2U)^{-3/4}$

n number of disk revolutions (overrollings)

p pressure

p_h maximum Hertzian pressure
$p_h = (3\mathrm{w})/(2\pi ab)$

p_0 constant (Roelands), $p_0 = 1.96\ 10^8$

P dimensionless pressure, $P = p/p_h$

R reduced radius of curvature
$R = (R_x^{-1} + R_y^{-1})^{-1}$

R_x reduced radius of curvature in x direction
$R_x = (R_{x1}^{-1} + R_{x2}^{-1})^{-1}$

R_y reduced radius of curvature in y direction
$R_y = (R_{y1}^{-1} + R_{y2}^{-1})^{-1}$

$R_x/R_y = \kappa^2(\mathcal{K} - \mathcal{E})/(\mathcal{E} - \kappa^2\mathcal{K})$

r_c relative oil film thickness
$r_c = H_{oil}/(\bar{\rho}H_{cff})$

r_{out} relative oil film thickness
$r_{out} = H_{oil}/(H_{outff})$

$\mathcal{R}_c$ central film thickness reduction,
$\mathcal{R}_c = H_c/H_{cff}$

$\mathcal{R}_{out}$ total film thickness reduction,
$\mathcal{R}_{out} = H_{out}/H_{outff}$

s inlet length $s \propto rs_{ff}$

$\overline{S}$ dimensionless inlet length
$\overline{S} \propto r\overline{S}_{ff} \propto r(L/M)^{1/2}$

s_{ff} fully flooded inlet length

$\overline{S}_{ff}$ dimensionless fully flooded inlet length
$\overline{S} = s_{ff}/a \propto (L/M)^{1/2}$

u_m mean velocity in x, $u_m = (u_1 + u_2)/2$

U dimensionless speed parameter
$U = (\eta_0 u_m)/(E'R_x)$

w external load

W dimensionless load parameter
$W = \mathrm{w}/(E'R_x^2)$

X, Y dimensionless coordinates
$X = x/a$, $Y = y/b$

X_0, X_e, Y_0
dimensionless domain boundaries

z pressure viscosity index (Roelands)

α pressure viscosity coefficient

$\bar{\alpha}$ dimensionless parameter $\bar{\alpha} = \alpha p_h$
$\bar{\alpha} = \frac{L}{\pi}\left(\frac{3M\pi^2\kappa}{2}\frac{(1+R_x/R_y)^2}{16\mathcal{E}^2}\right)^{1/3}$

δH_{in} inlet film thickness reduction

δH_{out} outlet film thickness reduction

λ dimensionless parameter
$\lambda = \pi\left(\frac{128}{3M^4}\frac{16\pi\mathcal{E}^5}{\kappa^4(1+R_x/R_y)^5\mathcal{K}^6}\right)^{1/3}$

$\overline{\lambda}_{oil}$ dimensionless oil profile wave length with respect to b

η_0 viscosity at ambient pressure

$\bar{\eta}$ dimensionless viscosity, $\bar{\eta} = \eta/\eta_0$

ν_1, ν_2 poisson ratio of bodies 1 and 2

ρ density

ρ_0 density at atmospheric pressure

$\bar{\rho}$ dimensionless density, $\bar{\rho} = \rho/\rho_0$

θ filling rate, ratio of the oil film thickness and the gap height

κ ellipticity ratio $\kappa = a/b$

$\mathcal{K}(m) = \int_0^{\pi/2} 1/\sqrt{1 - m^2\sin^2(\psi)}d\psi$

$\mathcal{E}(m) = \int_0^{\pi/2} \sqrt{1 - m^2\sin^2(\psi)}d\psi$

$\mathcal{S} = (\mathcal{E} - \kappa^2\mathcal{K})/(\mathcal{K} - \kappa^2\mathcal{K})$

2. STARVATION

The starved regime is characterised by a lack of lubricant in the inlet of the contact. An incomplete oil film profile is imposed in the inlet. Experimentally, this oil profile corresponds to the oil profile found in the outlet of the previous pass. From a numerical point of view, it constitutes a boundary condition and the meniscus position is treated as a free boundary problem. θ represents the ratio of the oil layer thickness (H_{oil}) and the total gap (H) between the surfaces ($\theta = H_{oil}/H$) see figure 2. In the inlet, the gap H between the surfaces decreases as X increases while the oil thickness H_{oil} remains constant. Hence the fractional film thickness θ increases steadily until a step occurs ($\theta = 1$). The lack of lubricant delays the pressure generation which starts at the point where the full film is formed (at the meniscus position). Figure 2 (obtained numerically with the model presented in appendix A), shows the variations of the different parameters: the pressure P, the fractional film θ, and the gap H between the surfaces along the X axis, in the plane $Y = 0$. The shape of the oil layer in the same plane is shown in figure 3.

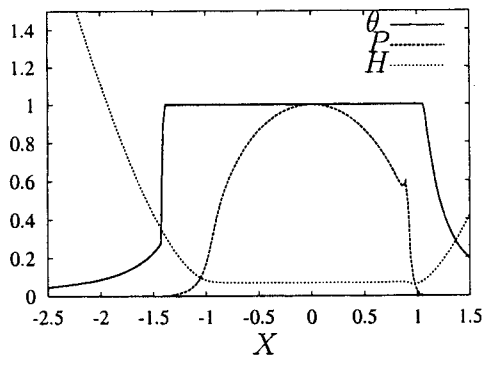


Figure 2. *Centreline profile $Y = 0$ of an elliptical contact $\kappa = 0.63$, $M = 300$, $L = 10$ and $r_c = 1$.*

Traditionally the central film thickness prediction is the main goal of numerical calculations. However, this work focuses on the oil film loss from inlet to outlet. Consequently, the ratio of the inlet oil layer thickness H_{oil}, the central film thickness H_c and the outlet oil layer H_{out} was studied, see figure 3. In this figure one observes a reduction of the oil layer thickness going from inlet to outlet. The difference between H_{oil} and H_{out} represents the oil removed from the track by a single contact pass. By varying the inlet oil layer H_{oil}, the evolution of the contact starvation as a function of multiple passes can be modelled.

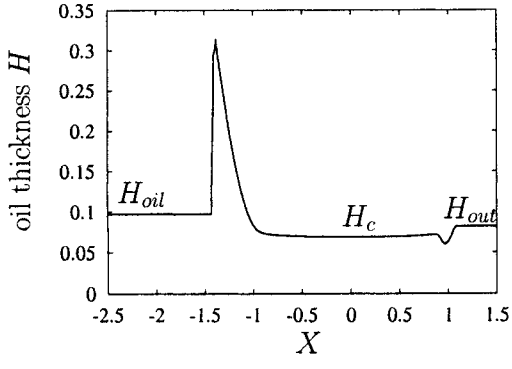


Figure 3. *Centreline oil profile $h(x, y = 0)$ for an elliptical contact $\kappa = 0.63$, $M = 300$, $L = 10$ and $r_c = 1$.*

As a first step the oil removal of a single pass is studied as a function of the amount of oil available in the inlet. In figure 4 the relative oil thickness r_c accounts for the fully flooded central film thickness and the relative compressibility. r_c ranges from 0 to ∞. The relative central film thickness $\mathcal{R}_c$ accounts for the fully flooded film thickness and ranges from 0 to 1. The function has two asymptotes. On the starved asymptote, the film thickness increases linearly with the oil layer $\mathcal{R}_c = r_c$. On the fully flooded asymptote, the central film thickness $\mathcal{R}_c = 1$ is independent of the oil layer thickness.

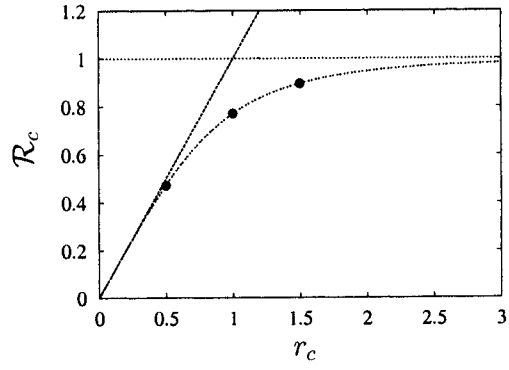


Figure 4. *$\mathcal{R}_c = f(r_c)$ for a circular contact with $M = 10$ and $L = 20$, $\gamma = 2.66$.*

Chevalier [2] studied the circular contact and showed that a good approximation to the function $\mathcal{R}_c = f(r_c)$ is given by equation (1).

$$\mathcal{R}_c = \frac{r_c}{\sqrt[\gamma]{1 + r_c^{\gamma}}} \tag{1}$$

In [1] Damiens et al. noticed the linear relation between the parameter γ and a dimensionless length $\bar{S} \propto (L/M)^{1/2}$ characteristic of the inlet zone (see figure 5). The parameter γ can be considered as a resistance to side-flow. With increasing ellipticity (decreasing κ), it becomes more difficult for the lubricant to flow around the contact and thus γ increases. For an infinitely long line contact, no side-flow is possible and γ will be infinite. This tendency is shown in figure

5 where γ is plotted as a function of $(M/L)^{1/2}$ for different ellipticities.

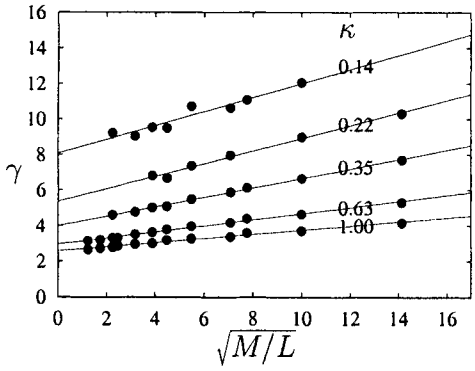


Figure 5. γ *as a function of* $\sqrt{M/L}$ *for different ellipticities:* $\kappa = 0.14, 0.22, 0.35, 0.63, 1.00$, $r_c \simeq 1$.

Using equation (1) recursively, γ can be directly related to the film thickness reduction as a function of the number of passes allowing experimental validation [1]. For the circular contact, good agreement was found between the numerical and experimental results. However, for the elliptical contact the agreement was less satisfactory. A number of factors can explain this difference:

- equation (1) is only an approximation
- γ oil removal occurs in inlet and outlet
- starvation can be non uniform across the contact

In [4], Notary and Lubrecht studied the starvation behaviour of the rigid isoviscous elliptical contact. They gave a more accurate approximation of the starvation curve using a fraction of two polynomials. However, the improved precision is obtained at the cost of a dozen constants. It is unlikely that a simpler relation can be found for the piezoviscous-elastic (EHL) starvation curve. Hence we will stick with equation (1)in spite of its mediocre performance to approximate the oil ejection in all the starved regime. The rest of this paper will treat the points 2 and 3.

3. EJECTION IN INLET AND OUTLET

Previous studies on starvation [2] have focussed on the lubricant loss from the inlet region. This is a valid approximation for fully flooded and mildly starved conditions, but not for the severely starved cases. As the contact becomes increasingly starved, the pressure distribution approaches Hertz. Hence, the pressure gradients in inlet and outlet become comparable, see figure 6. As the film thickness distribution is similar in inlet and outlet, this means that the oil removal in inlet and outlet regions become comparable.

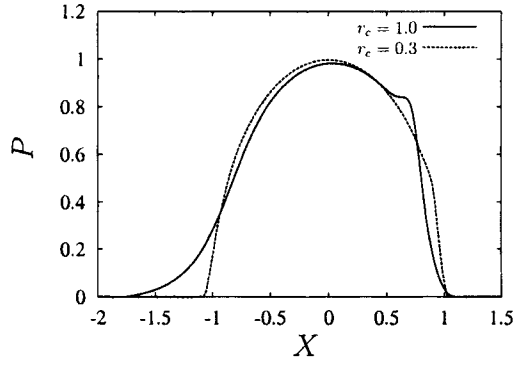


Figure 6. *Centreline pressure profile for 2 oil layer thicknesses* $r_c = 1.0$ *and* $r_c = 0.3$ *with* $\kappa = 0.35$, $M = 100$ *and* $L = 5$.

Figure 7 shows the relative amount of oil lost in the outlet region, compared to the overall oil lost. In this case the inlet covers the region from $X = -\infty$ to $X = 0$, whereas the outlet is defined by positive X values. On the central line ($Y = 0$), the inlet oil ejection is defined as: $\delta H_{in} = H_{oil} - \bar{\rho} H_c$. In the same way the outlet ejection is defined as: $\delta H_{out} = \bar{\rho} H_c - H_{out}$. Please note that $\bar{\rho} > 1$, hence, $\delta H_{out} > 0$ even if $H_c < H_{out}$. Since a pressure gradiant in the Y direction tends to empty the track, a positive amount of oil δH_{out} is removed from the centerline going from the contact center to the outlet meniscus. The percentage of the oil ejection taking place in the outlet can be defined as:

$$\%\text{outlet ejection} = 100 * \frac{\delta H_{out}}{\delta H_{in} + \delta H_{out}} \quad (2)$$

Figure 7 shows its variation as a function of $(M/L)^{1/2}$ for an elliptical contact ($\kappa = 0.35$) and two different starvation conditions. The parameter $(M/L)^{1/2}$ determines the ejection process only up to a first order. Furthermore, the values of the ejection are computed numerically, and contain a certain numerical error. This combination causes the variations observed in figure 7. Nevertheless, the global trend which is of interest in this section is clear and consistent. For large amounts of oil ($r_c \simeq 1$) only 5% of the lubricant ejection occurs in the outlet. Whereas roughly 40% of the lubricant ejection occurs in the outlet for thin oil layers ($r_c \simeq 0.3$). This shows the increased importance of outlet ejection under heavily starved conditions. Asymptotically, for values of r_c going to 0, the pressure distribution becomes symmetrical between inlet and outlet. As a result, the ejection in the inlet and outlet becomes comparable. A similar trend is observed in circular contacts.

As a consequence, the oil ejection in the outlet can be neglected for moderately starved conditions $r \simeq 1$. However, for severely starved contacts $r < 0.5$, the ejection in the outlet is almost as important as the one occurring in the inlet, and hence cannot be neglected.

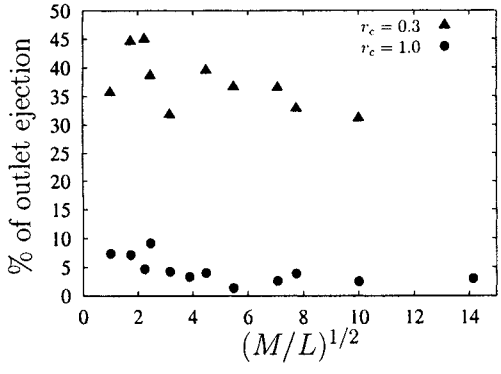


Figure 7. *Proportion of the outlet ejection for the oil layers $r_c = 0.3$ and $r_c = 1.0$ for $\kappa = 0.35$.*

4. OIL PROFILE SMOOTHING

In practice, the oil layer in the Y direction is not constant. Due to, for example, manufacturing defects, the oil layer can be close to zero at certain points (see Chevalier et. al. [5]). In such a case, it is important to know if the lubricant available in the neighbourhood of the defect is mobile enough to replenish the local lubricant shortage. Hence, it is important to study the lubrication coupling between points placed at different Y positions across the width of a starved elliptical contact.

This section uses a non smooth oil profile $H_{oil}(Y)$ described by equation (3) to feed the inlet of the contact. $\overline{\lambda}_{oil}$ is the dimensionless wave length of the oil profile, it is the real wave length divided by b, the half width of the contact in the Y direction. Note that the mean oil film thickness is r, and that locally zero oil film thickness exists in the inlet. In order to describe the lubricant mobility, one focuses on the ratio of the waviness amplitudes between inlet and outlet.

$$H_{oil}(Y) = r * \left[1 + \cos\left(\frac{2\pi Y}{\overline{\lambda}_{oil}}\right)\right] \quad (3)$$

The oil redistribution that occurs in the contact takes place in the low pressure region where Poiseuille flow occurs. Like the oil ejection, the redistribution should be observed in the inlet and outlet. The coupling C between two points of the profile decreases as their distance $(b\overline{\lambda}_{oil})$ increases. The further the points are from each other, the more independent they will be, the less the starvation level in one point will influence the lubrication in the other. The coupling will increase when the inlet length increases because as the inlet length increases, the low pressure region (in which the redistribution is assumed to occur) increases. Hence, the governing parameter in the coupling phenomena should be the ratio of these two distance. If the inlet length s is assumed to be proportional to the fully flooded inlet length multiplied by the amount of oil r on the track the coupling parameter can be expressed as :

$$C = \frac{s}{b\overline{\lambda}_{oil}} \propto \frac{ra(L/M)^{1/2}}{b\overline{\lambda}_{oil}} = \frac{r(L/M)^{1/2}}{\overline{\lambda}_{oil}/\kappa} \quad (4)$$

Figure 8 shows the evolution of the ratio of the waviness amplitudes in the inlet and outlet as a function of the coupling C. The points were calculated with the model presented in appendix A for average oil layer thicknesses of $r = 0.1, 0.3, 0.5$, ellipticities of $\kappa = 0.1, 0.2, 0.5, 1.0$, with the Moes numbers $M = 10, 30, 100, 300, 1000$ and $L = 5, 10, 20$, and oil profile wavelength of $\overline{\lambda}_{oil} = 0.1, 0.2, 0.5$.

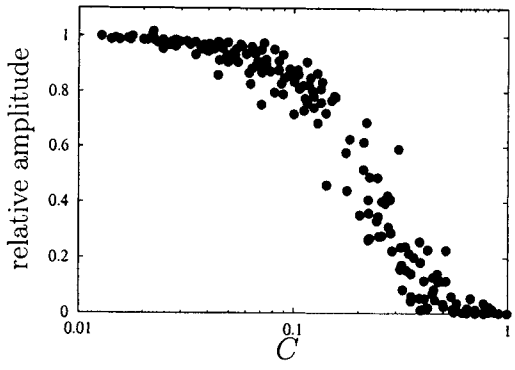


Figure 8. *Ratio of the amplitude in the inlet and outlet as a function of the coupling C*

The fact that the calculated points are more or less lying on a single curve indicates that the parameter C is governing the coupling under starved conditions. The ability of the parameter C to predict the coupling could be improved if a better approximation of the inlet length was used.

5. CONCLUSIONS

The starvation behaviour of circular and elliptical contacts has been studied and compared numerically. Increasing the ellipticity (decreasing κ) increases the resistance to side flow (γ). The oil ejection mechanisms are modified when the starvation degree increases. As starvation increases (r decreases), the relative film thickness reduction increases (γ decreases up to a factor 2 in elliptical contacts). The oil ejection location moves from a single site in the inlet (in fully flooded and moderately starved contacts) to an ejection shared between two sites of similar importance, inlet and outlet, as starvation increases (r decreases). A coupling (stronger for the circular contact case) seems to flatten the γ_{out} profiles, rendering the oil ejection more or less uniform across the contact. This lubrication coupling is an important parameter to understand local lubrication failure. A parameter C describing this coupling is found to depend on the contact operating conditions M and L, the contact ellipticity κ, the starvation level r and the lubricant defect width $\overline{\lambda}_{oil}$.

APPENDIX

A. MODEL EQUATIONS

To describe the lubricant flux in the contact, the Elrod model [6,7] is used. Wijnant [8] studied the starved elliptical contact and derived the corresponding dimensionless Reynolds equation (6). Capillary forces were neglected in this analysis.

The Elrod model divides the calculational domain into two sub-domains: In the first subdomain, the pressure P is zero and the lubricant film is incomplete. Two lubricant layers separated by an air layer stick to the rolling element and plane. The lubricant travels with the moving surfaces. The filling rate θ, which is the ratio of the thickness of the oil layers and the total gap between the surfaces, is smaller than 1.

In the second sub-domain, the film is complete, the pressure is positive and a lubricant side flow is observed.

in sub-domain 1, $0 < \theta < 1$ and $P = 0$ and in sub-domain 2, $\theta = 1$ and $P > 0$. The behaviour can be condensed in equation (5) :

$$P(1 - \theta) = 0 \tag{5}$$

The oil profile is imposed in the inlet of the contact. The position of the boundary (meniscus) between the two sub-domains is found by solving simultaneously the system of equations composed by the load equation (8), the modified Reynolds equation (6), the complementarity equation (5)

and the equation of elastic deformations (8). The problem is a free boundary problem.

The dimensionless parameters for an elliptical contact are defined in the nomenclature. The resulting dimensionless Reynolds equation reads:

$$\frac{\partial}{\partial X}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial X}\right)+\kappa^2\frac{\partial}{\partial Y}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial Y}\right) - \frac{\partial(\bar{\rho}\theta H)}{\partial X}=0 \qquad (6)$$
$$\mathcal{D}: X_0 \le X \le X_e, -Y_0 \le Y \le Y_0$$

With the boundary conditions: $P(X_0,Y) = P(X_e,Y) = P(X,-Y_0) = P(X,Y_0) = 0$ and $\theta(X_0,Y) = H_{oil}(Y)/H(X_0,Y)$.

The dimensionless gap between the surfaces is given by:

$$H(X,Y) = H_0 + \mathcal{S}X^2 + (1-\mathcal{S})Y^2 + \frac{1}{\pi\mathcal{K}}\iint_{\mathcal{D}}\frac{P(X',Y')\,dX'dY'}{\sqrt{\kappa^2(X-X')^2+(Y-Y')^2}} \qquad (7)$$

The dimensionless rigid body displacement H_0 is coupled to the force balance equation, which, for an elliptical contact reads:

$$\iint_{\mathcal{D}} P(X,Y)\,dXdY = \frac{2\pi}{3} \qquad (8)$$

The viscosity pressure relation proposed by Roelands [9] is used:

$$\bar{\eta}(P) = \exp\left\{\frac{\alpha p_0}{z}(-1+(1+\frac{Pp_h}{p_0})^z)\right\} \qquad (9)$$

with

$$\frac{\alpha p_0}{z} = \ln\eta_0 + 9.67 \qquad (10)$$

The compressibility is taken into account using the density pressure relation proposed by Dowson and Higginson [10]:

$$\bar{\rho}(P) = \frac{0.59\ 10^9 + 1.34Pp_h}{0.59\ 10^9 + Pp_h} \qquad (11)$$

REFERENCES

1. B. Damiens, C. H. Venner, P. M. E. Cann, and A. A. Lubrecht. Starved lubrication of elliptical e.h.d. contacts. *accepted at the ASME, Journal of Tribology*, 2003.
2. F. Chevalier, A. A. Lubrecht, P. M. E. Cann, F. Colin, and G. Dalmaz. Film thickness in starved ehl point contacts. *ASME, Journal of Tribology*, 120:126 - 133, 1998.
3. C. H. Venner and A. A. Lubrecht. *Multi-Level Methods in Lubrication.* Elsevier, ISBN 0-444-50503-2, 2000.
4. M. P. Notary and A. A. Lubrecht. Starved lubrication of isoviscous rigid circular contacts. *presented at the 29th Leeds Lyon symposium on Tribology.*
5. F. Chevalier, A. A. Lubrecht, P. M. E. Cann, and G. Dalmaz. Evolution of lubricant film defects in the starved regime. *Proceedings of the 24th Leeds-Lyon Symposium on Tribology*, pages 233 - 244, 1997.
6. H. G. Elrod. A cavitation algorithm. *ASME, Journal of Lubrication Technology*, 103:350 - 354, 1981.
7. H. G. Elrod and M. L. Adams. A computer program for cavitation and starvation problems. *Proceedings of the 1st Leeds-Lyon Symposium on Tribology*, pages 37 - 41, 1974.
8. H. Wijnant, Y. *Contact Dynamics in the field of Elastohydrdynamic Lubrication.* PhD thesis, University of Twente, Enschede, the Netherlands, ISBN : 90-36512239, 1998.
9. C. J. R. Roelands. *Correlation aspects of the viscosity-temperature-pressure relationship of lubricating oils.* PhD thesis, Technical University of Delft, The Netherlands, 1966.
10. D. Dowson and G. R. Higginson. A numerical solution to the elastohydrodynamic problem. *Journal of Mechanical Engineering Science*, 1:6 - 15, 1959.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Ordinary Shear Thinning Behavior and Its Effect Upon EHL Film Thickness

Scott Bair[a] and Farrukh Qureshi[b]

[a]Center for High Pressure Rheology, Georgia Institute of Technology,
George W. Woodruff School of Mechanical Engineering, Atlanta, GA 30332-0405, USA

[b]Physical & Analytical Sciences, The Lubrizol Corporation,
29400 Lakeland Boulevard, Wickliffe, OH 44092-2298, USA

Measurements are presented of ordinary shear thinning behavior of liquid lubricants using laboratory rheometers. Flow curves are constructed and compared with theoretical models. The data can be fitted to the well-known Carreau equation. This model is used to calculate rolling/sliding film thickness under EHD line contact for conditions typical of tribological contacts. This work provides useful insights regarding the ordinary shear thinning behavior of liquids typical of lubricants, and its impact on EHD film forming behavior of lubricants.

1. INTRODUCTION

The rather spectacular success of Newtonian based models in predicting the film thickness of mineral oils in rolling elastohydrodynamic line [1] and point [2] contacts may have made us less aware of the failure of these models in combined rolling/sliding contacts lubricated by liquids with high molecular weight additives and by neat liquids of only moderate molecular weight. Indeed, many pressure-viscosity coefficients published today are derived from rolling film thickness measurements and, of course, when these coefficients are utilized with Newtonian based formulae for the same kinematics, the film thickness prediction will be accurate. Unfortunately, only rarely are film thickness derived coefficients compared with viscosity derived coefficients. Jones [3] found that for a high molecular weight perfluoropolyalkyl ether, viscosity derived pressure-viscosity coefficients were about twice the magnitude of the film derived coefficients. Film thickness calculations [4] based upon the results of high-pressure rheometry clearly showed that shear-thinning of the Carreau form was responsible.

Typical flow curves for base oils and polymer blends are presented and compared with theory. Examples of film thickness calculations for ordinary shear-thinning are given to obtain some insight concerning real liquid behavior in contact.

2. EXPERIMENTAL MATERIALS AND TECHNIQUE

Three well-characterized liquids were utilized as idealized base oils. These are 1, 1-(1, 1, 3-Trimethyl-1, 3-propanediyl) bis-cyclohexane denoted as BCH, bis(2-ethylhexyl) phthalate denoted as DOP, and Dimethylpolysiloxane, a liquid polymer denoted as DMS. All were obtained from Sigma-Aldrich.

Idealized polymer-blended oils were prepared using squalane (2, 6, 10, 15, 19, 23-hexamethyl tetracosane) from Sigma-Aldrich as the base oil. The polymers were polyisobutylene from Polymer Standards Service. Their properties are listed in Table 1.

Table 1. Properties of Polyisobutylene used as an Additive in Squalane

Lot Number:	Pib 3	Pib 12027
Weight Average Molecular Weight /kg/kmol	12200	3580
Number Average Molecular Weight / kg/kmol	11200	2900
Peak Molecular Weight/kg/kmol	12200	3420
Polydispersity	1.08	1.23

Two high-pressure rotational rheometers were utilized here. It must be understood that substantially elevated pressure is essential to these measurements [5]. Shear-thinning is encountered when the shear rate, $\dot{\gamma}$, multiplied by a relaxation time, $\lambda=\mu/G$, is

$$\dot{\gamma}\lambda = Wi > 1 \tag{1}$$

where Wi is the Weissenberg number and G is a modulus associated with the longest relaxation time. In order to avoid the thermal softening limit to the shear stress that will preclude shear-thinning, the Nahme number,

$$Na = \frac{\beta\dot{\gamma}^2\mu h^2}{K} \tag{2}$$

at $\dot{\gamma} = 1/\lambda$ must be less than about one for simple steady shear. Here, β is the temperature-viscosity coefficient, h is film thickness and K is thermal conductivity of the liquid. Now a new dimensionless group may be defined that determines whether or not thermal softening will prevent shear-thinning in a steady simple shear experiment [5] by combining equations (1) and (2).

$$\mathcal{B} = \frac{\sqrt{Na}}{Wi} = hG\sqrt{\frac{\beta}{\mu K}} \tag{3}$$

In order to observe shear-thinning in steady simple shear $\mathcal{B}<1$. The smallest practical value of h is about 1 μm in a rheometer and the only other variable over which we have substantial control is μ, the limiting low shear viscosity. Viscosity may be made large by decreasing temperature at ambient pressure or by elevating pressure.

Decreasing temperature at ambient pressure is simply not helpful here as we shall see below that G for the base oils has a value greater than one atmosphere. Another way to express equation (1), the criteria for shear-thinning, is $\tau>G$, where τ is shear stress. Then for $Wi>1$, the shear stress must exceed the pressure at ambient conditions for the base oils and the liquid would be in a state of tension along the direction of one principal normal stress. This tension is known to result in fracture [6], a form of liquid failure [7]. Then we are left with high-pressure as the only practical means of observing shear-thinning in base oils.

The pressurized Couette rheometers used here have been described elsewhere [8, 9]. The shearing gap, h, was 2.1 μm in all cases.

3. EXPERIMENTAL RESULTS

Flow curves for the base oils are plotted in Figure 1 as log of reduced stress against log of shear rate. For each curve, a Newtonian regime of slope of one is apparent for reduced stress less than about one and a shear-thinning regime is observed at high reduced stress

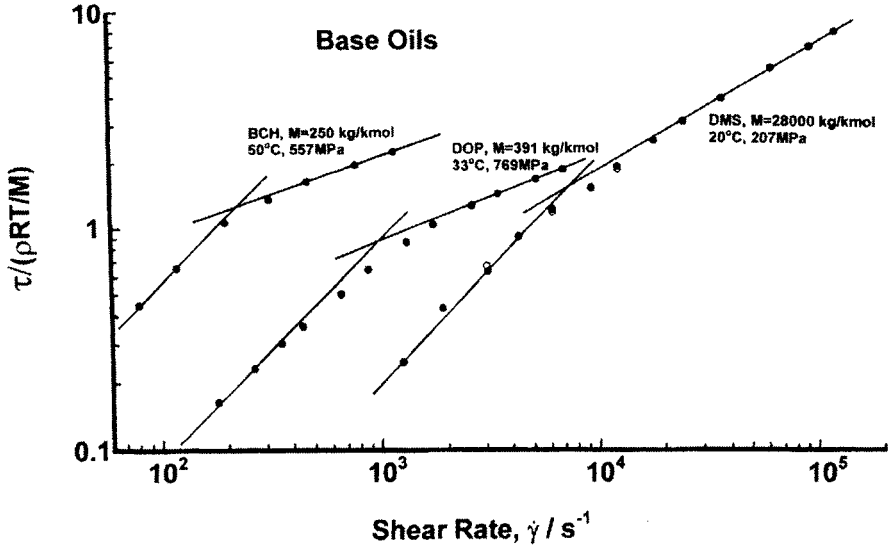


Figure 1. Reduced shear stress versus shear rate for base oils.

where the slope is $n<1$. For BCH, DOP, and DMS, n=0.41, 0.41, and 0.58, respectively. Kinetic theory [10] and non-equilibrium molecular dynamics simulations [11] find that the characteristic relaxation time, λ, as defined in equation (1) should be approximately proportional to molecular weight, M,

$$\lambda \approx \frac{\mu M}{\rho RT} \tag{4}$$

where ρ is mass density, R is the universal gas constant and T is absolute temperature. Then $G \approx \rho RT/M$ and shear-thinning response should occur for $\tau > G$. In Figure 1, the reduced shear stress is τ/G. Now $\rho RT/M$ is 1.08×10^7, 7.6×10^6 and 9.6×10^4 Pa for BCH, DOP and DMS, respectively. Although the shear stress at the transition to shear-thinning varies by two orders of magnitude, this method of reducing shear stress brings the transition to approximately the same position on the vertical axis for the three base oils. Clearly, our results are consistent with kinetic theory.

Flow curves for the polymer solutions are shown in Figure 2, again as log of reduced shear stress versus log of shear rate. Power-law response is apparent at large values of reduced stress and the power-law slope of the low molecular weight solution is n=0.87. For the high molecular weight solutions n varies from 0.85 to 0.94 increasing toward unity with decreasing concentration. For polymers in solution, the density in equation (4) should be the mass density of the polymer in solution. ρ must then be replaced by $c\rho$ where c is the polymer weight concentration and ρ is the density of the solution. Since a measure of molecular weight is required and these polymers are polydisperse, M is replaced by M_w, the weight average molecular weight so that

$$\lambda \approx \frac{\mu M_w}{c\rho RT} \tag{5}$$

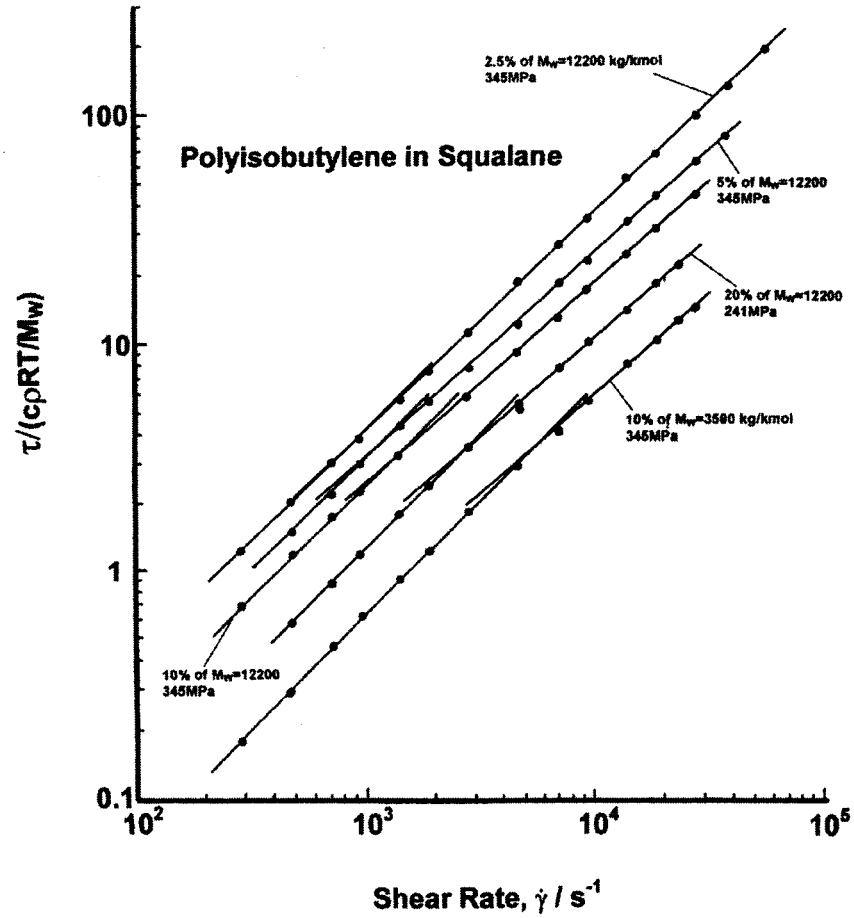


Figure 2. Reduced shear stress versus shear rate for polymer blends.

which leads to $G \approx c\rho RT / M_W$ and the reduced shear stress in Figure 2 is then τ/G. Among these solutions, G varies by a factor of ten, but this method of reducing shear stress brings the shear-thinning transition to essentially the same position along the vertical axis for the five blends. Clearly, these results are consistent with theory as well.

4. FILM THICKNESS CALCULATIONS

Although there have been numerous numerical simulations of EHL contacts lubricated with non-Newtonian liquids, few have assumed for the lubricant shear response the ordinary shear-thinning that can be observed in rheometers. One of the authors has developed a numerical scheme [12] for the calculation of central film thickness that can utilize any generalized Newtonian fluid model for the nonlinear shear response and any general pressure-viscosity response. It is an extremely simple Grubin style inlet zone analysis that, instead of utilizing the Reynolds equation, integrates the constitutive relation across the thickness of the film. The pressure boundary condition that we use for the Hertz boundary has been criticized [13]; however, the film thickness result is not sensitive to the pressure at the Hertz boundary and the results compare well with experiment [4, 12].

In the search for a calibration of this numerical calculation of film thickness consider the model used almost exclusively by tribology, the Prandtl-Eyring law, which is written for the generalized viscosity, η, as

$$\eta = \frac{\tau_O}{\dot{\gamma}} \sinh^{-1}\left(\frac{\mu\dot{\gamma}}{\tau_O}\right). \tag{6}$$

By 1958 Eyring [14] did not, however, believe that liquids behaved in this manner as his own data displayed power-law response [14].

As a comparison consider the sliding point contact results of Vergne, et al. [15]. Here the film thickness was calculated numerically over the entire contact assuming the constitutive equation (6) for various values of τ_o. Their results for central film thickness are shown as the solid points in Figure 3. A simple Grubin style analysis for the same Hertz pressure, P_H, the same pressure-viscosity behavior, and the same reduced radius of curvature as reference [15]was

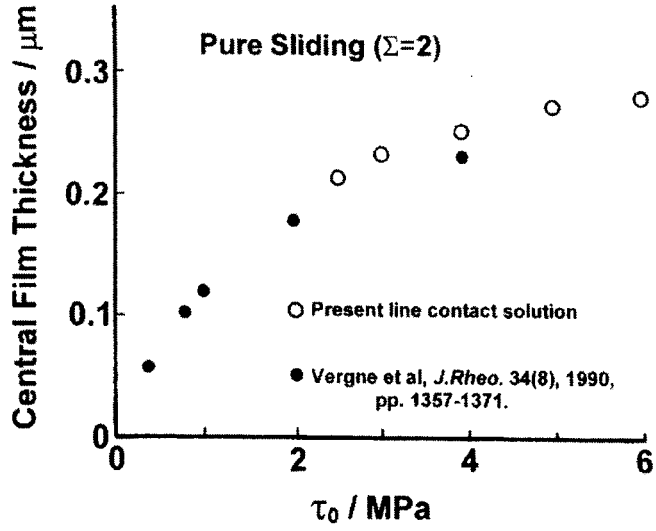


Figure 3. Comparison of film thickness calculations using Prandtl-Eyring.

performed. Since the line contact yields a slightly thicker film than the point contact, the rolling velocity was reduced from the 0.4 m/s used by Vergne et al. to 0.3 m/s to yield the same central film thickness under pure rolling. The results are plotted as the open points in Figure 3, also for pure sliding. Clearly, the trends are similar although this algorithm fails to converge for low values of τ_o.

The film thickness for the ordinary shear-thinning response of lubricants which can be measured is now calculated. There are many generalized Newtonian fluid models that will describe the shear response displayed in Figures 1 and 2 [16]. The Ree-Eyring model utilizes a series in inverse hyperbolic sine to approximate power-law behavior at high shear rate. In others [16] the power-law exponent, n, appears explicitly. Today the most widely used model outside of tribology is the Carreau equation [17] that was advanced to describe the results of molecular network theory.

$$\eta = \mu\left[1+\left(\lambda\dot{\gamma}\right)^2\right]^{(n-1)/2} \qquad (7)$$

If time - temperature - pressure superposition applies, n is constant and relationships for the pressure (and temperature) variation of low shear viscosity, μ, and relaxation time, λ, only must be supplied. For these examples it is assumed that μ varies exponentially with pressure with a pressure-viscosity coefficient of 20 GPa^{-1} and ambient viscosity of 50 mPa·s. The way in which λ varies with pressure (and temperature) is usually known as a shifting rule since it will determine the way in which a family of flow curves for one liquid can be shifted along the shear rate axis to superimpose into a single master flow curve. The Vinogradov-Malkin shifting rule [18] produces a master-curve by plotting η/μ against $\mu\dot{\gamma}$. If $\lambda = \mu/G$ then G must be a constant. Although there are other, perhaps more accurate, shifting rules, the Vinogradov-Malkin rule is utilized for its simplicity as well as the Carreau equation (7) for the example calculations. The reduced radius of the contacting cylinders was set at 0.02m and the Hertz pressure was 1 GPa and the rolling velocity was 1 m/s. The Newtonian solution is h=0.38 μm which compares with the Grubin formula [19] value of 0.42 μm and the Dowson and Higginson formula [19] value of 0.34 μm.

Examples of the film thickness calculations are plotted in Figures 4 and 5 for the conditions specified above for n=3/8, 5/8 and 7/8 and various values of G. For comparison some calculations were performed for the Prandtl-Eyring equation (6) that is preferred by tribologists. These results are shown as film thickness plotted versus slide to roll ratio, Σ, which is defined so that pure sliding represents $\Sigma = 2$ (one surface stationary). When G is 2 MPa and higher, which is typical of mineral oils, the effect of shear-thinning upon film thickness is minimal for pure rolling (Σ=0) and the effect of sliding is gradual except perhaps for n=3/8. For lower values of G film thinning is substantial even for pure rolling. Reducing the power-law exponent, n, reduces the film thickness overall and makes the film more sensitive to sliding.

Consider now the Prandtl-Eyring model at the bottom of Figure 5. In order to have a measureable effect on rolling film thickness, τ_o must be less than 2 MPa, but then the effect of sliding is no longer gradual, but extreme in a manner not seen experimentally.

5. CONCLUSION

Rheometers yield flow curves for lubricants that are consistent with some of the predictions of kinetic theory, and with sufficiently elevated pressure to prevent thermal softening and mechanical failure of the liquid film, shear-thinning of low molecular weight base oils may even be observed. Example calculations of film thickness using the Carreau equation are provided which correctly models ordinary shear-thinning. Clearly, full numerical solutions of the EHL point contact problem with an accurate description of shear-thinning should be ranked highly among the unfinished work in EHL. Some of the other rheology related issues in EHL that have not been sufficiently addressed are:

1) The effect of ordinary shear-thinning on traction.
2) The pressure-viscosity response at high-pressure which is greater-than-exponential.
3) The pressure-viscosity response at high-temperature and low pressure which is often linear with a pressure-viscosity coefficient that increases rapidly with temperature.

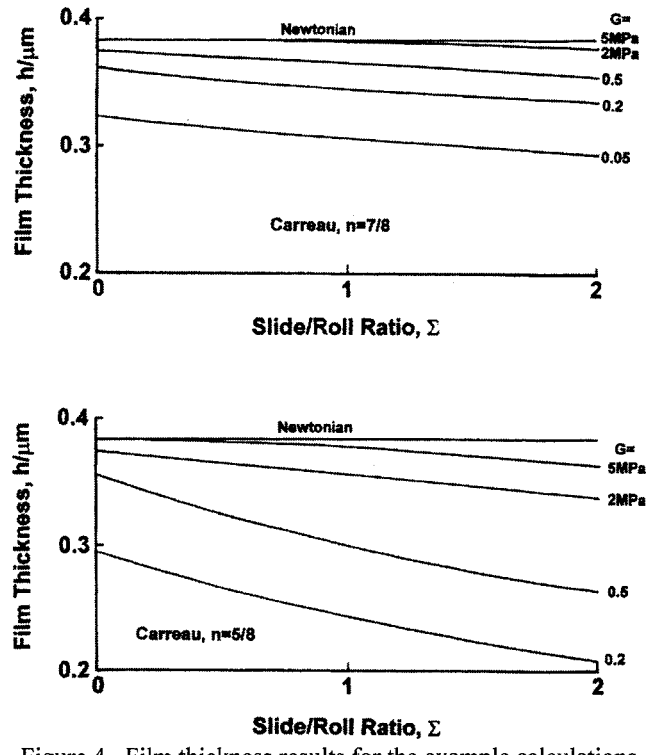


Figure 4. Film thickness results for the example calculations.

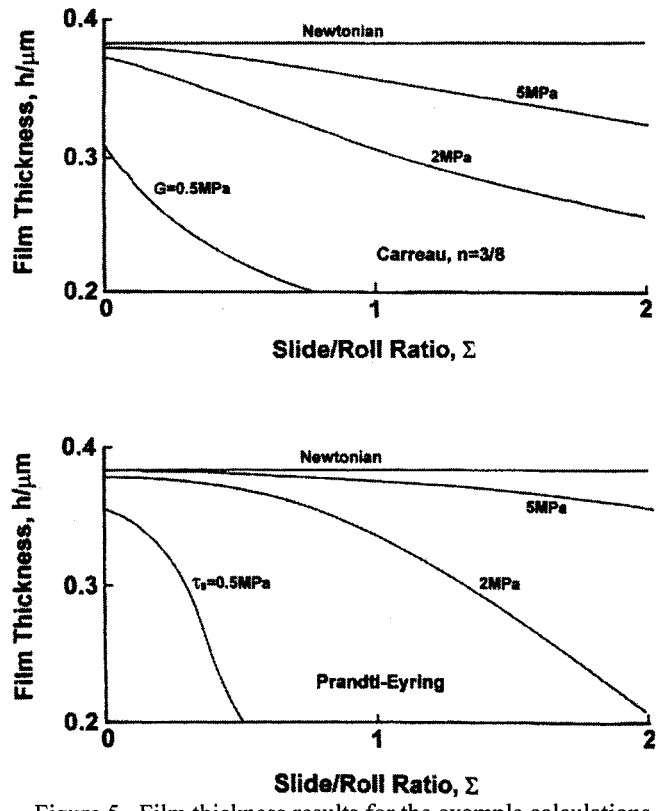


Figure 5. Film thickness results for the example calculations.

REFERENCES

1. Dyson, A., Naylor, H. and Wilson, A.R., "The Measurement of Oil Film Thickness in Elastohydrodynamic Contacts", Proc. Instn. Mech. Engrs., vol. 180, Part 3B, Elastohydrodynamic Symposium, 1965, pp. 119-134.
2. Koye, K.A. and Winer, W.O., "An Experimental Evaluation of the Hamrock and Dowson Minimum Film Thickness Equation for Fully Flooded EHD Point Contacts, Trans. ASME, J. Lubr. Techn., Vol. 103, No. 2, 1981, pp. 284-294.
3. Jones, W.R., "Properties of Perfluoropolyethers for Space Applications", STLE Trib. Trans., Vol. 38, No. 3, 1995, pp. 557-564.
4. Bair, S., Vergne, P. and Marchetti, M., "The Effect of Shear-Thinning on Film Thickness for Space Lubricants", STLE Trib. Trans., Vol. 45, No. 3, 2002, pp. 330-333.
5. Bair, S., "The Shear Rheology of thin Compressed Liquid Films", Proc. Instn. Mech. Engrs., Vol. 216, Part J, 2002, pp. 1-17.
6. Joseph, D.D., "Cavitation and the State of Stress in a Flowing Liquid", J. Fluid Mech., Vol. 366, 1998, pp. 367-378.
7. Archer, L.A., Ternet, D. and Larson, R.G., "Fracture Phenomena in Shearing Flow of Viscous Liquids", Rheologica Acta, Vol. 36, 1997, pp. 579-584.
8. Bair, S. and Winer, W.O., "A New High-Pressure, High-Shear Stress Viscometer and Results for Lubricants", STLE Trib. Trans., Vol. 36, No. 4, 1993, pp. 721-725.
9. Bair, S., "The High Pressure Rheology of Some Simple Model Hydrocarbons", Proc. Instn. Mech. Engrs.Vol. 216, Part J, 2002, pp. 139-149.
10. Bird, R.B., Curtis, C.F., Armstrong, R.C. and Hassager, O., Dynamics of Polymeric Liquids; Vol. 2; Kinetic Theory, Wiley-Interscience, New York, 1987, p. 123.
11. Moore, J.D., Cui, S.T., Cochran, H.D. and Cummings, P.T., "Rheology of Lubricant Base stocks: a Molecular Dynamics Study of C_{30} Isomers", J. Chem. Phys., Vol. 113, No. 19, 2000, pp. 8833-8840.
12. Bair, S., "Elastohydrodynamic Film Forming with Shear Thinning Liquids", ASME J. Tribology, Vol. 120, No. 2, 1998, pp. 173-178.
13. Greenwood, J.A. and Kavzlavich, J.J., "Elastohydrodynamic Film Thickness for Shear Thinning Lubricants", Proc. Instn. Mech. Engrs., Vol. 212, Part J, 1998, pp. 179-191.
14. Ree, F.H., Ree, T.R. and Eyring, H., "Relaxation Theory of Transport Problems in Condensed Systems", I.E.&C., Ind. & Engr. Chem., Vol. 58, 1958, pp. 1036-1040.
15. Vergne, F., Chaomleffel, J.P. and Dalmaz, G., "Elastohydrodynamic Lubrication of Point Contacts with a Ree-Eyring Fluid: Film Thickness and Traction", J. Rheology, Vol. 34, No. 8, 1990, pp. 1357-1371.
16. Bair, S. and Qureshi, F., "The Generalized Newtonian Fluid Model and Elastohydrodynamic Film Thickness", Proceedings of 2002 ASME/STLE Joint International Tribology Conference, 2002.
17. Carreau, P.J., Rheological Equations from Molecular Network Theories, Ph.D. Thesis, University Microfilms, Ann Arbor, 1968, p. 78.
18. Schurz, J., Vollrath-Rodiger, M. and Krassig, H., "Rheologische Untersuchungen an Polyacrylnitril-Spinnlosungen", Rheol. Acta, Vol. 20, 1981, pp. 569-578.
19. Dowson, D. and Higginson, G.R., Elasto-Hydrodynamic Lubrication, Pergamon Press, Oxford, 1966, p. 96.

This Page Intentionally Left Blank

SESSION XIX
TRIBOLOGICAL STUDIES OF DRY AND LUBRICATED SURFACES (3) – LUBRICATED

Chairman : Dr A V Olver

Paper XIX (i) **A Design Tool for Selection and Optimization of Surface Finish in Mixed Lubrication**

Paper XIX (ii) **Starved Lubrication of Isoviscous Rigid Circular Contacts**

Paper XIX (iii) **Gaseous Evolution of Hydrogen from Hydrocarbon Oil and Grease Lubricated Contacts**

Paper XIX (iv) **The Effect of Superimposed Low-frequency Modulation on Lubrication in Machining**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

A Design Tool for Selection and Optimization of Surface Finish in Mixed Lubrication

Dong Zhu

Eaton Corporation, 26201 Northwestern Highway, Southfield, Michigan 48037, USA

This paper presents a deterministic approach developed recently for the numerical simulation of mixed EHD lubrication in rolling/sliding point contacts based on 3-D roughness measured and digitized directly from real engineering surfaces. The computer program has been proven to be robust for the entire transition from full film to mixed EHL and all the way down to boundary lubrication, or even dry contact. The simulation is capable of giving detailed results at asperity scale (or entire distribution) for the hydrodynamic and/or contact pressure, film thickness (or gap), surface and asperity deformation, subsurface stresses, friction and flash temperature at the interface, as well as load sharing between the hydrodynamic film and asperity contacts. Example cases with eight different types of surface finish, including ground, shaved, honed, chemically polished and mechanically super-finished, are analyzed under the same operating conditions. Results are compared with those from ideally smooth surfaces. Different lubrication modes due to different surface finishing processes have been observed. The effect of surface roughness/topography on the lubrication and surface contact characteristics is found to be significant. The detailed results are apparently useful for performance evaluation, efficiency and life improvement, problem diagnosis, failure prevention and production cost minimization. The computer program developed appears to be a useful engineering tool for design optimization and selection of surface finish processes.

1. INTRODUCTION

Most heavy-duty mechanical components, such as gears, bearings, traction drives, cams and followers, piston skirts and rings, typically operate in mixed lubrication with deep surface asperity contacts, where significant frictional loss and various types of surface failure are originated. Today, fast advancing technology and aggravated global competition require manufacturers to continuously upgrade their product design, in order to transmit more power with reduced size/weight and maximized efficiency at minimized cost. This has been a great challenge to design engineers and researchers, and the challenge will certainly become tougher in the future. In continuously variable speed transmissions (CVTs), for example, with a compact geometric design the maximum Hertzian contact pressure between components may possibly reach 5.0-7.0 GPa, while the traction in the mixed EHL contacts must be precisely controlled and pitting failure minimized over wide ranges of speed, load and environmental temperature. With such tough design requirements and under such severe operating conditions, the selection and optimization of surface finish processes become extremely important. Those who want to stay in the leadership position in both the technology and the global market can easily find similar challenge in their product design, no matter what kind of components they produce.

In reality surface finish is typically obtained through machining or polishing processes, and there is a huge variety of surface roughness, topography and fabrication cost. The surface roughness and topography, differentiated by the finish processes, usually have great impacts on the following:

1. contact and lubrication characteristics, such as hydrodynamic and contact pressure/stress, surface deformation, load sharing, lubricant film thickness/gap, subsurface stresses and strains, etc.;
2. interfacial friction, power loss and operating efficiency;
3. surface temperature, thermal behavior and related failures such as scuffing;
4. production cost (as usually final finishing cost is a major or dominant part of total fabrication cost);

5. component life, reliability and warranty cost.

Due to limited computational power and relatively simple numerical algorithms in the past, early studies on the surface contact and lubrication were based either on an assumption of ideally smooth surfaces, or on stochastic models with limited numbers of surface roughness parameters. Those studies, such as presented in [1-7], have laid a foundation for the conventional contact and lubrication theories that have been widely used in industry. Predictions of contact and lubrication characteristics by those theories have so far been useful references for designers. However, since most engineering surfaces are quite rough, predictions based on the smooth surface assumption may not be satisfactory. The stochastic models with only a few roughness parameters, on the other hand, may not be sufficient to give the detailed results on the contact and lubrication characteristics for the real surfaces.

In the past 10-15 years computer technologies have been advancing greatly, so that sophisticated algorithms and large-scale computational analyses can now be executed on personal computers. This enables researchers to simulate rough surface contact and lubrication with deterministic approaches based on full scale 3-D digitized engineering surface topography instead of a few statistic parameters. Also, recently developed numerical algorithms can now overcome the difficulties associated with the EHL non-linearity/instability, which were very difficult or impossible to solve in the past. Early deterministic models include those presented by Lubrecht et al [8], Kweh et al [9], Ai and Cheng et al [10], Venner et al [11] and others, using either artificial surface irregularities of simple geometry or sinusoidal waves. Xu and Sadeghi [12] and Zhu and Ai [13] et al were pioneers employing full-scale 3-D measured roughness in the EHL analyses. It seems that, however, all the solutions published in these previous studies were still in the full film EHL regime with no significant surface asperity contacts presented.

Most recently, Jiang et al [14] published a deterministic EHL solution that handled hydrodynamic lubrication and asperity contacts (up to 12% of Hertzian contact area in presented cases) simultaneously. Shortly after that, Zhu and Hu [15,16] presented a unified approach that is capable of simulating the entire transition from the full film, mixed, down to boundary lubrication. The contact load and area ratios could go from zero up to 100%, while the film thickness ratio, λ, could be reduced from infinity down to zero. The numerical model has been proven to be robust with no convergence problem so far when handling rough engineering surfaces under severe operating conditions (see [17-19]). Based on this approach, a computer program has been developed to predict the EHL and mixed lubrication characteristics, friction, subsurface stresses and flash temperatures, using measured 3-D surface roughness/topography [17].

This paper presents a set of sample cases with eight different types of surface finish, including ground, shaved, honed, chemically polished and mechanically super-finished, under the same operating conditions. Results are compared with each other, also with those from ideally smooth surfaces. The objectives are to better understand the effect of surface finish process on contact and lubrication performance, and to develop an engineering tool for the design, selection and optimization of machining/polishing processes in industrial applications.

2. NUMERICAL SOLUTION METHOD

The present study is based on the deterministic model and the full numerical solution method described in detail in [15-17], which will not be repeated here in order to focus on case discussion. Since both contacting surfaces are rough and moving, the solution is time-dependent in nature. Typically a smooth surface solution under the same conditions is obtained first, then it is used as an initial value for the rough surface solution. After a sufficiently large number of time steps (typically 150-200 steps or more, corresponding to a rolling distance of 2.25~3.0*a or greater, see numerical examples in [18] and [19] as well as in Fig.2 below), the transient solution will become globally stabilized. For the purpose of investigating the effect

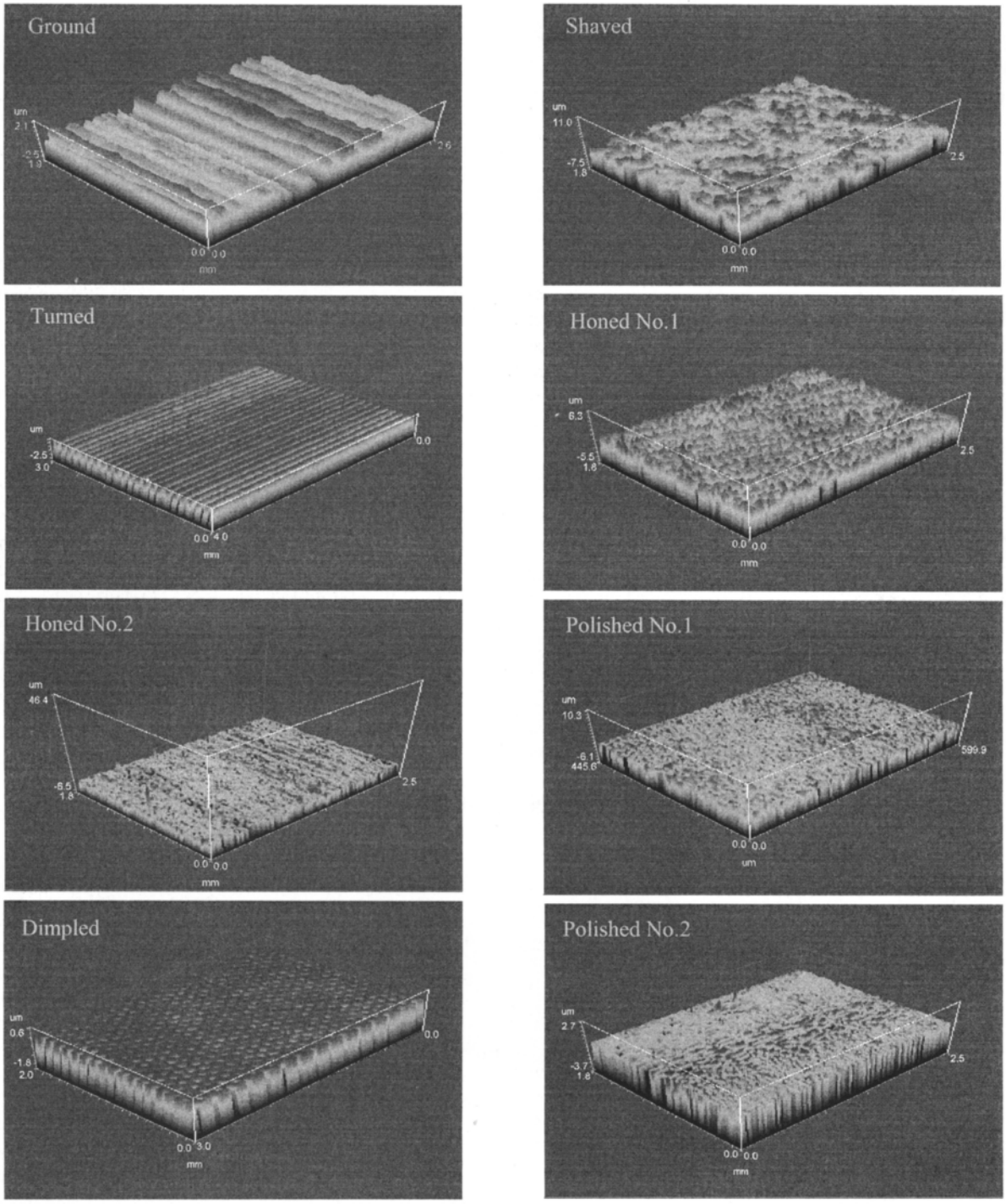


Figure 1. Sample Surfaces Used in the Present Study

of surface finish, all the cases presented below are under the same operating conditions and solved through the same number of time steps (200 steps specifically), which are started from the same smooth solution as the initial value.

Note that the EHL solutions for the film thickness and pressure distributions are isothermal in this paper for the purpose of reducing computing time, because the thermal effect on those is negligible when the speed is not high. However, the models used for the friction and the flash temperature within the contact zone take into account the thermal effect, and a detailed description can be found in [17].

3. CASES ANALYZED

Eight types of machined/polished surfaces are used in the present work in order to study the effect of surface finish/machining process on the contact and lubrication characteristics. In each case, both surfaces are of the same type, but not identical as they are from different measurements at different locations. Figure 1 shows samples of the optically measured 3-D topography for these eight types of surface, and Table 1 gives their RMS roughness data (two surfaces for each case). In all the analyzed cases the applied load is fixed at w=1600 N, the material properties are E'=219.8 GPa, η_o=0.01119 Pa.s, and α=14.94 GPa^{-1}. The geometric parameters are R_x=R_y= 19.005 mm, so that the Hertzian contact radius is a=0.5953 mm, and the maximum Hertzian pressure is P_h= 2.155 GPa. The rolling speed and the slide-to-roll ratio are U=7.5 m/s and S=0.25, respectively. Note that those surfaces represent a good variety of surface roughness and topography, as well as manufacturing technique and production cost. The RMS roughness ranges from 0.15 to 1.15 μm, which covers those obtained by the majority of final finishing processes currently used in industries.

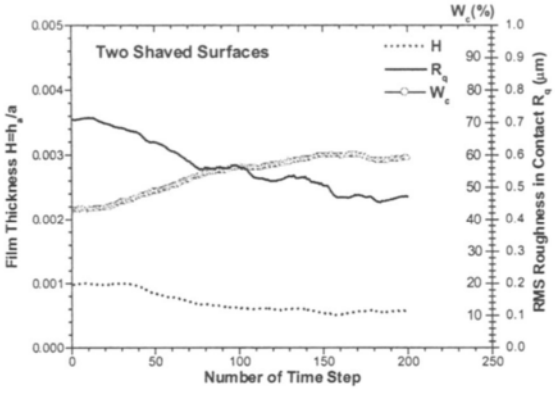


Figure 2. Transient Rough Surface Solution

Table 1. Eight Types of Machined/Polished Surfaces

Surface Type	RMS Roughness (μm)		Composite RMS Roughness (μm)
	Surface 1	Surface 2	
Ground	1.1400	1.1500	1.6193
Shaved	0.8864	0.9477	1.2976
Turned	0.8612	0.7882	1.1674
Honed 1	0.5312	0.7032	0.8813
Honed 2	0.4581	0.7116	0.8463
Polished 1	0.2802	0.3044	0.4137
Dimpled	0.2266	0.2201	0.3201
Polished 2	0.1499	0.1558	0.2162

In each case two rough surfaces of the same type are moved through the EHL conjunction with 25% of relative sliding. Computation is conducted under the same operating conditions through the same number of time steps. Figure 2 shows an example, in which two shaved surfaces were used, and the calculation was started from the smooth surface solution, described in detail below, as the initial value. It is observed that after about 150 time steps the transient solution becomes globally stabilized, but small fluctuations still exist due to moving surface irregularities.

4. RESULTS AND DISCUSSION

4.1. Smooth Surface Solution

The smooth surface solution under the operating conditions mentioned above is obtained first. It can be seen from Figure 3(a) and Figure 4(a) that the pressure distribution is quite smooth, close to the Hertzian with no pressure spike observed due to the heavy contact load. The two surfaces are completely separated with no direct contact, even though the lubricant film is very thin. Friction is calculated by using Bair and Winer's rheological model (see

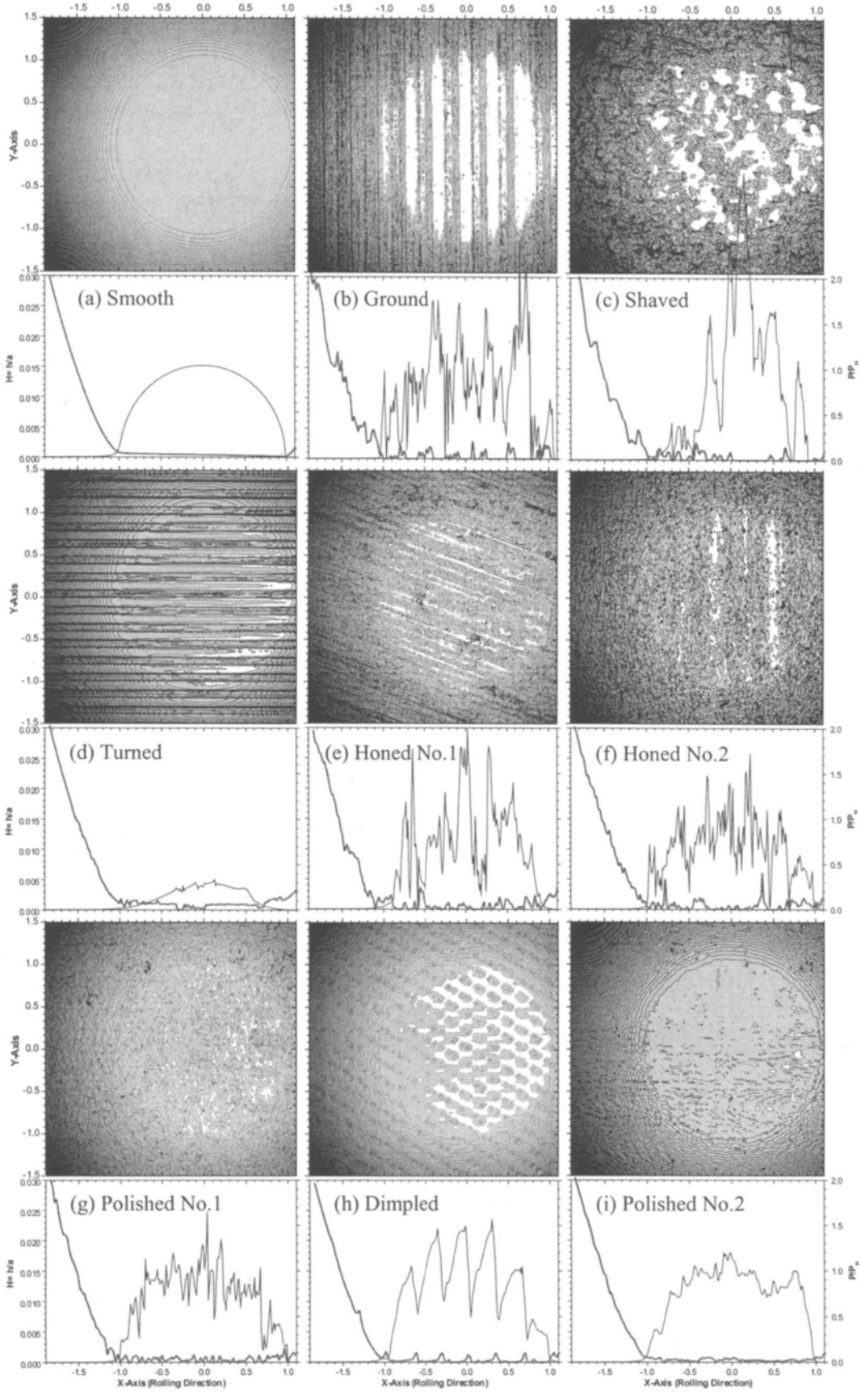


Figure 3. Film Thickness & Pressure Distributions

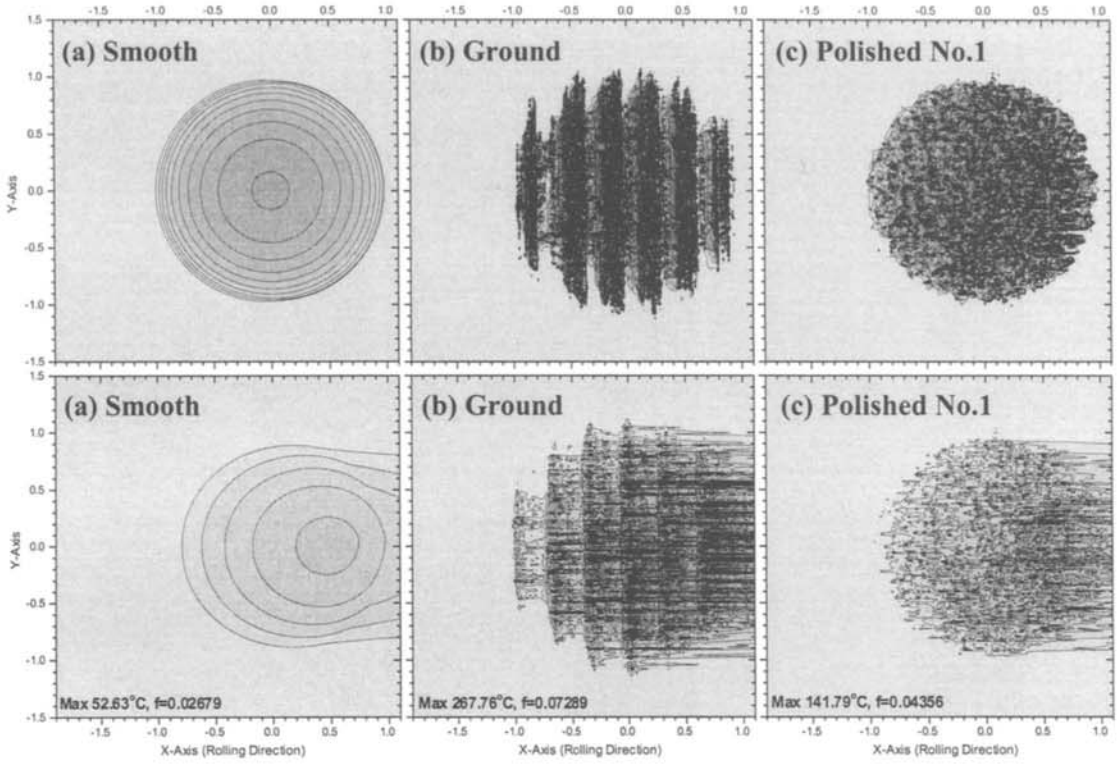


Figure 4. Sample Results for Pressure & Flash Temperature Distributions

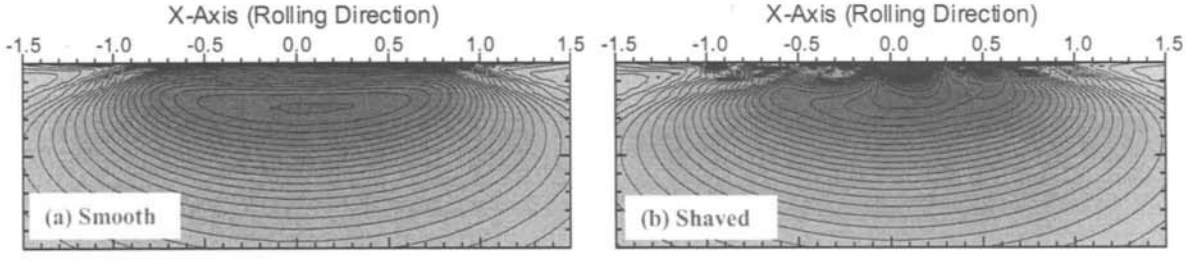


Figure 5. Sample Results for Max. Subsurface Shear Stress Distributions

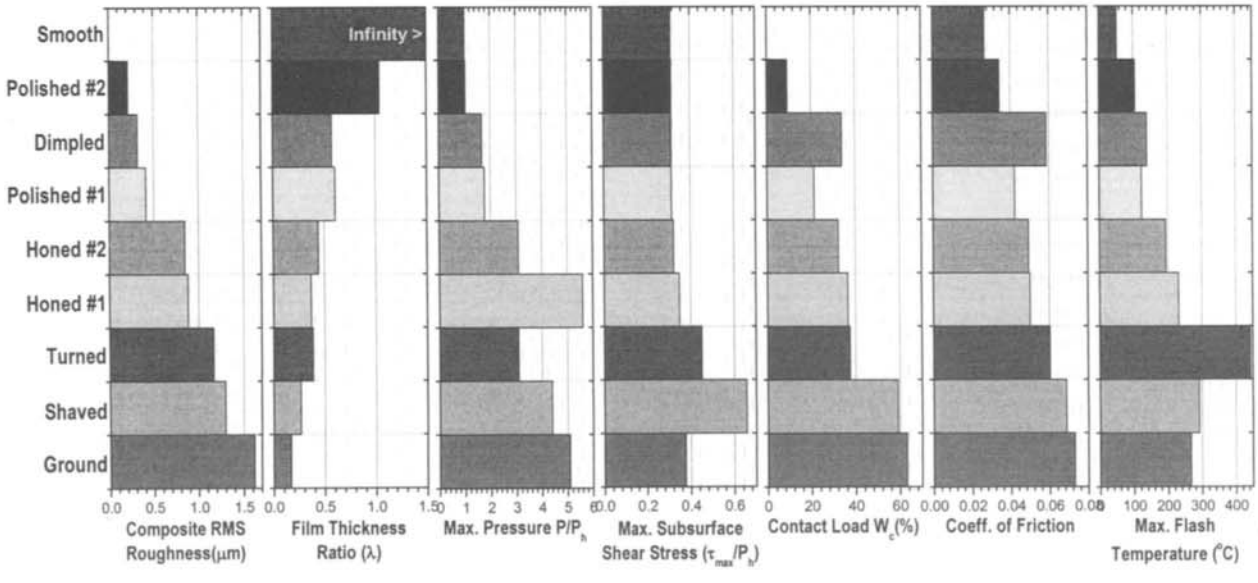


Figure 6. Summary of Results

[17]), and its coefficient is found to be 0.0268, which is quite low, as it is caused by hydrodynamic action only. The maximum subsurface shear stress obtained is $0.312*P_h$, very close to that from the classic Hertzian theory. The calculation of flash temperature is based on the theory of a moving heat source over a semi-infinite solid, and its more detailed description can be found in [17]. The maximum temperature increase on Surface 1 (which moves slower than Surface 2) based on 25% of sliding has been estimated as 52.6°C, and the temperature distribution is plotted in Figure 4(a).

Note that all the results mentioned above for the smooth surface EHL problem are steady state. They will be compared with eight rough surface solutions given below.

4.2. Eight Rough Surface Solutions

Detailed film thickness distributions and contact patterns as well as the solutions of pressure and film thickness along the center line (the x-axis) are shown in Figure 3. Note that in Figure 3(d) for the turned surfaces the center line happens to be located right in a valley between two longitudinal ridges. As a result, along the center line the pressure appears to be low and there is almost no contact (Actually at the ridges the pressure could be high and the contact could be quite severe). It is also important to point out that all the solutions for the rough surface EHL problems are no longer steady state, and the calculated results could change from one moment to another. After a sufficiently large number of time steps, however, the solutions are globally stabilized, and those transient changes usually do not affect the global characteristics that are well represented by the illustrated results.

Results of detailed pressure and flash temperature distributions from two sample cases can be found in Figure 4, in comparison with those from the smooth surface solution. Subsurface stresses can also be calculated with the present computer program, and, as an example, Figure 5 illustrates the maximum shear stress distribution for the shaved surfaces, compared with that for the smooth surfaces. It has been observed from these cases that the pressure, film thickness and flash temperature distributions obtained are directly correlated to the original surface micro-geometry. The surface roughness and topography could significantly affect the contact/EHL characteristics.

4.3. Summary and Comparison of the Results

The results for all the analyzed cases, one for the smooth surfaces and eight for the rough surfaces, are summarized in Figure 6. Note that those cases are listed from the top to the bottom in the order of increasing RMS roughness. The first one on the top, for example, is the case for the smooth surfaces, while the last one on the bottom is for Ground with the largest composite RMS roughness, 1.6193 μm (see the first bar chart on the left). The calculated film thickness ratios, λ, defined as the average film thickness divided by the composite RMS roughness, are plotted in the second bar chart from the left. Although the results of film thickness are not explicitly given, it can be readily figured out that the different machining/polishing processes have only a relatively limited influence on the average film thickness. In most cases h_a/a is in the range from 0.00037 to 0.00057, except for the longitudinally turned surfaces, 0.00075, and for the dimpled, 0.00031. Since the film thickness is relatively constant, the λ ratio, that is commonly used to describe the lubrication status, is affected directly by the surface characteristics. For the ground surfaces it can be as low as 0.167 while for the process of polished #2 it becomes 1.038 (Note that the λ ratio is infinite for the smooth surfaces).

The third and the fourth graphs from the left show the results of the maximum pressure peak height and the maximum subsurface shear stress, which are directly related to the possible surface failures such as pitting due to fatigue. It is observed that generally the greater the roughness, the higher the pressure/stress would be. However, there might be some exceptions, e.g. Honed #1 has caused the highest pressure peak probably due to significant secondary roughness. The polished surfaces with reduced RMS roughness, on the other hand, are likely to provide a positive impact on the performance and life of various mechanical components.

The third bar chart from the right in Figure 6 is for the contact load ratio, W_c, which is defined as the ratio of load supported by the surface contact to the total load. This parameter describes the load sharing between the hydrodynamic film and the asperity contacts. It can be seen that when the surfaces are perfectly smooth, the two surfaces are completely separated, and the contact load is zero. This is in the full film EHL regime. For Polished #2, the contact load could be as low as 9.3%, while for Ground as high as 63.2% (which can be considered nearly boundary lubrication). The lubrication status appears to be significantly influenced by the surface roughness and topography.

The two graphs on the right are for the results of friction and flash temperature increase. They are also greatly affected by the surface characteristics, and, generally speaking, the smoother the surfaces, the lower the friction and the temperature. Under the same operating conditions, the friction coefficient could be as high as 0.073 (Ground), or as low as 0.034 (Polished #2). The flash temperature increase, in most case, appears to be closely correlated with the friction, which is a direct source of the heat generation. However, the temperature increase is also related to the heat transfer. The dimpled surfaces seem to be able to provide high friction while still keeping the flash temperature relatively low probably due to enhanced convection. This type of surface might be good for some special applications such as traction drives. The longitudinally turned surfaces have caused the highest flash temperature, due to the accumulated frictional heat in the direction of motion that is coincident with the orientation of heavily loaded ridges. The calculated maximum temperature rise greater than 400°C may be an interpretation of possible lubrication breakdown and scuffing failure.

It is important to note that Polished #1 and Polished #2 have given the results that are reasonably close. However, the cost of Polished #1 could be only a fraction of that for Polished #2. Therefore, the first one could be better preferred in engineering practice when the production cost is a concern.

5. CONCLUSIONS

1. A deterministic approach is developed that is capable of analyzing practical cases based on measured 3-D roughness under severe operating conditions. The simulation could cover the entire transition from full film to mixed and even boundary lubrication.
2. The obtained results show that different surface roughness/finishing processes have a relatively limited influence on the average film thickness, but greatly affect the λ ratio, the pressure peak height, the subsurface stress, the load shearing, the friction and the flash temperature.
3. Generally surfaces with smaller roughness may provide better performance and probably longer life as well. However, there might be some exceptions, especially for cases with special requirements.
4. The computer program developed appears to be a useful tool for design optimization and finishing process selection.

NOMENCLATURE

a = radius of Hertzian contact zone
h = local film thickness (or gap)
H = h/a, dimensionless film thickness
h_a = average film thickness in the center of contact
p = pressure
p_h = maximum Hertzian pressure
R_q = RMS roughness in contact zone after deformation
U = rolling velocity
W_c = contact load ratio (load supported by asperity contacts divided by total load)
λ = h_a/σ, film thickness ratio
σ = composite RMS roughness before deformation

REFERENCES

[1] Dowson, D., and Higginson, G. R., Elastohydro-dynamic Lubrication, Pergamon Press (1966).
[2] Hamrock, B.J., and Dowson, D., Isothermal Elastohydrodynamic Lubrication of Point Contacts, Part I — Theoretical Formulation, Part II — Ellipticity Parameter Results, Part III

— Fully Flooded Results, Part IV — Starvation Results, *ASME J. of Lubrication Technology*, **98** and **99** (1976-77).

[3] Johnson, K.L., Contact Mechanics, Cambridge University Press, Cambridge, UK (1985).

[4] Greenwood, J.A., and Tripp, J.H., 1970-1971, The Contact of Two Nominally Flat Rough Surfaces, *Proc. Instn. Mech. Engrs.*, **185**, Part 1 (1970-71), 625-633.

[5] Christensen, H., 1969-70, "Stochastic Models for Hydrodynamic Lubrication of Rough Surfaces," *Proc. Instn. Mech. Engrs.*, Vol. 184, Part 1, No. 55, pp.1013-1026.

[6] Patir, N., and Cheng, H.S., 1978, "An Average Flow Model for Determining Effects of Three-Dimensional Roughness on Partial Hydrodynamic Lubrication," *ASME Journal of Lubrication Technology*, Vol.100, pp.12-17.

[7] Zhu, D., and Cheng, H.S., 1988, "Effect of Surface Roughness on the Point Contact EHL," *ASME Journal of Tribology*, Vol.110, pp.32-37.

[8] Lubrecht, A. A., ten Napel, W. E., and Bosma, R., 1988, "The Influence of Longitudinal and Transverse Roughness on the Elastohydrodynamic Lubrication of Circular Contacts," *ASME Journal of Tribology*, Vol. 110, pp. 421-426.

[9] Kweh, C. C., Evans, H. P., and Sindle, R. W., 1989, "Micro-Elastohydrodynamic Lubrication of An Elliptical Contact with Transverse and 3-Dimensional Roughness," *ASME Journal of Tribology*, Vol. 111, pp.577-584.

[10] Ai, X., Cheng, H.S., and Zheng, L., 1993, "A Transient Model for Micro-Elastohydrodynamic Lubrication with Three Dimensional Irregularities," *ASME Journal of Tribology*, Vol.115, pp.102-110.

[11] Venner, C. H., and Lubrecht, A.A., 1996, "Numerical Analysis of the Influence of Waviness on the Film Thickness of a Circular EHL Contact," *ASME Journal of Tribology*, Vol.118, pp.153-161.

[12] Xu, G., and Sadeghi, F., 1996, "Thermal EHL Analysis of Circular Contacts With Measured Surface Roughness," *ASME Journal of Tribology*, Vol.118, pp.473-483.

[13] Zhu, D., and Ai, X., 1997, "Point Contact EHL Based on Optically Measured Three-Dimensional Rough Surfaces," *ASME Journal of Tribology*, Vol.119, pp.375-384.

[14] Jiang, X., Hua, D.Y., Cheng, H.S., Ai, X., and Lee, S.C., 1998, "A Mixed Elastohydrodynamic Lubrication Model With Asperity Contact," *ASME Journal of Tribology*, Vol.121, No.3, July.

[15] Zhu, D., and Hu, Y.Z., "The Study of Transition from Full Film Elastohydrodynamic to Mixed and Boundary Lubrication," *Proceedings of 1999 STLE/ASME H.S. Cheng Tribology Surveillance*, pp.150-156 (1999).

[16] Hu, D., and Zhu, D., "A Full Numerical Solution to the Mixed Lubrication in Point Contacts," *ASME Journal of Tribology*, **122**, pp.1-9 (2000).

[17] Zhu, D., and Hu, Y.Z., "A Computer Program Package for the Prediction of EHL and Mixed Lubrication Characteristics, Friction, Subsurface Stresses and Flash Temperatures Based on Measured 3-D Surface Roughness," *Tribology Transactions*, **44**, pp.383-390 (2001).

[18] Zhu, D., and Hu, Y.Z., "Effects of Rough Surface Topography and Orientation on the Characteristics of EHD and Mixed Lubrication in Both Circular and Elliptical Contacts," *Tribology Transactions*, **44**, pp.391-398 (2001).

[19] Zhu, D., "Effect of Surface Roughness on Mixed EHD Lubrication Characteristics," presented at the 2nd World Tribology Congress, Vienna, Austria, September 3-7, 2001, and to appear in *Tribology Transactions*.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Starved Lubrication of Isoviscous Rigid Circular Contacts

M.-P. Noutary[a], A.A. Lubrecht[b]

[a]Centre de Mathématiques, INSA de Lyon, France.

[b]Laboratoire de Mécanique des Contacts, UMR CNRS 5514, INSA de Lyon, France.

A very large proportion of modern bearings is grease lubricated for life. Grease lubrication results for the majority of bearing operating conditions in a starved lubricant supply. This starved lubrication limits the film thickness generated and as a result provokes a reduction of the bearing life. Consequently, the topic of starved lubrication of elliptical EHL contacts is of more than academic interest. The numerical solution of starved elliptical EHL contacts is marked by an additional parameter (ellipticity) and additional convergence difficulties. Hence, the solution of the starved elliptical problem under simpler conditions is of interest. The current study extends the analysis of Kapitza to starved contacts and expresses the results in terms of the starvation parameter γ. The analytical results are compared with full numerical results presented in the literature.

1. Introduction

Over the last decades the performance of rolling element bearings has increased tremendously due to improvements in steel processing, manufacturing techniques and lubrication. However, it is also true that these performance improvements have been partially obtained by a reduction of the safety factors. As a very large proportion of these bearings are grease lubricated (for life), the lubricating conditions are far from fully flooded, and most likely partially or fully starved. The operation in the starved regime implies a reduction of the lubricant film thickness separating the rolling elements from the inner and outer ring. Such a reduction in film thickness translates directly in a reduction of operating life. As the safety margins on modern bearings have decreased, it is important to accurately predict the film thickness under these starved conditions.

The film thickness of fully flooded smooth EHL contacts problem was calculated numerically by Dowson et al. [1,2]. It was the early work of Wedeven et al. [3], Chiu [4] and Pemberton et al. [5] that defined the starved lubrication problem. They studied starvation visually using optical interferometry, and defined it in terms of the meniscus position.

The analytical work of [6] established a relation between the meniscus position and the film thickness reduction for line contacts. In numerical studies of the point contact problem [7] the inlet meniscus position was again adopted to describe starvation.

Modeling the degree of starvation using the inlet meniscus position is not very physical, as the inlet meniscus position is a dependent variable, the independent variable being the amount of oil present on the surfaces in the inlet.

Consequently, more recent work by Chevalier et al. [8,9] uses the physical parameter r, the ratio of the inlet oil thickness and the fully flooded film thickness. This work showed that the film thickness under starved lubrication conditions could be predicted using a single parameter γ. The value of this parameter ranged from 2 to 5 for typical EHL conditions. Recently, Damiens et al. [10] showed that the variation of the γ parameter was governed by the inlet length parameter $\sqrt{M/L}$ and the ellipticity $e = R_y/R_x$. As the influence of the ellipticity is a non-trivial problem, it was felt that a thorough analysis of the starved Rigid Isoviscous problem was of interest. That analysis should then be extended to elliptical contacts to understand the elliptical EHD problem. Consequently, this paper extends the classical Kapitza [11] analysis to starved conditions and the analysis of elliptical contacts is programmed.

2. Nomenclature

a	inlet meniscus position, [m]
A	dimensionless inlet meniscus position $A = a/\sqrt{h_0 R_x}$
e	ellipticity $e = R_y/R_x$
h	film thickness, [m]
H	dimensionless film thickness, $H = h/h_0$
h_0	minimum film thickness, [m]
h_{0ff}	fully flooded min. film thickness, [m]
h_{oil}	oil thickness in inlet, [m]
p	pressure, [Pa]
p_0	reference pressure [Pa] $p_0 = 12\eta_0 u_m \sqrt{R_x}/(h_0\sqrt{h_0})$
P	dimensionless pressure, $P = p/p_0$
r	relative oil inlet thickness $r = h_{oil}/h_{0ff}$
$\mathcal{R}$	relative oil film thickness $\mathcal{R} = h_0/h_{0ff}$
R_x	reduced radius of curvature in x, [m]
R_y	reduced radius of curvature in y, [m]
u_m	mean velocity in x, $u_m = (u_1 + u_2)/2$, [m/s]
w	load, [N]
X, Y	dimensionless coordinates, $X = x/\sqrt{h_0 R_x}$, $Y = y/\sqrt{h_0 R_x}$
γ	starvation parameter
η_0	viscosity at ambient pressure, [Pa s]
ρ_0	density at ambient pressure, [kg/m^3]
θ	filling rate or relative oil quantity, [−]

3. Fully Flooded Circular Case

For the stationary isoviscous incompressible rigid lubrication case the Reynolds equation reads

$$\frac{\partial}{\partial x}(h^3\frac{\partial p}{\partial x}) + \frac{\partial}{\partial y}(h^3\frac{\partial p}{\partial y}) - 12\eta_0 u_m \frac{\partial h}{\partial x} = 0 \qquad (1)$$

The film thickness h, for the circular case, is given by:

$$h(x,y) = h_0 + \frac{x^2}{2R_x} + \frac{y^2}{2R_x} \qquad (2)$$

The following dimensionless variables are introduced $X = x/\sqrt{h_0 R_x}$, $Y = y/\sqrt{h_0 R_x}$, $P = p/(12\eta_0 u_m \sqrt{R_x}/(h_0\sqrt{h_0}))$, $H = h/h_0$.
The dimensionless Reynolds equation and film thickness equation become:

$$\frac{\partial}{\partial X}(H^3\frac{\partial P}{\partial X}) + \frac{\partial}{\partial Y}(H^3\frac{\partial P}{\partial Y}) - \frac{\partial H}{\partial X} = 0 \qquad (3)$$

$$H(X,Y) = 1 + \frac{X^2}{2} + \frac{Y^2}{2} \qquad (4)$$

The Reynolds equation can be rewritten to make the Laplace operator $\Delta = \partial^2/\partial X^2 + \partial^2/\partial Y^2$ appear:

$$H^3\Delta P + 3H^2(X\frac{\partial P}{\partial X} + Y\frac{\partial P}{\partial Y}) - X = 0 \qquad (5)$$

Using polar coordinates $X = r\cos(\theta)$, $Y = r\sin(\theta)$ and introducing $G(r,\theta) = P(X,Y)$ one obtains:

$$H^3\Delta G + 3rH^2\frac{\partial G}{\partial r} = r\cos(\theta) \qquad (6)$$

with

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \qquad (7)$$

and

$$H(r) = 1 + \frac{r^2}{2} \qquad (8)$$

One looks for a solution of the form $G(r,\theta) = g(r)\cos(\theta)$, where $g(r)$ satisfies:

$$g''(r) + [\frac{1}{r} + \frac{3r}{H(r)}]g'(r) - \frac{1}{r^2}g(r) = \frac{r}{H^3(r)} \qquad (9)$$

This equation has a particular solution

$$g_\infty(r) = -\frac{r}{5H^2(r)} \qquad (10)$$

which is plotted in Figure 1. This is the solution obtained by Kapitza [11]. It is named g_∞ because of the fully flooded boundary condition $\lim_{r\to\infty} g_\infty(r) = 0$.
The fully flooded dimensionless pressure distribution reads:

$$P_\infty(X,Y) = -\frac{X}{5H^2(X,Y)} \qquad (11)$$

and $P_\infty(X,Y)$ is plotted in figure 2.
The dimensionless fully flooded load carrying capacity is defined as

$$W_\infty = w\sqrt{h_{0ff}}/(12\eta_0 u_m R_x \sqrt{R_x}) \qquad (12)$$

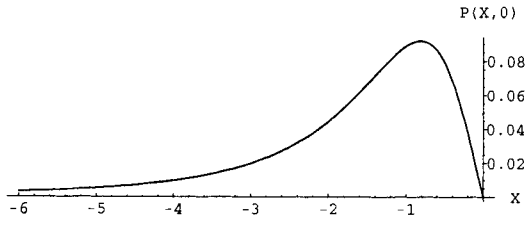


Figure 1. *Fully flooded pressure profile* $P_\infty(X, Y = 0)$.

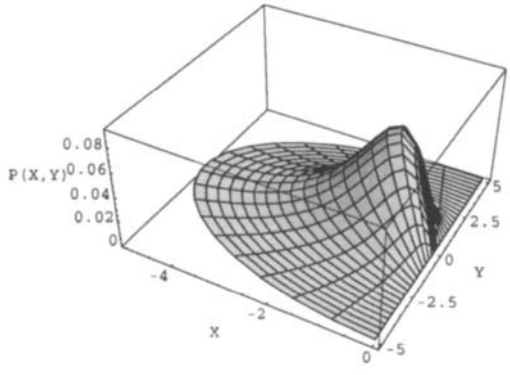


Figure 2. *Fully flooded pressure distribution* $P_\infty(X, Y)$.

$$W_\infty = \int_{-\infty}^{+\infty} [\int_{-\infty}^{0} P_\infty(X, Y)\, dX\,] dY \tag{13}$$

where the integration takes place over the zone of positive pressure ($P > 0$, $X < 0$).
This integral gives

$$W_\infty = \frac{\pi}{5}\sqrt{2} \approx 0.888\ldots \tag{14}$$

From this result the fully flooded minimum film thickness h_{0ff} can be obtained as:

$$h_{0ff} = \frac{288\pi^2}{25} \frac{\eta_0^2 u_m^2 R_x^3}{w^2} \tag{15}$$

4. Starved Circular Results

The solution of the homogeneous equation

$$g''(r) + (\frac{1}{r} + \frac{3r}{H(r)})g'(r) - \frac{1}{r^2} g(r) = 0 \tag{16}$$

can be found using Mathematica [12] and is given in the appendix. The homogeneous solution can be added to the particular solution to obtain the required starved boundary condition $g(r = A) = 0$ ($A = a/\sqrt{h_0 R_x} > 0$). The circular shape of the inlet meniscus seems closer to the physical reality than the straight line proposed in [7]. The function $P(X, 0)$ is plotted in figure 3, with $P(A = 4) = 0$, i.e. $P(-4, 0) = 0$. Please note that the pressure for $X < -4$ is negative, and that this part of the solution is not physical. The same is true for the solution for $X > 0$. In both zones the parameter θ, the filling rate, has to be introduced in the Reynolds equation, as suggested by Elrod [13,14].
The starved load carrying capacity W is a function of the inlet meniscus position A and given by the integral

$$W(A) = \iint_{\mathcal{A}} P(X, Y)\, dX\, dY \tag{17}$$

where the domain $\mathcal{A}$ is given by $X^2 + Y^2 \leq A^2$, $X \leq 0$. A graph of $W(A)$ as a function of A is given in figure 4. Note that the value of $W(A)$ reaches the fully flooded asymptote $\pi\sqrt{2}/5 \approx 0.888$ for large values of A.
The definition (12) can be extended to define W under starved lubrication conditions. Using this

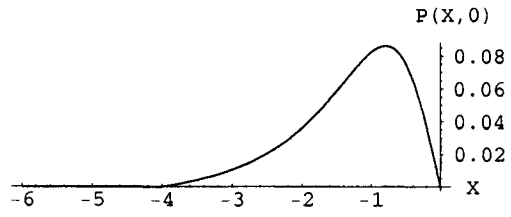


Figure 3. *Starved pressure profile* $P(X, Y = 0)$, *for* $A = 4$.

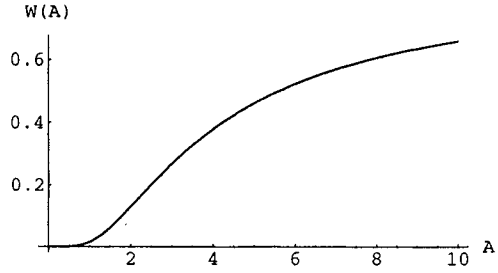


Figure 4. *Dimensionless load carrying capacity* W *as a function of the meniscus position* A.

extension, (12) and (14), the minimum film thickness ratio, $\mathcal{R} = h_0/h_{0f}$, can be defined as:

$$\mathcal{R}(A) = \frac{25}{2\pi^2} W^2(A) \tag{18}$$

As was argued in previous work, the inlet meniscus position A is of limited interest, as it is determined by the thickness of the oil layer in the inlet h_{oil}. In order to link h_{oil} with the inlet meniscus position a the flow continuity equation is analysed on the negative x-axis ($\theta = 0$).

$$h_{oil}u_m = h(-a,0)u_m - \frac{h^3(-a,0)}{12\eta_0}\frac{\partial p}{\partial x}(-a,0) \tag{19}$$

On the x-axis side flow is neglected as $\partial p/\partial y \approx 0$. Introducing the dimensionless oil film thickness $H_{oil} = h_{oil}/h_0$ one obtains

$$H_{oil} = H(-A,0) - H^3(-A,0)\frac{\partial P}{\partial X}(-A,0) \tag{20}$$

For a given dimensionless inlet meniscus position A, the pressure distribution and the pressure derivative are known, thus H_{oil} can be determined.
Hence it is possible to express the relative film thickness $\mathcal{R} = h_0/h_{0f}$ as a function of the relative inlet oil film thickness $r = h_{oil}/h_{0f}$, by combining equations (17) and (19), see figure 5.

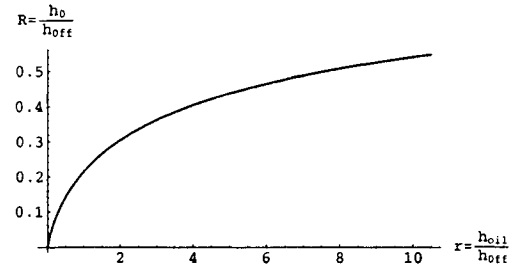


Figure 5. *Dimensionless film thickness* $\mathcal{R} = h_0/h_{0f}$ *as a function of the dimensionless oil inlet thickness* $r = h_{oil}/h_{0f}$.

In [9] this relation was approximated by

$$\mathcal{R}_1 = \frac{r}{\sqrt[\gamma]{1+r^\gamma}} \tag{21}$$

and for the isoviscous rigid regime a value of γ slightly smaller than 0.5 was found. From this analytical solution the best fit for the isoviscous rigid conditions is obtained with $\gamma = 0.47$. However, a function that describes $\mathcal{R}(r)$ more accurately is given by:

$$\mathcal{R}_2 = \frac{r}{1 + r^{1/5} + 1.8r^{3/5} - 0.2r^{4/5} + r} \tag{22}$$

Figures 6 and 7 show the comparison of the analytical solution and the two approximations. It is clear from figure 6 that both approximations reflect the overall trend of the analytical solution. From figure 7 it is obvious that the more complex equation (22) gives the best fit. However, it should be remembered that the parameter γ is a variable in the case of starved EHL contacts,

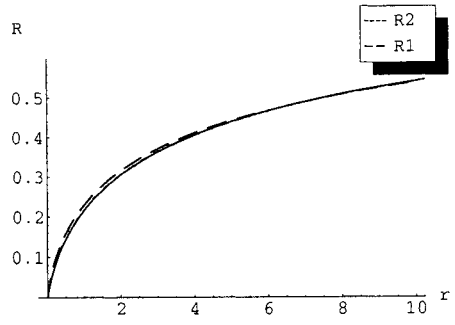


Figure 6. *Comparison of the two approximations* $\mathcal{R}_1(\gamma = 0.47)$, $\mathcal{R}_2$ *and the analytical solution.*

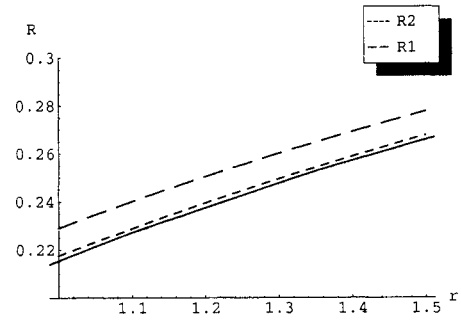


Figure 7. *Zoom of figure 6* ($\gamma = 0.47$).

depending on the operating conditions. It would be very difficult to determine the variation of all coefficients and powers (6) as a function of the operating conditions. Hence, even though the precision of equation (22) exceeds that of equation (21), the additional complexity of (22) makes it less suitable for an application to starved EHL contacts.

5. Conclusion

In this paper the starved lubrication of the isoviscous rigid circular contact was treated analytically. The dimensionless load carrying capacity was determined as a function of the dimensionless inlet meniscus parameter A. As the pressure distribution is known, the inlet meniscus position can be related to the thickness of the oil layer in the inlet. The combination of the two equations gives the relative oil film thickness $\mathcal{R}$ as a function of the relative inlet oil layer r. The analytical study results in a relation that is close to the relation obtained numerically, with $\gamma \simeq 0.5$ for the isoviscous rigid case. A more accurate equation describing the starved lubrication of the isoviscous rigid contact was found (22), however, its increased precision comes at the cost of an increased number of parameters.

REFERENCES

1. **Dowson, D., and Higginson, G.R.**, 1959, "A Numerical Solution to the ElastoHydrodynamic Problem," *J. of Mech. Eng. Science*, **1**, 1, pp. 6-15.
2. **Hamrock, B.J., and Dowson D.**, 1976, "Isothermal Elastohydrodynamic Lubrication of Point Contacts, part I, Theoretical Formulation," *ASME J. Lub. Tech.*, **98**, pp. 223-229.
3. **Wedeven, L.D., Evans, D., and Cameron, A.**, 1971, "Optical Analysis of Ball Bearing Starvation," *Trans. ASME, Series F*, **93**, pp. 349-363.
4. **Chiu, Y.P.**, 1974, "An Analysis and Prediction of Lubricant Film Starvation in Rolling Contact Systems," *ASLE Trans.*, **17**, 1, pp. 22-35.
5. **Pemberton, J., Evans, D., and Cameron, A.**, 1976, "A Mechanism of Fluid Replenishment in ElastoHydrodynamic Contacts", *WEAR*, **37**, pp. 184-190.
6. **Wolveridge, P.E., Baglin, K.P., Archard, J.F.**, 1970-1971, "The Starved Lubrication of Cylinders in Line Contact," *Proc. Inst. Mech. Engrs.*, **181**, pp. 1159-1169.
7. **Hamrock, B.J., and Dowson D.**, 1977, "Isothermal Elastohydrodynamic Lubrication of Point Contacts, part IV, Starvation Results," *ASME J. Lub. Tech.*, **99**, pp. 15-23.
8. **Chevalier, F., Lubrecht, A.A., Cann, P.M.E., Colin, F., Dalmaz, G.**, 1995, "Starvation Phenomena in EHL Point Contacts: Influence of the Inlet Flow Distribution", *Proceedings of the 22st Leeds-Lyon*

Symposium on Tribology.

9. **Chevalier, F., Lubrecht, A.A., Cann, P.M.E., Colin, F., Dalmaz, G.**, 1998, "Film Thickness in Starved EHL Point Contacts", *ASME JoT*, **120**, pp. 126-133.
10. **Damiens, B., Venner, C.H., Cann, P.M.E., and Lubrecht, A.A.**, 2002, "Starved Lubrication of Elliptical EHD Contacts," *paper submitted to the ASME JoT.*
11. **Kapitza, P.L.**, 1955, "Hydrodynamic Theory of Lubrication during Rolling", Zh. Tek. Fiz., **25**, pp. 747-762.
12. **Wolfram, S.**, 1999, "The MATHEMATICA Book", Cambridge University Press, *ISBN 0-521-64314-7.*
13. **Elrod, H.G., and Adams, M.L.**, 1974, "A Computer Program for Cavitation and Starvation Problems," *Proceedings of the 1st Leeds-Lyon Symposium on Tribology*, pp. 37-41.
14. **Elrod, H.G.**, 1981, "A Cavitation Algorithm," *ASME J. Lub. Tech.*, **103**, pp. 350-354.

Appendix

The solution of the homogeneous differential equation (15) is given by

$$\begin{aligned}\frac{g(r)}{r} &= {}_2F_1(2-\sqrt{\frac{5}{2}}, 2+\sqrt{\frac{5}{2}}; 2; -\frac{r^2}{2}) \\ &= \frac{\Gamma(2)}{\Gamma(2-\sqrt{5/2})\Gamma(2+\sqrt{5/2})}\int_0^1 t^{1-\sqrt{5/2}} \\ &\qquad (1-t)^{\sqrt{5/2}-1}(1+t\frac{r^2}{2})^{-2-\sqrt{5/2}}\,dt \\ &= \sum_{k=0}^{\infty}\frac{(2-\sqrt{5/2})_k(2+\sqrt{5/2})_k}{(2)_k}\frac{(-r^2/2)^k}{k!}\end{aligned}$$

where ${}_2F_1()$ is the hypergeometric function and $(a)_k$ is defined as

$$(a)_k = a(a+1)(a+2)\cdots(a+k-1)$$

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Gaseous Evolution of Hydrogen from Hydrocarbon Oil and Grease Lubricated Contacts

C.Newlands[a], A.Olver [b] and N.Brandon[c]

[a&b] Tribology Section, Mechanical Engineering Department, Imperial College of Science, Technology and Medicine, Exhibition Road, London SW7 2BX

[c] Electrochemical Engineering Group, Chemical Engineering, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2AZ

The level of hydrogen surrounding lubricated rubbing contacts has been investigated using mass spectral analysis. Two mineral oils of differing viscosity grades (SAE 10 and 40) and three grease base stocks have been investigated. The tests were carried out using a modified rolling four-ball tester operating under extreme pressure conditions. The effect of test temperature was investigated by introducing heating. The resulting levels of hydrogen have been compared with the test parameters and the factors important for gaseous evolution of hydrogen are discussed.

1. INTRODUCTION

It has been known since the early sixties that hydrogen can enter steel while undergoing wet lubricated rubbing in a 4-ball fatigue tester [1]. The hydrogen embrittles the steel and causes premature failure by speeding rolling contact fatigue. These studies relied on out of contact detection or extraction of hydrogen by driving the hydrogen off at a very high temperature (+800 °C). The use of tritiated water and subsequent measurement of the radioactivity of the used driving ball showed that the bulk of hydrogen came from the hydrocarbon lubricant. Atomic hydrogen can enter the steel and diffuses through interstitial holes. The breakthrough time (t_b) estimated by equation 1.1 [5] gives an indication of the speed of diffusion and some examples are given in table 1.

$$t_b \approx L^2 \div \left(15.3 \times D_{eff}\right) \qquad (1.1)$$

The diffusion coefficient for steels [2-4] varies widely and this may be due to sample history influencing the entry and diffusion of hydrogen. Coefficients for pure iron at room temperature have been reported to vary between 3×10^{-9} and 5×10^{-4} $cm^2 s^{-1}$ [2].

Table 1
Examples of Diffusion in Iron Based Materials.

Material	(D_{eff}) Effective Diffusion Coefficient (cm^2s^{-1})	(t_b) Breakthrough Time/(min) for 2 mm (L) thick membrane.
Pure Iron	1.1×10^{-7} [2]	39.6
AISI 4140	5×10^{-7} [3]	87
AISI 52100	1.1×10^{-8} Estimated[4]	3960

The measurement of the hydrogen entering a contact is traditionally difficult with the fluxes and amounts of hydrogen involved being very small. Typically a rise of just over 1 part per million of hydrogen spread through out the ball is observed [4] during a typical high-pressure 4-ball test. This is very close to the operating sensitivity of the hot extractive detection equipment and therefore it is difficult to resolve the effects at work.

Table 2
Properties of the Greases used.

Grease Name	Thickener Type/ Formula / Proportion of Thickener (%)	Base oil type / Kinematic Viscosity at 40°C (cSt)
Di-Urea	Organic / R_1-NHCONH-R-NHCONH-R_2 [R = alkyl group] / 7%	Paraffinic / 200
Lithium-12- Hydroxystearate	Carboxylate-soap / $(CH_3(CH_2)CHOH(CH_2)_{10}COOLi$ / 7%	Paraffinic / 200
Lithium Complex	Mixture of Lithium-12-Hydroxystearate and dilithium azelate / 7%	Paraffinic / 200

Hydrogen causes embrittlement and damage in several ways [6]. Hirth reports that 4 basic processes are involved in hydrogen embrittlement and that they all play a role:

1) Hydrogen is transported internally by interstitial or dislocation movement. Hydrogen can lock dislocations and slow plastic flow within the material and therefore crack tips cannot be blunted so effectively. Hydrogen can also speed up dislocation movement and effectively be carried to the crack tip where it embrittles via adsorption.
2) The adsorption of hydrogen onto freshly created surfaces at the crack tip reduces the surface energy and therefore makes crack growth more thermodynamically favourable.
3) The hydrogen can recombine to form molecular hydrogen within voids and substantial pressures can build up leading to crack initiation.
4) The hydrogen can react with carbides at high temperature to create voids of insoluble methane gas (hydrogen attack). This leads to crack initiation by creating high-pressure voids and weakening the metal structure by depleting the amount of carbon in the lattice.

Currently there is uncertainty surrounding the roles of certain extreme pressure (EP) additives that can contain sulphur. Sulphur has been shown [7] in aqueous electrochemistry to enhance the entry of hydrogen by adsorbing onto the surface and affecting the kinetics involved in the recombination of hydrogen thus leading to a higher coverage of hydrogen. It has also been shown [8] that organic sulphur EP molecules similar to dimethyl disulphide produce hydrogen during the reaction in vacuo with pure iron. The reaction mechanism is proposed by the researchers to follow the following simplified route:

$$R\text{-}S\text{-}S\text{-}R \rightarrow 2\ R\text{-}S\text{-}H \rightarrow 2Fe\text{-}S\text{-}R + H_2 \qquad (1.2)$$

R is an alkyl or phenyl group. The disulphide molecule becomes a thiol and then ejects hydrogen at the surface to become a thiolate. This then leads to the formation of iron sulphide (the desired solid boundary lubricant). This reaction therefore should yield some hydrogen during the 4-ball tests.

The present paper reports a gaseous detection system that will go some way to determining the operating factors that lead to creation of hydrogen within a rubbing lubricated contact. This method combined with an in-situ method that is under development will yield information on the partitioning of the hydrogen between the steel and the surrounding environment. This may lead to a greater understanding of the physical processes at work within the contact.

2. METHOD

The amount of hydrogen evolved by two mineral oils, SAE 10 (3.02e-2 Pa.s 35.5 cSt @ 40°C), SAE 40 (1.47e-1 Pa.s, 173.0 cSt @ 40°C) and three grease base stocks (table 2) are investigated by running within a rolling 4-ball test rig. The lubricants were tested under 4 conditions:

1) As received.
2) With 5% distilled water added.
3) With 2% of an EP sulphur additive (a di-mercapto-thiadiazole derivative).
4) With both 5% water and 2% EP additive.

The work was carried out using an AOL test machine, which was converted for use as a 4-ball tester. Using diagrams obtained from IP test method 300 [9] a race holder and ball collet were manufactured. The race holder was adapted to allow the fitting of tubes that allowed the sampling of gas from within the race holder. The sampling set-up is shown in figure 1.

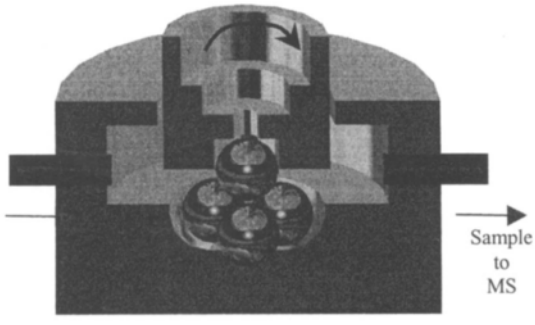


Figure 1 Diagram of Gas Sampling Set-up in the Modified 4 Ball Race Holder.

12.7 mm diameter grade 10 AISI 52100 steel balls were used throughout. The tests were conducted in a similar manner to the IP test method [9]. The exceptions are listed here: The 1.1 kW motor is single phase rotating at 50 Hz. A hydraulic cylinder applied the load and the temperature was recorded using a thermocouple inserted in the ball-race keyway. Temperature was recorded every five minutes along with any events in the run such as turning the motor on, loading up the sample or pitting being detected by audible rough running.

The test procedure consisted of 10 minutes to establish a background level of hydrogen, then the motor was started and run at light load 0.5 kN (P_0=2.6 Gpa) for 5 minutes. This was then followed by the load being increased to 5.3 kN (P_0=5.6 Gpa) for the remaining duration of the test. P_0 is the hertzian average contact pressure. The sampled gas was collected via a fine, trace-heated capillary into a quadrupole mass spectrometer at a rate of 20 cm^3 min^{-1}. The quadrupole mass spectrometer was set up to perform measurements on selected masses.

All test equipment and samples were ultrasonically cleaned in toluene and then in acetone, prior to being loaded into the rig. The grease

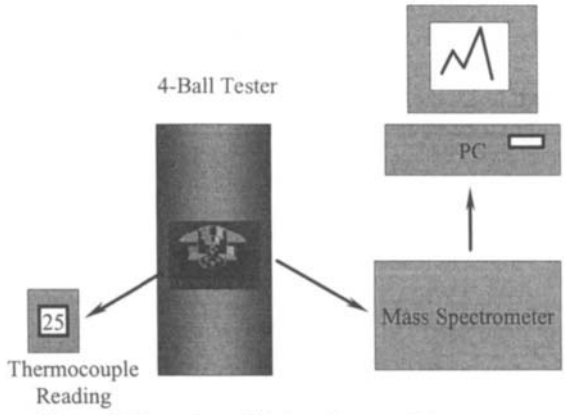


Figure 2 Data Acquisition System Diagram.

lubricants were added directly to the race way by means of a syringe. 0.2 cm^3 of grease was used in each test. A 5 cm^3 volume of oil lubricant was added directly to the race holder. When more than this amount was used it tended to be flung from the contact and could block the sample tube.

3. RESULTS

The results of the 4 ball tests are presented here in chart form and show how the concentration of hydrogen develops with time. Events of each test are included in table form. Each chart contains 4 plots showing the effects of different additives or the effect of imposed heating.

3.1. Mineral Oils

Tests on mineral oils are split between ambient tests concentrating on the action of additives, and temperature controlled tests.

3.1.1 Ambient Tests

Two mineral oils having the viscosity grades, SAE 10 and 40 were tested under four conditions involving the presence of water or the extreme pressure additive. The mass spectrometry results for each fluid are shown in figures 3 and 4.

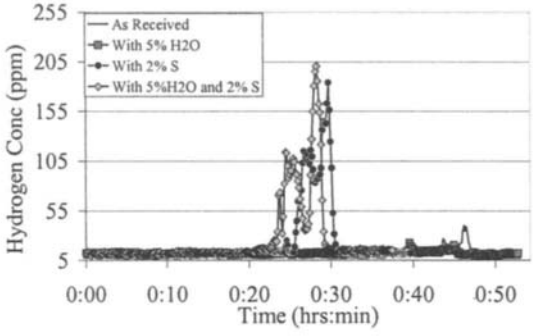


Figure 3 SAE 10 Mineral Oil, Hydrogen Concentration Surrounding the 4-ball Test Head.

Test conditions such as maximum raceway temperature and fatigue severity are recorded in tables 3 and 4.

Table 3
Test Events Corresponding to the SAE 10 4-ball tests

Additive	Test start (min)	Pitting Start (min)	Motor Stop (min)	Max T (°C)/ Pit Type
As Received	10	46	46	84 / A
5% H_2O	10	45	45	74 / D
2% EP	10	30	30	87 / B
5% H_2O +2% EP	10	28	28	87 / B

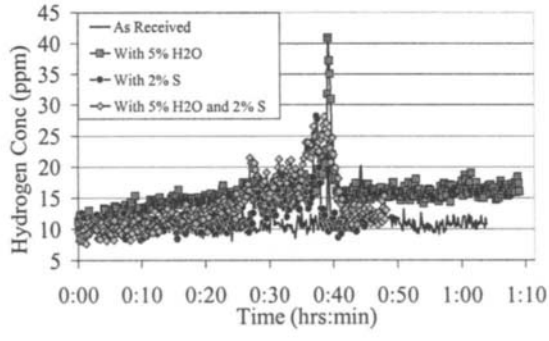


Figure 4 SAE 40 Mineral Oil, Hydrogen Concentration Surrounding the 4-ball Test Head.

The tests were carried out until an elevated level of audible vibration was observed. This coincided with an increase in the sampled concentration of hydrogen surrounding the test head. After disassembly of the rig it was found that in the majority of cases the vibration and sharp increase in hydrogen was associated with fatigue pitting. In some cases no pit was found.

Table 4

Test Events Corresponding to the SAE 40 4-ball tests

Additive	Test start (min)	Pitting Start (min)	Motor Stop (min)	Max T (°C)/ Pit Type
As Received	10	N/A	50	80/D
5% H_2O	10	39	39	85/A
2% EP	10	37	37	88/A
5% H_2O +2% EP	10	N/A	40	96/D

3.2. Heated Tests

In these tests the rig head was heated using a thermostatic controller. Care was taken to ensure minimal over shoot of the temperature compared to the set point and that the distribution of temperature was even throughout the sample holder. The test hydrogen concentration results are shown in figure 5. The motor was started at 50 minutes.

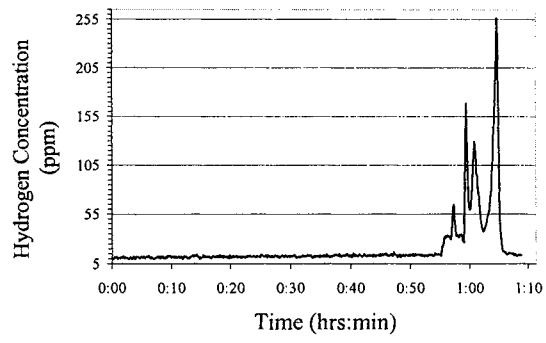


Figure 5 Heated Test Using SAE 10 in the As Received Condition

The test involved increasing the temperature under static conditions from 20 °C to 105 °C. This resulted in an average increase of just 3 ppm of hydrogen.

3.3. Grease Base Stocks (Ambient)

The grease base stocks have yielded broadly similar levels of hydrogen evolved during a test. Therefore only an example of typical behaviour is shown in figure 6 and the basic characteristics of the all grease tests are recorded in tables 5 to 7.

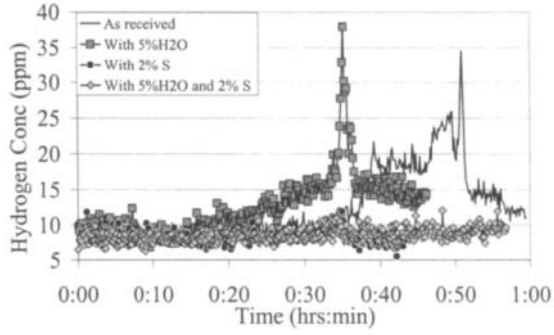


Figure 6 Typical Chart for Di-Urea Grease Showing Hydrogen Concentration Surrounding the 4-ball Test Head.

Table 5
Characteristic Events for Di-Urea Grease.

Di-Urea Grease	Pre-Pit Level of H_2 (ppm) / Time (m:s)	Maximum Concentration of H_2 (ppm) / Time (m:s)	Background Average Concentration H_2 (ppm) over first 10 min	Maximum Temperature (°C) /Pit type
As received	25 / 49:33	34 / 51:16	8.6	111 / C
+ H_2O	17 /33:30	37 / 35:00	9.3	85 / B
+ EP	11.9 / 35:14	11.9 / 35:14	8.7	83 / D
+ H_2O + EP	10.9 / 34:30	10.9 / 34:30	7.8	74 / D

Table 6
Characteristic Events for Lithium Hydroxy Stearate Grease.

Lithium Hydroxy Stearate Grease	Pre-Pit Level of H_2 (ppm) / Time (m:s)	Maximum Concentration of H_2 (ppm) / Time (m:s)	Background Average Concentration H_2 (ppm) over first 10 min	Maximum Temperature (°C) /Pit type
As received	23.1 / 38:03	29.6 / 42:00	10.8	81 / B
+ H_2O	25.4 / 32:24	33.4 / 40:47	9.9	81 / C
+ EP	13.3 / 25:01	15.9 / 30:16	9.0	60 / B
+ H_2O + EP	9.86 / 28:32	9.86 / 28:32	8.8	72 / D

Table 7
Characteristic Events for Lithium Complex Grease.

Lithium Complex Grease	Pre-Pit Level of H_2 (ppm) / Time (m:s)	Maximum Concentration of H_2 (ppm) / Time (m:s)	Background Average Concentration H_2 (ppm) over first 10 min	Maximum Temperature (°C) /Pit type
As received	17.5 / 34:57	26.2 / 35:56	9.4	79 / C
+ H_2O	18.5 / 17:12	22.5 / 34:14	10.6	83 / C
+ EP	15.6 / 53:34	22.8 / 56:50	8.0	85 / C
+ H_2O + EP	12 / 29:42	12 / 29:42	8.7	80 / D

3.4. Investigation of Fatigue Pitting Severity

The level of pitting/spalling, in the tests carried out, is characterised by four basic types:

A) Several small pits have spalled away and are contained within the wear scar. An example is shown in figure 7

B) The pit size is large, numerous and spread throughout the wear scar. An example is shown in figure 8.

C) The pit size is very large and has spread outside the wear scar causing flakes of material to become dislodged from the crown or body of the ball. An example is shown in figure 9.

D) There is no obvious pit. The test may have been stopped due to audible vibration that did not result from the pitting of the driving ball

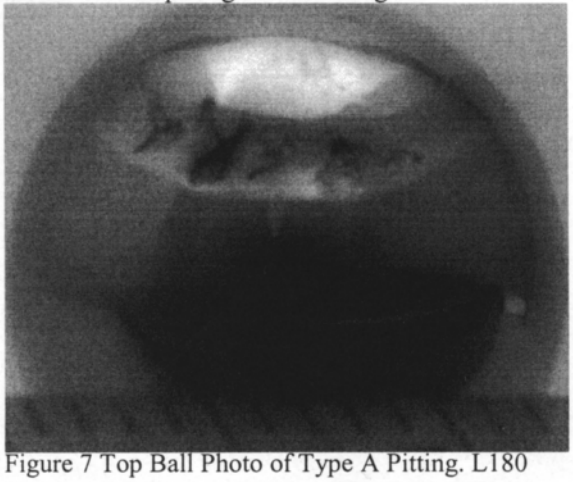


Figure 7 Top Ball Photo of Type A Pitting. L180 Lubricant used. Note Cluster of Pits within Wear Scar. Minor Divisions on Scale Correspond to mm.

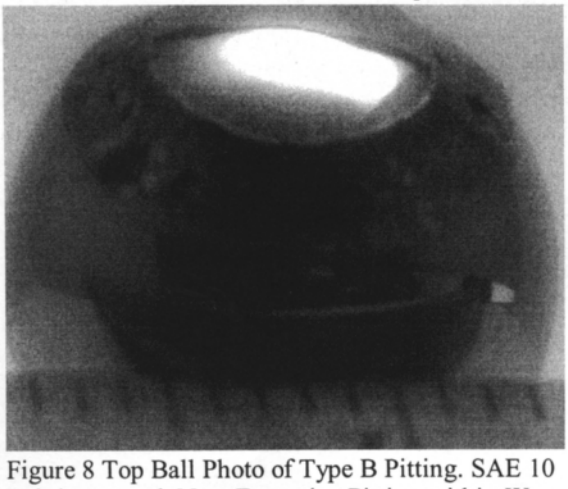


Figure 8 Top Ball Photo of Type B Pitting. SAE 10 Lubricant used. Note Extensive Pitting within Wear Scar. Minor Divisions on Scale Correspond to mm.

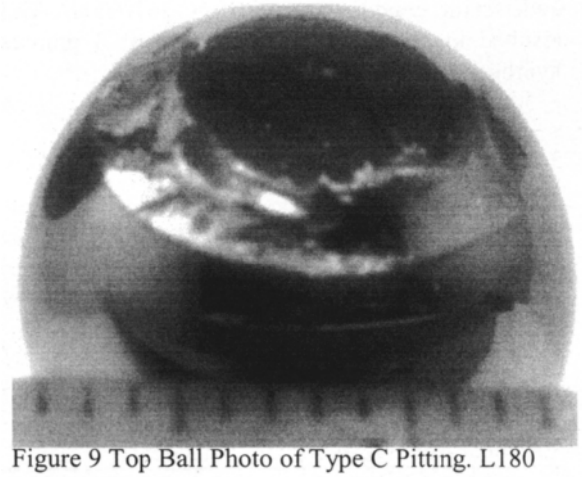


Figure 9 Top Ball Photo of Type C Pitting. L180 with 5% H_2O Lubricant used. Note Brittle Flaking on Crown. Divisions on Scale Correspond to mm.

4. DISCUSSION

The experimental work has shown some trends concerning factors that affect hydrogen evolution in rubbing lubricated contacts. However it is likely that the evolution of hydrogen is dependent on many factors that are coupled. For example temperature, film thickness, friction, wear and plastic deformation are all mutually interdependent.

It is noted that there are two different distinct periods of hydrogen evolution shown by the test data. The first period occurs during normal running and can be seen as a steady increase in the concentration that starts from the start of rolling. This type of hydrogen is referred to as the pre-pit level of hydrogen and is discussed in sections 4.1 – 4.3. Examples of this can be seen in figures 4 and 6. The other period of hydrogen evolution is that associated with fatigue pitting (figure 3). This is a more random event and will be discussed in section 4.4

4.1. Lubricant Viscosity

The difference in viscosity of the two mineral base oils might be the factor that accounts for the difference between the levels of hydrogen generated during the tests (figures 3 and 4), but can equally be due to the severity of pitting. Therefore it is not possible to conclude anything about the viscosity from the maxima associated with pitting The pre-pit

levels are difficult to resolve for SAE 10 oil and therefore no conclusion is drawn. It is noted that SAE 40 oil behaves in a similar way to the greases and this may be due to the similarity in the viscosity of the base oil.

4.2. Temperature

From Figure 5 it is clear that heating the test up causes the greatest evolution of hydrogen out of all the tests under pitting conditions. However, heating the lubricant without rubbing evolves very little hydrogen. Figure 5 shows an increase in background hydrogen due to heating of roughly 3 ppm. This is very small when compared to the pre-pit levels of hydrogen seen in the grease tests. It is also noticed that during the raised temperature test on SAE 10 that the pre-pit levels of hydrogen are similar to the other tests. Therefore it appears that rubbing is a very important factor for the evolution of hydrogen during the pre-pit phase of testing.

4.3. Effect of water and Extreme Pressure (EP) Additive

The pre-pit levels of hydrogen recorded in tables 5 to 7 show that all greases show reduced levels of hydrogen prior to pitting when the EP additive is present, compared to the other conditions. This reduction in pre-pit hydrogen level could be due to surface conditions promoting greater absorption of hydrogen (poisoning). Another possibility is that the surface film of iron sulphide is passivating the surface and preventing the breakdown of lubricant on the fresh metal surface.

The effect of water is very slight. Any effect might be due to its effect on film thickness and heat transfer.

4.4. Effect of Pit Severity

The size and geometry of the pit running through the contact will have some bearing on the conditions within the contact. Rough surfaces with sharp asperities will have much higher temperatures, pressures and wear associated when run through a lubricated contact. Therefore it might be expected that the presence of pits will increase the amount of hydrogen evolved by increasing the severity of the conditions and therefore more lubricant cracking will occur. These extreme conditions associated with asperity contact might be greater for oils of low viscosity. This is due to the expectation that more metal to metal contact will take place due to the thinner films involved. This could lead to more lubricant degradation taking place and therefore could explain the high levels of hydrogen seen from SAE 10 mineral oil during pitting. Another possible explanation for the peak of hydrogen seen during pitting is the release of hydrogen from the freshly exposed substrate and spalled material. This could be due to hydrogen charged debris entering the contact and having the hydrogen effectively squeezed out. A large amount of fresh surface is also produced and might catalytically increase the amount of cracking. The test samples have been ranked according to a phenomenological scale (see section 3.4). The levels of hydrogen observed in the peaks associated with pitting do not always match with the apparent severity of pitting. Type C pitting produces the roughest surfaces and the largest amount of debris but does not produce the largest amount of hydrogen.

Type C pitting involves a flake of material being removed from the crown of the ball or (more rarely) below the wear track. This failure is reminiscent of brittle flaking failure [10] that has been seen in grease-lubricated bearings that undergo rapid acceleration-deceleration tests. The authors showed that the brittle flaking mechanism was associated with an average hydrogen concentration approximately above 1 ppm found over a 20 mm by 6 mm section of raceway in the load zone and that this concentration was the threshold for this type of failure. If the slow diffusion of hydrogen in high strength bearing steels is considered (table 1) the hydrogen levels just below the wear track must be larger to give an average concentration of 1ppm over the section examined. In the present work no samples run using mineral oil showed type C pitting but several tests run with grease showed this failure. This might indicate a greater degree of hydrogen absorption and embrittlement under grease lubricated running.

5. CONCLUSIONS

The sampling of gaseous hydrogen from the 4-ball test head has shown the following important conclusions:

- The tests were unable to resolve the true effect of lubricant viscosity on the gaseous evolution of hydrogen.
- High temperatures up to 105 °C do not evolve appreciable amounts of hydrogen in the absence of rubbing for SAE 10 mineral oil.
- The presence of the sulphur containing extreme pressure additive appeared to decrease the pre-pit level of hydrogen in all greases tested. This may be due to an embrittling condition that involves greater amounts of hydrogen entering the steel, or a layer of iron sulphide could be preventing the catalytic action of fresh surface.
- Some grease tests showed type C (Brittle Flaking) modes of failure that might indicate that the balls were acting as a stronger sink for hydrogen. No mineral oil tests showed this mode of failure.
- Pitting failure is associated with increased levels of gaseous hydrogen. This could be due to evolution by harsh running conditions that cause local spikes in temperature, pressure, plastic deformation and fluid shear thus causing local cracking of the lubricant or entrainment of hydrogen charged debris into the high pressure contact and subsequent liberation of hydrogen.

To fully understand the role of hydrogen in rubbing lubricated contacts an in-situ technique currently being developed needs to be applied to the same lubricants to try and establish the partitioning of hydrogen between steel and gaseous phases.

Acknowledgements

The Authors would like to thank EPSRC and GKN Technology for their assistance with funding and would also like to thank Dr. Phillipa Cann for useful discussions and the supply of grease base stocks.

References

1. Grundberg, L., Jamieson, D.T. & Scott, D., 1963, Hydrogen penetration in water accelerated fatigue of rolling surfaces, The Philosophical Magazine, A Journal of Theoretical and Apllied Physics, Vol. 8, No. 91, pp. 1553-1568.
2. Galaktionowa, N.A., 1981, Hydrogen-Metal systems databook, Published by Ordentlich.
3. Chen, C.A, Wu, S, Ni R.F & Bai, C.M, 1999 , Study on the bulk diffusion of hydrogen in stainless steels, Surface and Interface Analysis, Vol. 28, pp. 77-80.
4. Ciruna, J.A. & Szieleit, H.J., 1973, The effect of hydrogen on the rolling contact fatigue life of AISI 52100 and 440C steel balls, Wear, Vol. 24, pp. 107-118.
5. BS standard 7886:1997, Method of measurment of hydrogen permeation and the determination of hydrogen uptake and transport in metals by an electrochemical technique.
6. Hirth, J.P., 1980, Effects of hydrogen on the properties of iron and steel, Metallurgical Transactions A, Vol. 11A, pp. 861-890.
7. Newman J.F. and L.L. Shrier, Role of Hydrides in Hydrogen entry into Steel from Solutions Containing Promoters, 1969, *J.I.S.I.*, 207, 1369
8. Javier.L., Blunt T., Kotvis P., Riga A. & Tysoe W.T., 1998, Surface Chemistry and Extreme-Pressure Lubricant Properties of Dimethyl Disulfide. *J. Phys. Chem. B,* Vol.102 pp. 1703-1709
9. IP designation 300/82, (1987), Determination of Rolling Contact Fatigue Failure – Modified Four Ball Machine Method, Standard Methods for Analysis and Testing of Petroleum and Related Products, (!992), Vol 2, pp. 300.1-300.7.
10. Tamada, K. & Tanaka, H., 1996, Occurrence of brittle flaking on bearings used for automotive electrical instruments and auxiliary devices, Wear, Vol. 199, pp. 245-252.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

The effect of superimposed low-frequency modulation on lubrication in machining

E. Olgun, W. D. Compton, and S. Chandrasekar

Center for Materials Processing and Tribology, School of Industrial Engineering – Purdue University
1287 Grissom Hall, West Lafayette, IN 47907-1287, USA

A study has been made of the effect of an externally imposed, low-frequency modulation on the action of lubricants in machining. It is shown that in conventional machining, lubricant action is essentially confined to the edges of the chip-tool contact, with little or no change in the friction condition over much of the rake face. However, the effectiveness of lubricant action is significantly enhanced when a controlled low-frequency modulation of sufficient amplitude is imposed in the direction of cutting. Measurements show that the friction coefficient on the rake face is reduced by a factor of two to three in the presence of such a modulation. The nature of metal transfer onto the rake face is also substantially changed by the modulation.

1. INTRODUCTION

The nature of contact between chip and tool along the rake face has a major influence on the mechanics of machining. The friction condition on the rake face is an important parameter controlling chip-tool contact length and shear plane angle [1,2,3], thereby influencing the deformation, cutting forces, and energy consumption [4]. Additionally, the frictional interaction between the chip and tool along the rake face plays an important role in determining tool wear and the effectiveness of cutting fluids. Thus, any modification of the friction condition along the rake face, for example through better lubrication or use of a low friction coating, will have important ramifications for the machining process.

The chip-tool interface represents a unique tribological contact with characteristics quite different from those of conventional sliding systems [5]. This interface is characterized by high normal and shear stresses whose distribution is highly non-uniform. The freshly generated chip surface, especially at the location where it is generated, is atomically clean, devoid of contaminants, and potentially chemically active [3]. The real area of contact between tool and chip is essentially equal to the apparent area of contact over much of the interface. It has been suggested that under these conditions, a lubricant is unlikely to penetrate into much of the tool-chip contact region [1,6,7,8]. Consequently, the effectiveness of a lubricant will be controlled by the degree to which it can penetrate into the tool-chip contact region.

While models explaining the action of a metal-cutting lubricant are few, all of these suppose that the rate-controlling step in lubrication is transport of lubricant molecules into the tool-chip contact region [8,9]. Mechanisms proposed for fluid access to the region of intimate contact along the rake face include fluid flow through a network of connected capillaries between tool and chip along the rake face [6,10]; fluid penetration through voids associated with built-up edge formation [1]; and flow of fluid through fissures in the chip [11]. Vapor phase lubrication of the tool-chip contact has also been suggested and demonstrated [12].

A potentially attractive way for creating 'space' in the tool-chip contact that could enable a fluid to penetrate into this region is by physical perturbation of the contact. In a recent paper [13] it was demonstrated that a controlled, low-frequency modulation, applied normal to the rake face, has a major influence on the contact conditions (adhesion and friction) at the tool-chip interface. Physical separation of the tool from the chip in the contact region could be achieved if the modulation was of sufficient amplitude. The present study was undertaken with the objective of determining what effect, if any, such a super-imposed modulation would have on the effectiveness of lubricant action along the tool-chip interface. As will be seen, the application of the modulation greatly facilitates

lubricant penetration into the tool-chip contact with attendant, dramatic decreases in the coefficient of friction in 2-D orthogonal cutting.

2. EXPERIMENTAL

Figure 1 shows a schematic of the experimental configuration developed to carry out 2-D orthogonal cutting with modulation applied in the direction of cutting. A wedge-shaped tool is mounted in an L-bracket fixture on a vertical linear slide that is driven by a specially designed linear motor. This linear slide is capable of providing continuous feed motion to the tool in the vertical direction (z-axis) along with a low frequency, sinusoidal modulation superimposed onto the feed motion. Details of the characteristics of the linear drive and its construction are described elsewhere [13]. The workpiece is rigidly mounted directly beneath the tool on a 4-component, piezoelectric force dynamometer (Kistler Model 9272) using a stable custom built fixture. A Newport positioning stage equipped with a micrometer is used to accurately set a depth of cut on the workpiece. This depth of cut is provided by a translation of the stage along the x-axis (Fig. 1). The actual depth of material removed is measured after cutting by a dial gage. Repeatability of this depth measurement is 2 μm. Thus, the motion of the vertical linear slide determines the cutting velocity for the experiment while the Newport positioning stage provides the depth of cut.

When a modulation is superimposed in the cutting direction onto the tool (Fig.1), it causes a variation in the in the instantaneous cutting velocity V_c' that is given by:

$$V_c' = V_c - 2\pi f\left(\frac{a}{2}\right)\sin(2\pi ft)$$

where V_c is the main cutting velocity, i.e. the continuous feed velocity provided by the drive, a is the peak-to-peak modulation amplitude, and f is the modulation frequency. If the product of the modulation amplitude and modulation frequency is sufficiently large, then the instantaneous cutting velocity will reach zero and subsequently negative values over each cycle of modulation and the tool reverses direction for a portion of the cut.

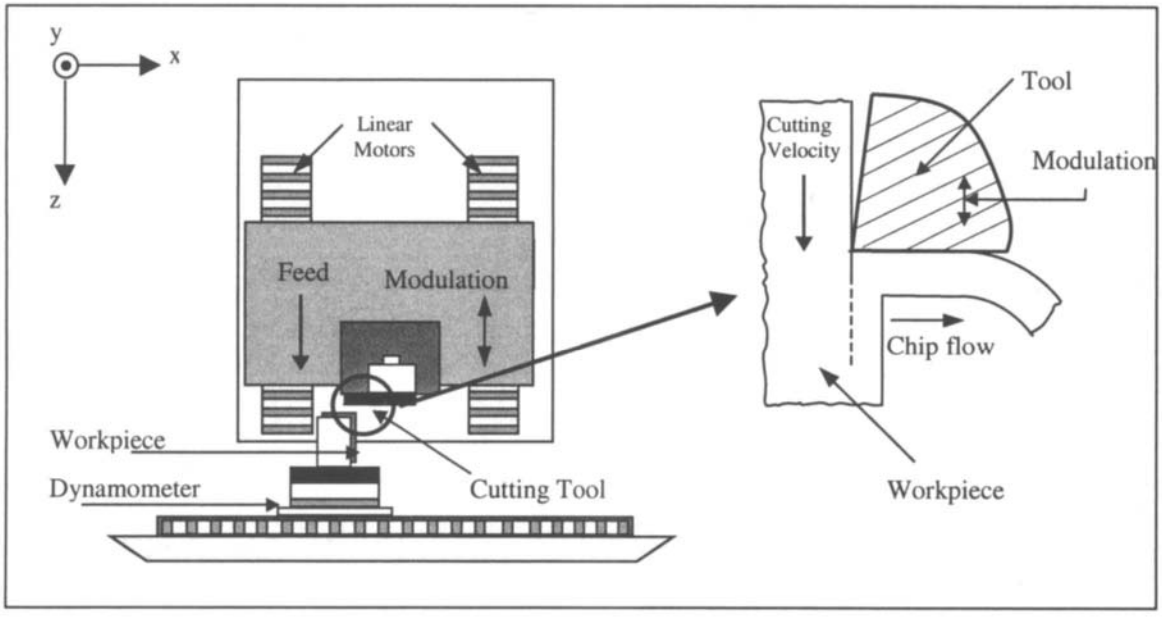


Figure 1. Schematic of the experimental set-up for 2-D orthogonal cutting with low frequency modulation. Inset shows the modulation direction with respect to the tool-chip interface.

Rectangular workpieces of 100 mm length, 50 mm thickness, and 1.6 mm width (2 mm in case of lead) were used in the experiments. The workpiece materials were commercially pure lead (99.95%), aluminum 6061, and Oxygen-Free High-Conductivity (OFHC) copper (99.95%). Mechanical properties of the workpiece materials are given in Table 1. The materials were chosen for their varying hardness, ductility, and strain hardening properties. Of the three materials, OFHC copper strain hardens the most, while lead, which typically recrystallizes at room temperature, is significantly softer than the aluminum and the copper. Furthermore, the interaction of copper with a high-speed steel tool surface is known to be quite different from that of Al and Pb in conventional cutting [13]. Each machining cut was made along the width of a workpiece and over its entire length; the width of the tool cutting edge was substantially greater than the width of the workpiece so that a state of 2-D, plane strain deformation prevailed.

High speed steel tools (M-42 grade), having zero degree rake angle, twenty degree clearance angle, and a controlled contact length of 0.3 mm along the rake face, were used in the experiments. The tools were ground on the rake and flank faces to ensure a sharp cutting edge (~2 μm cutting edge radius) and a negligible flank-land. Furthermore, a groove was machined on the rake face of each tool with a CBN grinding wheel such that the tool-chip contact length along this face was fixed at 0.3 mm [13]. In the present experiments, this contact length was less than the natural contact length corresponding to the zone of intimate contact between tool and chip for the specific cutting conditions used. The tools were cleaned prior to each cutting experiment.

The cutting and thrust forces on the workpiece were measured using the piezoelectric dynamometer. Since the tools have a zero degree rake angle, the normal force on the rake angle face is equal to the cutting force while the friction force is equal to the thrust force. The ratio of the friction force to the normal force is the coefficient of friction (μ) along the rake face. A National Instruments data acquisition system was used to digitize the force data for processing and data analysis. A scan rate of 4000 per second was used to acquire the force data.

2.1. Lubrication

In some of the experiments, a lubrication system (Unist Standard Cool-Lubricator) was used to provide a continuous supply of lubricant onto the rake face. This supply, in the form of an air-mist, was controlled by an air metering screw. The cutting fluid, Cool-Lube 2210, was delivered to the tool rake face as a spray through a mix nozzle that was coupled to a co-axial air hose. Cool-Lube 2210 (Triglycerol & Propylene glycol esters of C8 & C10 acid) is a vegetable-based metal cutting lubricant.

2.2. Post-Cutting Analysis

Post-cutting analysis involved microscopic observations of the contact zone on the rake face of the tool after cutting under different conditions, roughness measurements on the newly cut workpiece surfaces, and recording the amount of material side flow on the workpiece. These observations proved to be useful in explaining the effect of different cutting conditions on the cutting process with different workpiece materials. Microscopic observations of the tool rake and workpiece surfaces were done before and after each experiment to obtain insight into the manner in which the junction between the tool and chip separates. These were performed with a Nikon Epiphot 200 optical microscope equipped with a Sony DXC-30 camera. Observations were also done on regular tools (not of controlled-contact type) used in cutting experiments. The microscope observations discussed later are from these regular tools.

Table 1. Mechanical properties of work materials

Workpiece Material	Tensile Strength [MPa]	Yield Strength [MPa]	Young's Modulus [GPa]	Elongation at break [%]	Brinell Hardness [kg/mm^2]
Aluminum 6061	310	275	69	12	95
Lead	18	9	14	55	4
OFHC Copper	224-314	54-270	129	45	90-100

The roughness of the newly cut workpiece surfaces was measured using a Form Talysurf 50 profilometer (Rank Taylor Hobson). Arithmetic average surface roughness (R_a) values were recorded at five different locations along the length of the workpiece and averaged to obtain a mean value for the R_a. This mean R_a value is reported.

2.3. Cutting Conditions

All experiments were done at a cutting velocity of 2 mm/s and a modulation frequency of 75 Hz. The depth of cut was chosen to be 200 μm for aluminum and lead, and 130 μm for copper due to power limitations of the slide. For each workpiece material, two sets of experiments were conducted. In one set of experiments, the workpieces were cut without the application of a lubricant. In the second set of experiments, the lubricant was applied during cutting. Each set of experiments involved the use of a low-frequency modulation. Several amplitudes were used ranging from 0-80 μm. For comparison, conventional orthogonal cutting experiments were also carried out dry and in the presence of the lubricant.

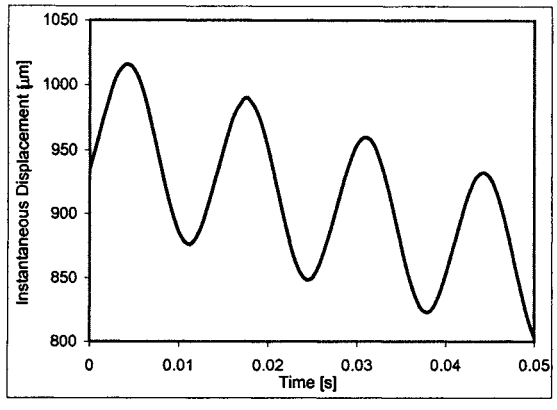


Figure 2. Sample capacitance probe output showing displacement of the tool with respect to the workpiece during cutting. Workpiece material: Aluminum 6061; modulation frequency: 75 Hz; dry cutting. The modulation amplitude is 65 μm as derived from the measurement.

The modulation amplitude during cutting was measured using a capacitance probe (Capacitec HPC-75-13090). Figure 2 is a typical output of the capacitance probe, which shows the displacement of the tool with respect to the workpiece in a modulation assisted cutting experiment. The modulation amplitude (half peak-to-peak value) as determined form the capacitance probe output in Figure 2 is 65 μm. In all of the experiments, the modulation amplitude was directly obtained from the capacitance probe measurements. The actual depth of cut was measured using a dial gauge. It is this depth of cut that is reported in the results. In this manner, all parameters of interest were either measured or controlled.

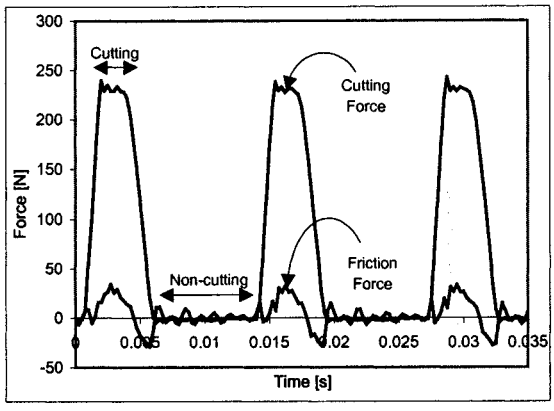
3. RESULTS

3.1. Forces and Friction Coefficient

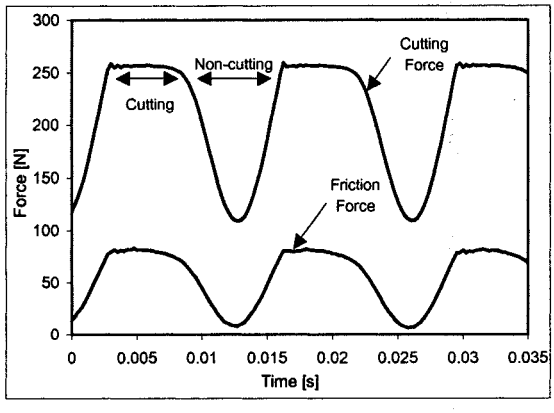
Figures 3a) and 3b) show typical cutting and friction force traces recorded when cutting Al 6061 with modulation. The forces in the figure were recorded over three cycles of modulation. The force signals in each cycle of modulation can be broadly divided into two parts – one corresponding to the cutting part and the other to the non-cutting part – as labeled in Figures 3a) and 3b). During the cutting part, which corresponds to the duration in which the forces have their highest values, the tool is engaged with the workpiece and material removal takes place. The part of the cutting cycle when the friction and cutting forces are zero or negative (Fig. 3a)) corresponds to the period of time during which the tool is physically separated from the chip. This separation will occur only if the modulation amplitude is sufficiently high; this will henceforth be referred to as high-amplitude modulation. In contrast, if the modulation amplitude is not high enough for the tool to separate from the chip, the forces will not go to zero during the non-cutting part of the modulation cycle (Fig. 3b)). This latter instance will be referred to as low-amplitude modulation. The period of oscillation of the force traces in Fig. 3 corresponds to the modulation frequency of 75 Hz.

The friction and cutting force values in modulation-assisted cutting were obtained by averaging the respective force values in the cutting part of the modulation cycle (Fig. 3). Typically, this averaging was done over 750 consecutive cycles of modulation. For this purpose, a simple numerical routine was written in MATLAB® that filtered out the 'flat' top portion of the force data and subsequently computed the average over the 750

cycles. It is these averaged force values that are, henceforth, reported for each cutting condition. For the case of conventional cutting, where there is no oscillation of the forces as in the case of modulation, the average forces were computed over a time duration that was about the same as that corresponding to the 750 cycles of modulation.



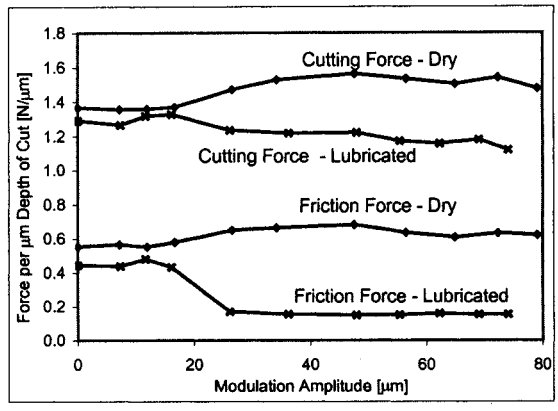
a)



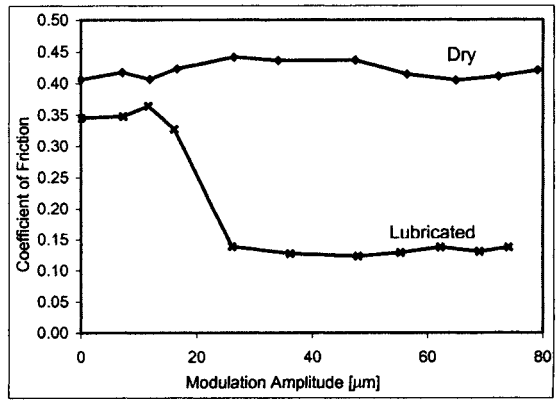
b)

Figure 3. Sample force trace from lubricated cutting of Al 6061 with modulation. Three cycles of modulation are shown. a) Modulation Amplitude: 74 μm. The cutting and non-cutting portions are marked in the figure. b) Modulation Amplitude: 16 μm.

Figures 4a) and 4b) show the variation of forces and coefficient of friction (μ) with modulation amplitude, respectively, for Al 6061. Results are shown for both dry and lubricated cutting. The force and μ values for a modulation amplitude of 0 μm correspond to the case of conventional cutting. It is seen from Figure 4a) that the cutting and friction forces in lubricated cutting are smaller beyond a modulation amplitude of ~25 μm. This decrease in the force value is especially significant (~50%) with the friction force, suggesting that lubricant action along the tool-chip interface is enhanced by the use of modulation. It should be noted that only for modulation amplitudes in excess of 25 μm does the normal (cutting) force go to zero for a portion of the modulation cycle as in Fig. 3a). That is, the amplitude is sufficiently high for the tool to break contact with the chip.



a)



b)

Figure 4. Variation of a) forces and b) coefficient of friction (μ) with modulation amplitude for dry and lubricated cutting of Al 6061. Modulation amplitude of 0 μm corresponds to conventional cutting. Modulation frequency: 75 Hz.

Further and quite dramatic evidence of the interactive effects of modulation and lubrication can be seen in the coefficient of friction values shown in

Figure 4b). Here, μ is seen to decrease quite significantly for lubricated cutting from a value of ~0.35 to about 0.14 when the modulation amplitude exceeds 25 μm. The coefficient of friction variation for lubricated cutting in Fig. 4b) is characterized by two distinct regions – a region of relatively high but constant μ (~0.35) at the lower modulation amplitudes, where contact between the chip and tool is not broken, and a region of low μ (~0.14) at the higher modulation amplitudes, where this contact is broken. In contrast, very little change is seen in the forces or μ with modulation amplitude in dry cutting (Figs. 4a) and 4b)). Also, the effect of the lubricant in conventional cutting is relatively small with μ having a value of ~0.4 in dry cutting as against a value of 0.35 in lubricated cutting (Fig. 4b)). This would appear to suggest that lubricant action (and lubricant penetration) in the zone of intimate contact between chip and tool is negligible in the absence of modulation. This is consistent with observations made of the rake face contact when cutting with optically transparent tools [14,15].

The interactive effect of modulation and lubrication on the friction at the tool-chip interface was found to be similar in lead and OFHC copper (Fig. 5). As with Al 6061, μ is seen to become significantly smaller, by a factor of two or three, at modulation amplitudes beyond 25 μm. For example, in the lubricated cutting of copper, μ decreases from 0.34 to about 0.12 while with lead, μ decreases from a value of 0.38 at low amplitudes of modulation to about 0.18 at the high amplitudes. These observations are consistent with the hypothesis that the fluid (lubricant) penetrates and lubricates the zone of intimate contact between chip and tool when the amplitude of modulation is sufficiently large to break this contact during the cutting. However, copper did show one important difference in its cutting characteristics in the presence of modulation when compared to Pb and Al 6061. This difference pertains to the effect of modulation on μ in dry cutting. Figure 5 shows that the μ value in dry cutting of copper decreases from 0.45 to 0.36 at modulation amplitudes in excess of 40 μm. Such a decrease, though small, was not observed in the dry cutting of aluminum and lead. This suggests that the nature of the chip-tool contact may be fundamentally different in the case of copper.

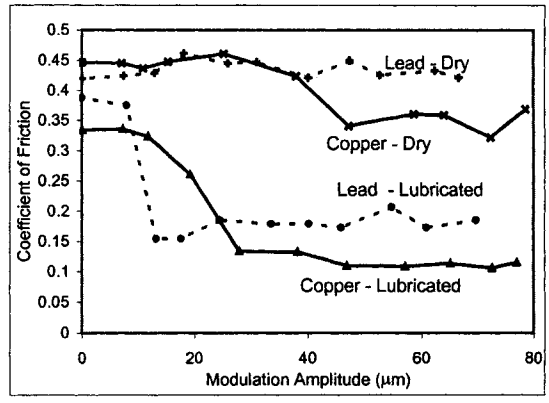


Figure 5. Variation of coefficient of friction (μ) with modulation amplitude for dry and lubricated cutting of OFHC copper and lead. Modulation frequency: 75 Hz.

3.2. Rake Face Condition

Figure 6a) shows an optical microscope image of the rake face of a regular HSS tool (not controlled-contact) after conventional dry cutting of the Al 6061. The tool-chip contact region is seen to be composed of two distinct zones: 1) a zone adjoining the cutting edge and extending about two-thirds of the length of the contact, that is devoid of visible metal deposits and 2) an adjoining zone bordering the edges of the tool-chip contact that is characterized by metal transfer from the chip onto the tool rake face. The former zone is a zone of intimate contact between tool and chip with the real and apparent areas of contact being equal while the latter zone is characterized by intermittent contact between tool and chip [14-16]. When conventional cutting is carried out in the presence of the lubricant, no metal transfer appears to take place in the zone of intermittent contact (Fig. 6b)). This suggests that the lubricant is active in this zone along the edges of the tool-chip contact.

Figure 6c) shows an optical microscope image of the rake face of an HSS tool after dry cutting of Al 6061 at a high amplitude of modulation. The rake face image appears to be the inverse of Fig. 6a) with metal deposits in the zone of intimate contact adjoining the cutting edge and negligible metal transfer along the edges of the tool-chip contact. Metal transfer in the zone of intimate contact in this case is likely a consequence of rupture taking place within the chip material, than at the tool-chip

interface, when the tool is retracted from the chip during a cycle of modulation. When the same modulation cutting experiment is carried out in the presence of the lubricant, only a small amount of metal transfer is visible on the tool rake face in this zone of intimate contact (Fig. 6d)). This suggests that in the presence of modulation, the lubricant penetrates the intimate contact zone and lubricates this zone. This is also consistent with the significant reduction in μ observed in lubricated cutting at a high amplitude of modulation. The rake face observations when cutting lead were found to be the same.

There was one important difference observed in the rake face condition when cutting copper. In none of the different instances of cutting, be it with or without lubrication, or with or without modulation, was any metal deposit visible on the rake face of the HSS tool. This may be seen in Figures 7a) and 7b) which show rake face images after conventional and modulation-assisted cutting of copper without the use of lubricant. These should be contrasted with Figures 6a) and 6c). No metal deposits are visible in the intimate contact zone or in the adjoining region in Fig. 7. Indeed, it was difficult in these experiments to discern the true extent of the tool-chip contact region. These observations suggest that even in the dry modulated cutting of copper, the nature of the tool-chip interface is quite different from that in the cutting of lead and Al 6061. The interface appears to be weaker than the chip material when subjected to a tensile loading, a consequence of which is that rupture occurs at the interface than within the chip when the tool is retracted during the modulation. It is this difference in the nature of the interface which may be responsible for the observed variation of μ with modulation in dry cutting of copper (Fig. 5).

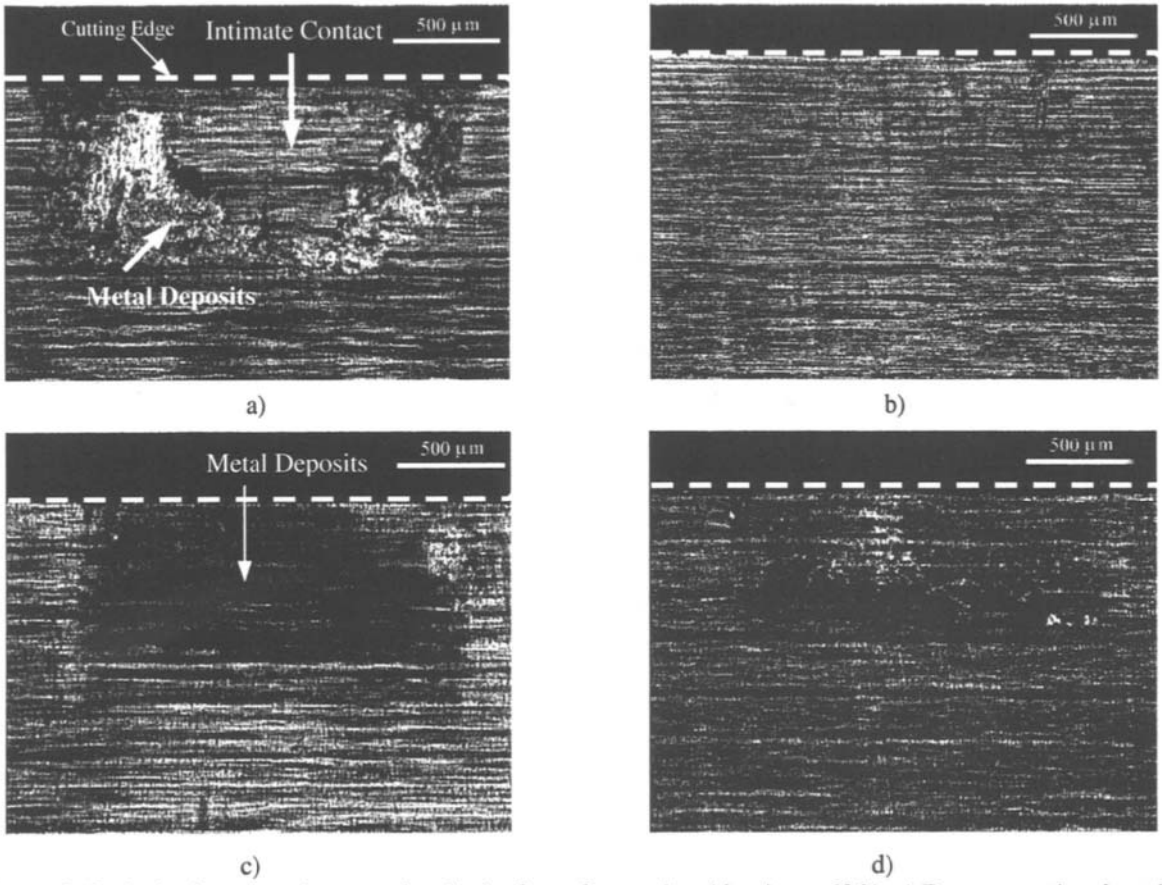


Figure 6. Optical microscope images of tool rake face after cutting Aluminum 6061. a) Dry conventional cutting b) Lubricated conventional cutting c) Dry, high-amplitude modulation cutting d) Lubricated, high-amplitude modulation cutting. Modulation frequency: 75 Hz, modulation amplitude: 70 µm.

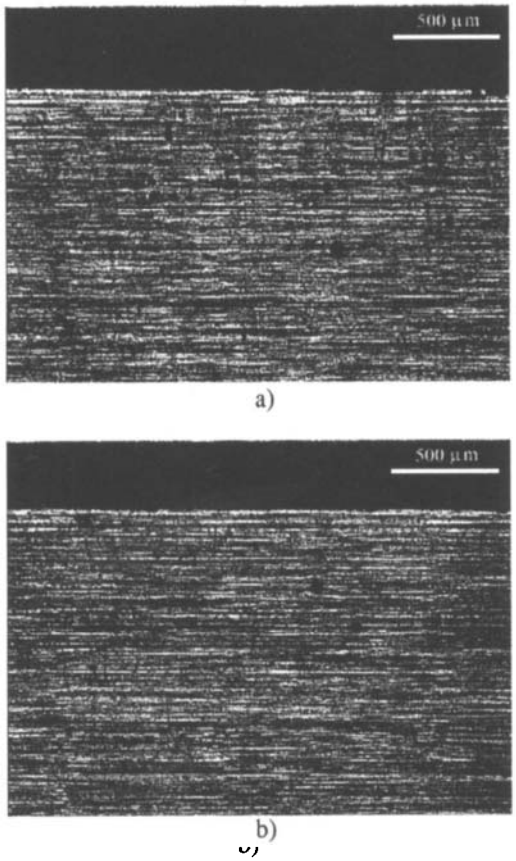


Figure 7. Optical microscope images of the tool rake face after cutting OFHC copper. a) Dry conventional cutting b) Dry, high-amplitude modulation cutting. Modulation frequency: 75 Hz, modulation amplitude: 70 µm. Note the complete absence of any material deposit along the rake faces in contrast to Figs. 6a) and 6c)

3.3. Surface Finish and Material Side Flow

Two other striking effects of high-amplitude modulation were observed in the surface finish on the workpiece and the side flow of material around the workpiece edges. In copper, for example, the surface roughness (R_a) on a machined surface cut with modulation and lubricant was 1.69 ± 0.1 µm as against an R_a value of 3.08 ± 0.3 µm recorded on a conventionally machined surface. Improvements in surface finish, due to the application of modulation, were also observed in lead and Al 6061 though the extent of this improvement was smaller than in copper. The use of modulation was also found to significantly reduce the extent of side flow of material along the transverse edges of the workpiece, in comparison to conventional cutting. This would suggest that the extent of burrs developed the near workpiece edges is likely to be significantly less in modulation assisted cutting processes. Such a reduction in burrs has indeed been observed in holes made using a modulation-assisted drilling process [17].

4. DISCUSSION

Our experiments have shown that when a low frequency modulation of sufficiently high amplitude is applied normal to the rake face of the tool, the tool-chip contact is repeatedly broken during each cycle of modulation, thereby, enabling the fluid to penetrate into this contact region. Under such conditions, the fluid is found to effectively lubricate the contact. This occurs even when the fluid is applied in the form of a spray mist as in the present experiments, with the fluid volumetric rate being only about one milli-liter per minute. Even smaller volumetric flow rates for the lubricant may suffice as long as the tool-chip contact is repeatedly broken. The effectiveness of this lubricating action is best demonstrated in the dramatic decreases observed in the coefficient of friction at the rake face (Figs. 4b) and 5). As may be seen from these figures, there is a two– to three-fold reduction in the coefficient of friction at the high amplitudes of modulation.

The observations have also shown that lubricant action in the zone of intimate contact between tool and chip is only marginally effective in conventional cutting or even when cutting with a low amplitude of modulation. These conclusions are based on a study of the coefficient of friction and of metal transfer on the rake face noted in the optical microscope images (Figs. 6 and 7). Furthermore, these observations indicate that lubricant action under these conditions of cutting is almost entirely confined to the edges of the tool-chip contact, a region where the contact is far from intimate [14-16]. A particularly intriguing aspect of our observations is the evidence from coefficient of friction measurements (Figs. 4b) and 5), which suggests that there is insufficient penetration of the lubricant into the chip-tool contact even when cutting at the lower amplitudes of modulation. This would appear to be at variance with a suggested mechanism of lubrication of the tool-chip interface involving fluid penetration into

this interface through a network of inter-connected capillaries along the contact [6,10].

Ongoing work is aimed at exploring in greater detail the mechanism of fluid penetration into the chip-tool interface, direct observation of fluid action at the interface using optically transparent tools, and the role of the workpiece-tool material combination in influencing the contact condition along the chip-tool interface.

5. CONCLUSION

When a low frequency modulation of sufficiently high amplitude is used to perturb the tool-chip contact in a direction normal to the rake face, the cutting fluid is found to penetrate this interface and provide effective lubrication of the contact. The effectiveness of this lubrication is seen in the two- to three-fold decrease observed in the coefficient of friction along the rake face. The lubricant is found to be only slightly effective in conventional cutting or even when cutting with super-imposed modulation of relatively low amplitude. The lack of effectiveness of the fluid in these instances appears to be due to insufficient penetration of the fluid into the chip-tool contact region. The results suggest that the application of modulation in conjunction with only a minute amount of application of lubricant in the form of an air mist, can greatly enhance the efficiency of cutting processes through reduction in tool-chip friction and tool wear. The attendant economic and environmental benefits could be significant.

6. ACKNOWLEDGEMENT

This research was supported by Ford Motor Company.

REFERENCES

1. Childs, T.H.C., 1972, *Proc. Inst. Mech. Engrs.* (U.K.), 186, 717-726.
2. Merchant, M.E., 1945, Mechanics of the metal cutting process, *J. Appl. Phys.*, **16**(5), 267-275.
3. Shaw, M.C., 1984, *Metal Cutting Principles*, Oxford Series on Advanced Manufacturing, Clarendon, Oxford.
4. Zorev, N.N., 1963, Interrelationship between shear processes occurring along tool face and on shear plane in metal cutting, *Proceedings of the International Production Engineering Research Council*, ASME, Pittsburgh, Pennsylvania, 42-49.
5. Finnie, I., and Shaw, M.C., 1956, The friction process in metal cutting, *Transactions of the ASME*, **77**, 1649-1657.
6. Merchant, M.E., 1957, Cutting fluid action and the wear of cutting tools, *Proceedings of Conference on Lubrication and Wear*, Institution of Mechanical Engineers, London, 566-574.
7. Bailey, J.A., 1975, Friction in metal machining – mechanical aspects, *Wear*, **31**, 243-275.
8. Williams, J. A. and Tabor, D., 1977, The Role of Lubricants in Machining, *Wear*, **43**, 275-292.
9. Smith, T., Naerheim, Y. and Lan, M. S., 1988, Theoretical Analysis of cutting fluid interaction in machining, *Tribol. Int.*, **21**, 239-247.
10. Williams, J. A., 1977, The action of lubricants in metal cutting, *J. Mech. Engng. Sci.*, **19**, 202-212.
11. Cassin, C. and Boothroyd, G., 1965, Lubricating action of cutting fluids, *J. Mech. Engng. Sci.*, **7**, 67-79.
12. Wakabayashi, T., Williams, J. A., Hutchings, I. M., 1995, The kinetics of gas-phase lubrication in the orthogonal machining of an aluminum alloy, *Proc. Instn. Mech. Engrs.*, **209**, 131-136.
13. Ackroyd, B., Akcan, N. S., Chhabra, P., et al., 2001, Exploration of contact conditions in machining, *Proc. Instn. Mech. Engrs.*, **215**(4), 493-507.
14. Madhavan, V., Chandrasekar, S., Farris, T. N., 2002, Direct observations of the tool-chip interface in the low speed cutting of pure metals, *Journal of Tribology*, **124**(3), 617-626.
15. Ackroyd, B., Chandrasekar, S., and Compton, W. D., 2002, A model for the contact conditions at the chip-tool interface in machining, accepted for publication in *Journal of Tribology*.
16. Doyle, E.D., Horne, J.G. and Tabor, D., 1979, Frictional interactions between chip and rake face in continuous chip formation, *Proceedings of the Royal Society of London*, 366, 173-187.
17. Ramaswamy, D., Tucker, A., Olmez, A., el al., 2000, A study of the effect of process parameters on surface finish of drilled holes, *Transactions of the NAMRI / SME*, **28**, 227-232.

This Page Intentionally Left Blank

SESSION XX
ENGINE TRIBOLOGY (2)

Chairman :	Professor J F Booker
Paper XX (i)	**Optimisation of Piston Assembly Tribology for Automotive Applications**
Paper XX (ii)	**Optimised Additive - Surface Material Combinations for Reduced Friction and Wear in Internal Combustion Engines**
Paper XX (iii)	**Prediction of Total Engine Friction Power Loss from Detailed Component Models**
Paper XX (iv)	**Experimental Evaluation of Engine Valve Train Friction Under Motored and Fired Conditions**
Paper XX (v)	**Inertia, Shear-Thinning and Thermal Effects on Connecting Rod Bearing Behaviour**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Optimisation of Piston Assembly Tribology for Automotive Applications

M. Priest
Institute of Tribology, School of Mechanical Engineering, The University of Leeds, UK

Current trends in engine design, such as higher operating temperatures and the enhanced use of exhaust gas recirculation (EGR) in diesel engines, and legislative and consumer pressures on lubricants, such as the requirement for reduced levels of phosphorous and extended oil drain intervals, are detrimental to the piston assembly. Future engines will require more wear resistant surfaces in the piston assembly with minimal scope for rapid running-in of the components to establish optimisation on behalf of the engine designer and compatibility issues for the lubricant formulator as iron based materials become less common. In this paper a review is presented of the current understanding of piston assembly tribology in order to contribute to methods of future design optimisation.

1. INTRODUCTION

The piston assembly is a key engineering system in the reciprocating internal combustion engine and central to its efficient and reliable operation. It is at the heart of the engine, forming a vital link in transforming the energy generated by combustion of the fuel and air mixture into useful kinetic energy.

A typical piston and piston ring pack from a modern automotive engine is shown in Figure 1. The top two rings are referred to as the compression rings and utilise gas pressure to supplement the inherent elasticity of the ring to maintain an effective combustion chamber seal. The top compression ring is the primary gas seal and, as the ring nearest the combustion chamber, encounters the highest loads and temperatures. It usually has a barrel-faced profile with a wear resistant coating such as chromium, flame-sprayed molybdenum, metal or cermet mixtures on the periphery and occasionally on the flanks. The second compression ring, which is sometimes referred to as the scraper ring, is designed to assist in limiting upward oil flow in addition to providing a secondary gas seal. As such it has a taper-faced, downward scraping profile which is not normally coated. The bottom ring in the pack is the oil-control ring which has two running faces, or lands, and a spring element to enhance radial load. As its name suggests, the role of this ring is to limit the amount of oil transported from the crankcase to the combustion chamber and by design it has no gas sealing ability.

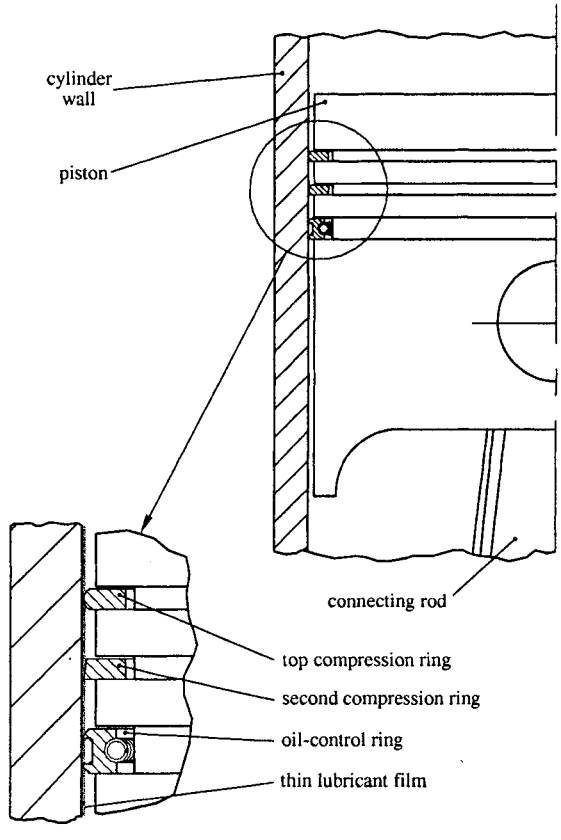


Figure 1: Piston assembly from an automotive engine

The periphery of the oil-control ring lands and, occasionally, the flanks are often chromium plated. Neale [1] gives an excellent summary of the piston ring types in common use.

The primary role of the piston ring pack is to maintain an effective gas seal between the combustion chamber and the crankcase. The rings of the piston ring pack, which together effectively form a labyrinth seal, achieve this by closely conforming to their grooves in the piston and to the cylinder wall. The small quantity of gas that does find its way into the crankcase, blow-by, is normally piped back to the inlet valve and fed back into the cylinder.

In addition to causing a dramatic increase in pressure, the combustion event generates a large amount of heat. Much of this thermal energy is convected into the piston causing a marked increase in the temperature of the piston, which is dissipated by heat transfer to adjacent components and the engine coolant. The secondary role of the piston ring pack is to transfer this heat from the piston into the cylinder wall and thence into the coolant.

The final function of the piston ring pack is to limit the amount of oil that is transported from the crankcase to the combustion chamber. This flow path is probably the largest contributor to the oil consumption of an engine and leads to an increase in harmful exhaust emissions as the oil mixes and reacts with the other contents of the combustion chamber [2]. The desire to extend service intervals of engines and minimise harmful exhaust emissions to meet ever more stringent legislative requirements, means that the permissible oil consumption levels of modern engines are very low compared to their predecessors of twenty years ago [3].

The piston ring pack must fulfil these three roles with a minimum of frictional power loss, most notably at the sliding interface with the cylinder wall, and a minimum of wear in order to maximise component life. Unfortunately, the piston assembly is one of the largest sources of friction in the internal combustion engine over the normal range of engine speeds and loads encountered in service [4-7]. Exact figures vary from engine to engine, but typically the piston assembly, comprising both the piston rings and the piston skirt, accounts for 40-50% of total engine friction. Piston ring pack friction losses are greater than those of the piston skirt at low to moderate engine speeds but the situation may be reversed at high engine speeds due to the large wetted area of the piston skirt contributing to viscous friction [8].

In terms of wear, there is insufficient understanding of the interaction with the lubrication process. So, even though manufacturers can produce rings that have an excellent life expectancy, these components may be far from optimum from a lubrication and friction standpoint [9].

Interactions between the piston, piston rings, cylinder and the lubricant are complex and sensitive to a whole range of engine configuration and operating conditions. The need to maximise gas sealing in the combustion chamber and minimise oil consumption with low friction and wear presents the engine developer with a number of directly conflicting design parameters.

Those technology drivers for the automotive industry relevant to engine tribology were recently addressed by Korcek *et al* [10] and are summarised in Figure 2.

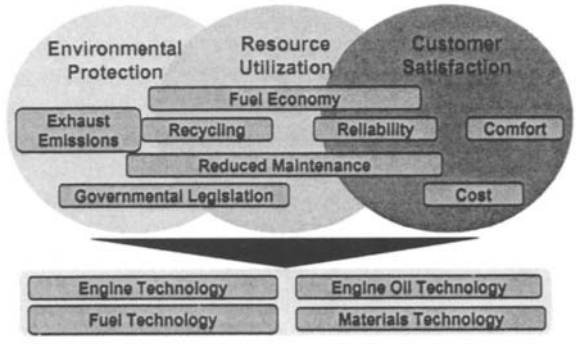


Figure 2: Technology drivers for a major automotive vehicle manufacturer, after Korcek *et al* [10]

The need to produce vehicles which use resources wisely, have minimal impact on the environment and that customers will want to own and drive presents major challenges to engine designers and lubricant formulators. For example, the automotive industry is currently developing the next generation of engine lubricants, with important and potentially conflicting implications for piston assembly tribology. The new lubricant specification, known as ILSAC GF-4, places strict demands on fuel economy and its retention as the oil ages in service and the level of phosphorus permitted in the formulation. Phosphorus is known to contaminate catalytic after-treatment devices in the exhaust,

reducing their efficiency, but in the form of zinc dialkyldithiophosphate (ZDDP) is the most cost-effective anti-oxidant and anti-wear agent currently available.

The conflicting requirements for enhanced fuel economy, emissions system compatibility and extended useful lubricant life demands improvements in the

- Thermodynamic performance of the engine
- Mechanical design of the piston assembly
- Performance of the lubricant within strict compositional constraints
- Novel materials and coatings to control wear with less opportunity for running-in. A shift away from iron based materials is likely.

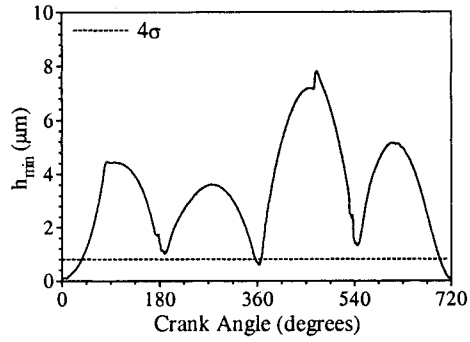
A major personal reason for writing this paper was to contribute to methods of future design optimisation for piston assemblies.

2. SOFTWARE TOOLS

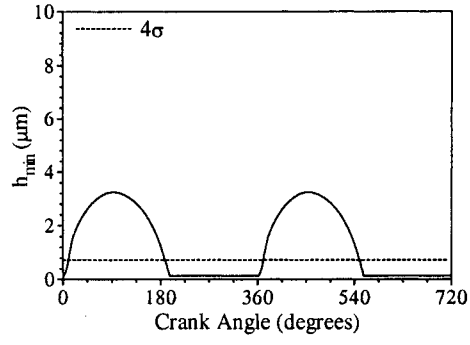
The tribological design of modern piston assemblies for automotive applications is greatly assisted by the use of sophisticated predictive mathematical models.

Numerical analyses for the piston and the ring pack are generally discrete models. Predictive analyses for the dynamics and lubrication of sliding interface between the piston skirt and the cylinder wall have been proposed by a number of research groups, for example [8, 11-13]. The primary outputs are piston secondary motion, tilting and lateral translation within the bore, and friction. Attempts have also been made to predict piston slap, as a precursor of piston derived noise [14].

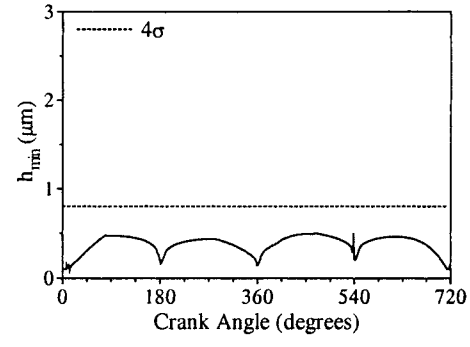
Piston ring lubrication models are numerous, for example [15-19], reflecting the ongoing problems encountered by industry with piston ring lubrication and associated phenomena. They provide vital information on lubricant film thicknesses, component dynamics, friction losses and lubricant transport paths. Figure 3 presents some exemplar film thickness results from such an analysis for the top compression ring and oil-control ring of a diesel engine [9]. The data is presented as the minimum lubricant film thickness between the piston ring and cylinder at any instant (h_{min}) plotted against crank angle where 0 degrees is top dead centre firing. Also shown is the upper limit of surface contact between the asperities on the two opposing surfaces, which is set equal to four times the composite root mean square surface roughness (σ) in accordance with the stochastic mixed lubrication and empirical boundary lubrication models incorporated in the software. This limit is the transition between the mixed and full fluid film lubrication regimes.



(a) Top compression ring, fully-flooded lubrication



(b) Oil-control ring, fully-flooded lubrication



(c) Top compression ring, starved lubrication

Figure 3: Predicted piston ring film thickness (h_{min}) for a diesel engine after 628 hours running at constant speed and load [9]
(4σ is the upper limit of surface contact)

Figures 3 (a) and (b) show the film thickness behaviour of the top compression ring and oil-control ring in isolation, assuming a plentiful supply of lubricant at the inlet, so called fully-flooded lubrication. In reality, however, the performances of adjacent rings are interdependent as the lubricant available at the inlet is to a great extent that left behind on the cylinder by the preceding ring. Figure 3(c) shows the effect this interaction has on the predicted film thickness for the top compression ring, drastically reducing the overall film thickness magnitudes throughout the engine cycle. As a consequence, this consideration of oil flow continuity within the ring pack is known as starved lubrication.

Such software tools for pistons and piston rings are in widespread use in the automotive industry and commercial software is available from a number of vendors. The analyses continue to become more sophisticated, building incrementally upon previous versions. For example, piston ring models have been enhanced to consider detailed lubricant rheological characteristics such as shear thinning [20]. Figure 4 highlights the significance of such behaviour.

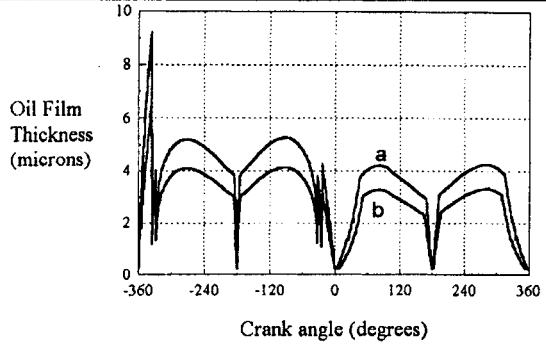


Figure 7: Predicted top compression ring oil film thicknesses including lubricant starvation effects (a) without shear thinning, (b) with shear thinning.

Figure 4: Influence of detailed lubricant rheology on piston ring performance, after Taylor *et al* [20]

The complex dynamics of piston rings within their grooves in the piston have also been considered [21]. Figure 5 illustrates the predicted influence on inter-ring gas pressures developed between adjacent piston rings, which strongly affects radial load, and the axial motion between the top ("up") and bottom ("down") of the groove. The correlation with experimental data also shown in Figure 5 is encouraging.

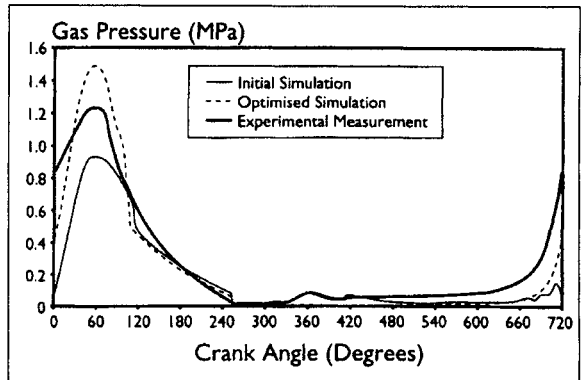


Fig 15.2 Measured and Predicted Gas Pressures Between Top and Second Rings in a 2.5 litre DI Diesel Engine at 1500 rev/min Full Load.

(a) Inter-ring gas pressures

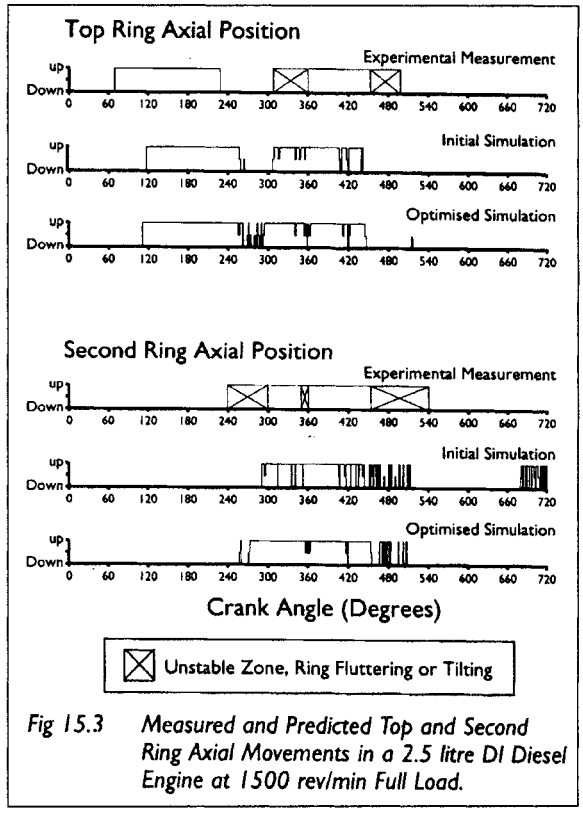


Fig 15.3 Measured and Predicted Top and Second Ring Axial Movements in a 2.5 litre DI Diesel Engine at 1500 rev/min Full Load.

(b) Axial ring motion

Figure 5: Consideration of axial ring dynamics in piston ring performance, after Mierbach *et al* [21]

The validated predictions of fluttering in Figure 5(b), an unstable axial reciprocation of the ring within the groove, are invaluable to engine designers as this phenomenon is associated with a loss of gas and oil flow control in the piston assembly.

As a final example of recent advances in piston ring modelling, the detailed of modelling of Tian [22] is exemplified by the predictions of variation in piston ring film thickness around the circumference of the piston of Figure 6.

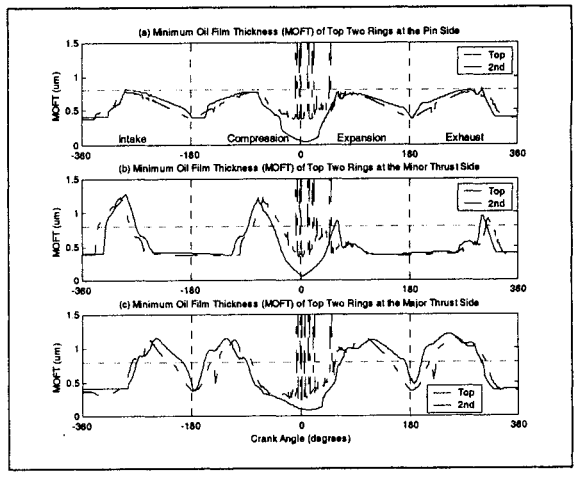


Figure 6: Predicted variation of piston ring film thickness around the circumference of the piston, after Tian [22]

3. APPLICATION TO THE REAL SYSTEM

However sophisticated such models become, their origins remain in mechanical dynamics and fluid film lubrication analyses. As such they can only perform realistic predictions when provided with component geometry and surface texture after running-in and the local lubricant performance characteristics at the engine running time of interest. Unfortunately such data is not readily available to the engine designer.

The main reason for this shortcoming is the complexity of the operating environment of the piston assembly as illustrated in Figure 7.

There are clearly a large number of inter-related mechanical, metallurgical, rheological, chemical and environmental parameters influencing the performance of piston rings in an internal combustion engine.

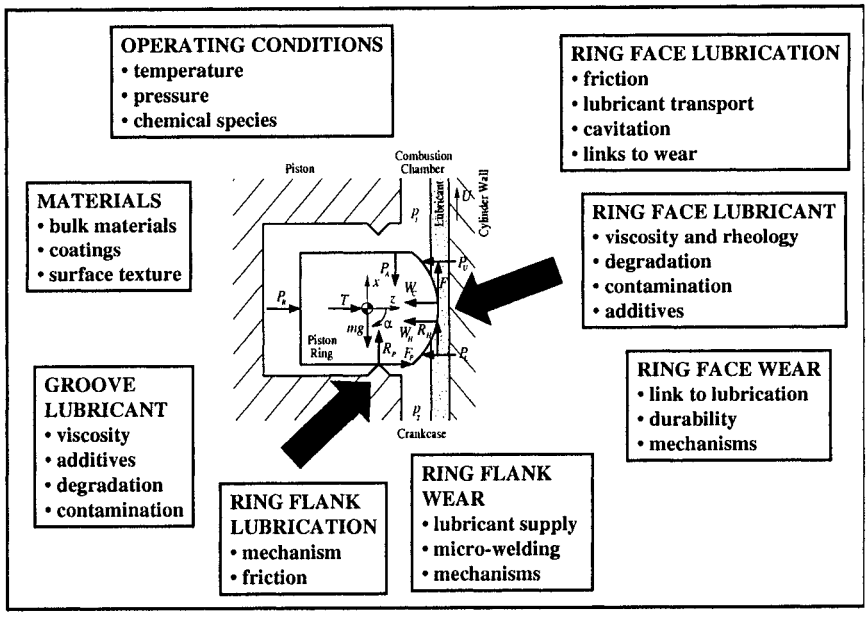


Figure 7: Complex operating environment of the piston ring

At the ring face in Figure 7 there are uncertainties surrounding the nature of the lubricant present in the interface and how this changes with running time. Similarly the complex interactions between lubrication and wear at this juncture are challenging to predict with any confidence.

At the ring flank there is even less capability to predict long term performance and occasional reported instances of piston material transfer to the ring flank in service, so-called "micro-welding", is indicative of the severity of the tribology at this interface [23, 24].

The highly dynamic operating conditions in the piston assembly and the large range of materials and surface treatments used for pistons, piston rings and cylinder bores, simply adds to the difficulty of optimising design using predictive software tools. As examples of ways to improve this situation, two complexities at the ring face are considered further.

3.1. Wear and Lubrication at the Piston Ring Face and Cylinder Bore Interface

Optimisation of piston assembly tribology is currently achieved through early-life wear of the tribological surfaces, the process of running-in. Most notably at the cylinder wall and the piston ring outer surface that slides against it, the piston ring face. This reveals not only the desire to minimise manufacturing costs but also the shortfall in our understanding of the tribology of this system.

The interactions between lubrication and wear during the running-in process are complex with significant changes in the

- Overall shape of the piston ring face
- Surface texture of the piston ring and the cylinder wall
- Surface texture of the piston skirt, to a lesser extent.

The modelling of wear at the piston ring / cylinder wall interface receives relatively sparse attention in the literature with just a handful of notable analyses [9, 25-29]. An example from the work of the current author [9] is given in Figure 8. The measured and predicted wear of the top compression ring from a diesel engine after 120 hours running at constant load and speed are compared. This is a barrel-faced, chromium-plated, cast iron ring. The measured new and worn ring profiles are shown overlaid using two different methods. Firstly by geometry, visually matching the unworn outlying regions of the profile, and secondly by mass, converting the weight loss of the ring to an evenly distributed volume loss around the ring circumference. The predicted ring profile, after 120 hours simulated running with multiple interactions between wear and lubrication models is also shown and this correlates well with the measured data. In terms of assessing the significance of this wear on performance, the new ring was predicted to operate for 41% of the engine cycle in the mixed or boundary lubrication regime whereas for the worn profile this fell to just 7% [9]. The friction power loss was also predicted to decrease, by some 22%.

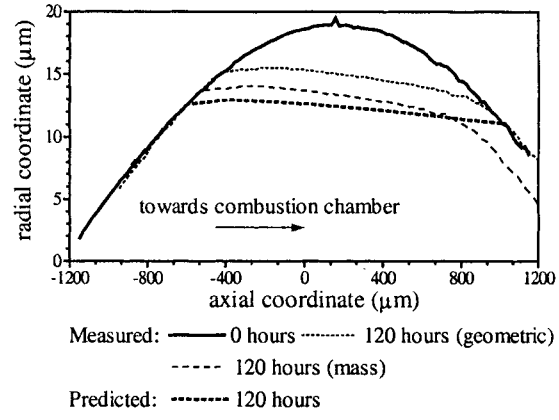
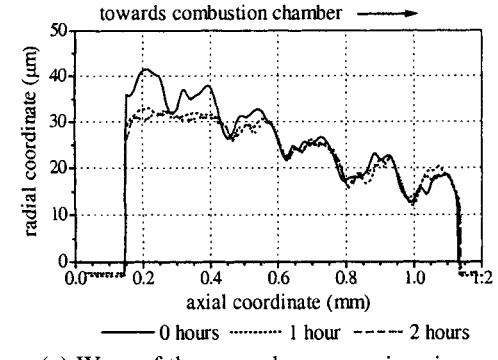


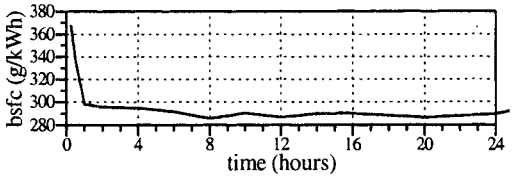
Figure 8: Predicted and measured top compression ring profiles, after Priest *et al* [9]

Such models offer the possibility of predicting piston assembly performance variation with running time and their development should continue. They are currently limited by over-reliance on empirical wear data from laboratory tribometers, interpretation of such data in the mixed lubrication regime and the inability to predict at the microscopic scale, for roughness as well as profile change.

To add further difficulties, many of the coatings and surface finishing processes used in engines have surface topographies, even after running-in, which are not modelled satisfactorily using standard fluid film and mixed lubrication models [30]. New strategies are therefore required to analyse the response of such surfaces. To illustrate this point, consider the rapid, early-life wear of the second compression ring measured in a modern gasoline engine, Figure 9(a). This is a plain cast iron, Napier scraper ring with compression and oil-control functions. The complex surface topography of this piston ring as new, with the fine turning marks deliberately left in place, and after 2 hours rapid wear is apparent. Figure 9(b) gives the measured brake specific fuel consumption (bsfc) of the engine for the early stages of running and shows a dramatic fall during the first hour.

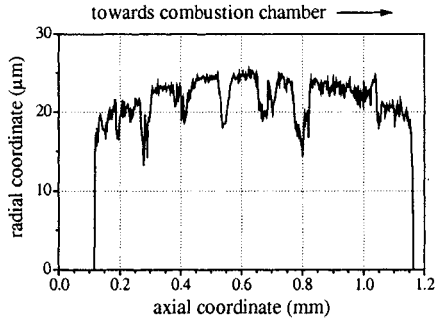


(a) Wear of the second compression ring

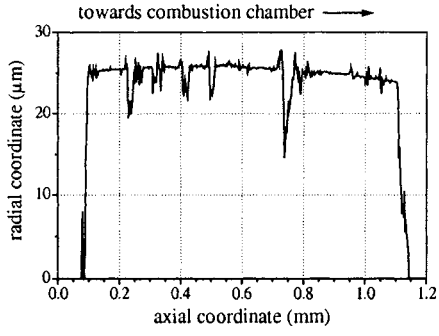


(b) Brake specific fuel consumption (bsfc)

Figure 9: Early-life measured wear and fuel economy of a gasoline engine, after Priest *et al* [30]



(a) Initial, 0 hours



(b) After 120 hours

Figure 10: Measured topographical changes in the top compression ring of a gasoline engine, after Priest *et al* [30]

Although, there are other mechanisms involved, such as changes in the lubricant, it is argued that this dramatic reduction in fuel consumption, and hence friction, is mainly attributable to the rapid wear of the second compression ring [30]. The distinctive topography of the surface of this ring is not readily analysed used typical mixed and boundary lubrication models, which are stochastic.

Similar problems are faced when considering the top compression ring of the same engine; a barrel-faced, spheroidal graphite cast iron ring with a flame-sprayed molybdenum coating, Figure 10. The open porous structure of the coating produces discrete features which persist even after 120 hours running [30].

3.2. Lubricant at the Piston Ring Face

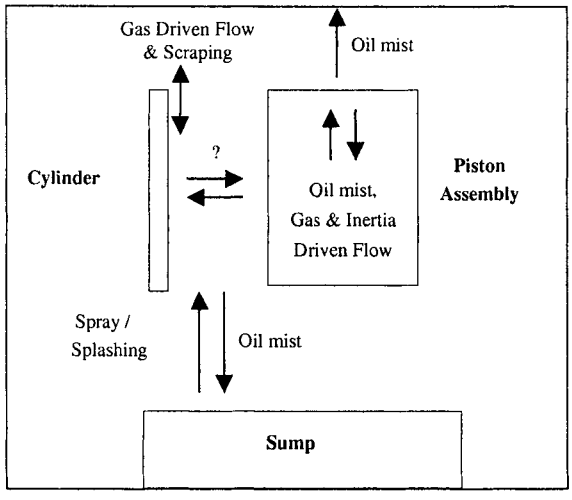
Much effort has been devoted to sensitising models of piston assembly tribology to bulk lubricant properties such as shear thinning and piezoviscosity. However, precise characterisation of the lubricant present in the piston assembly presents difficulties with respect to

- Knowledge of local temperatures
- Lubricant composition with the likelihood of fuel dilution, soot contamination and local thermal degradation
- Interaction of the lubricant with the surfaces, most notably anti-wear behaviour and boundary friction modification
- Background deterioration in the sump bulk lubricant properties with time.

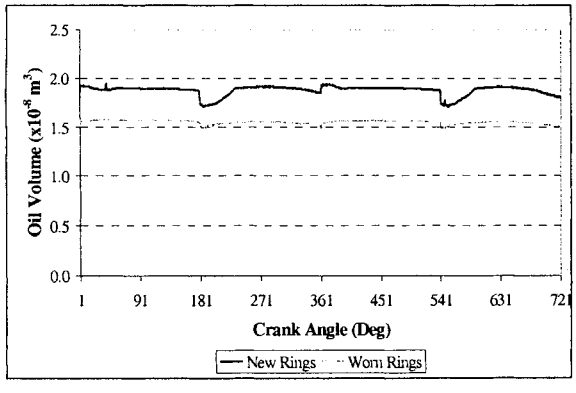
Degradation of the lubricant, especially around the top compression ring, has been a major cause for concern in recent years. Saville *et al* [31], Picken *et al* [32] and Taylor and Bell [33] all examined the chemical and physical changes in samples of lubricant drawn from the top ring zone of firing diesel engines. Clear differences were found in the degraded samples when compared to fresh lubricant, most notably a significant increase in viscosity. The degradation process is related to the residence time of the oil in the top ring groove. This in turn is dependent on the oil circulation mechanism in the piston assembly and hence the oil consumption and blow-by patterns [34].

Recent research associated with the current author has improved the prediction of lubricant transport within the piston assembly as a means to

predict lubricant residence times [35]. The quantification and interaction of all conceivable oil transport mechanisms in the piston assembly was considered, Figure 11(a). As shown in Figure 11(b), the model is capable of predicting differences in the volume of oil present in the piston assembly of a gasoline engine with new and worn piston rings. Comparison of the flow predictions with residence time experiments using a hydrocarbon marker and sampling of oil extracted from the top ring zone of the same gasoline engine has been undertaken [36]. The results overall are encouraging although a need for more research, especially into oil mist as a transport mechanism has been identified.



(a) Interaction of transport mechanisms



(b) Predicted variation of oil volume in the piston assembly

Figure 11: Detailed analysis of lubricant transport in the piston assembly of a gasoline engine, after Gamble *et al* [35]

4. OPTIMISATION

Given the complexity of the piston assembly system and particularly the important role of lubricant chemistry in controlling wear and lubricant rheology, empirical data from laboratory and engine tests will remain a key input to the design optimisation process in the medium term.

As has been demonstrated in this paper, sophisticated software tools to predict tribological performance of the piston assembly have been developed and continue to be enhanced. However, the demands from industry for new solutions to the problems posed by this complex system are growing more rapidly. There is a danger that as the models become ever more complicated and detailed, the developers and users lose perspective in terms of the overall aim of designing successful piston assemblies.

Therefore, in parallel, more systematic engineering design tools must be developed to capture the capability of and utilise the results from the new generations of detailed mathematical models. This will ensure rapid technology transfer from research to design.

5. CONCLUSIONS

- Validated mathematical models of piston assembly tribology have proved excellent aids to the designer for a tribological system that presents many design conflicts.
- The extremely complicated operating environment of the piston assembly presents major challenges when developing such predictive models.
- Input data and modelling problems exist with respect to the evolution with time of the component surfaces, form and roughness, and the lubricant, both physical and chemical behaviour.
- Research is progressing rapidly but industry demands are changing even more quickly. Future generations of engines will demand novel tribological solutions.
- To this end, alongside research, engineering tools to achieve design optimisation must be developed.

REFERENCES

1. Neale, M.J., "Drives and Seals, A Tribology Handbook", Butterworth-Heinemann, Oxford, 1994
2. Gazzard, S.T., Eastham, D.R., Jakobs, R.J. and Lunsford, R.L., "Piston System Design for Low Emissions", Leading Through Innovation, T&N Symposium 1995, Paper 20, 1995
3. Munro, R., "Emissions Impossible - The Piston & Ring Support System", SAE Paper 900590, 1990
4. Monaghan, M.L., "Engine Friction - A Change in Emphasis", Instn. Mech. Engrs., 2nd BP Tribology Lecture, 1987
5. Monaghan, M.L., "Putting Friction in its Place", 2nd Int. Conf.: Combustion Engines - Reduction of Friction and Wear, IMechE conf. pub. 1989-9, Paper C375/KN1, 1989, pp. 1-5
6. Andersson, B.S., "Company Perspectives in Vehicle Tribology - Volvo", 17th Leeds-Lyon Symposium on Tribology - Vehicle Tribology, Elsevier (Tribology Series 18), 1991, pp.503-506
7. Taylor, R.I. and Coy, R.C., "Improved Fuel Efficiency by Lubricant Design: A Review", Jour. Engineering Tribology, Proc. Instn. Mech. Engrs., Part J, 214, J1, 2000, pp.1-15
8. Chittenden, R.J. and Priest, M., "Analysis of the Piston Assembly, Bore Distortion and Future Developments", Chapter 10, Engine Tribology, ed. C.M. Taylor, Elsevier, Amsterdam, 1993, pp.241-270
9. Priest, M., Dowson, D. and Taylor, C.M., "Predictive Wear Modelling of Lubricated Piston Rings in a Diesel Engine", Wear, 231, 1, 1999, pp.89-101
10. Korcek, S., Sorab, J., Johnson, M.D. and Jensen, R.K., "Automotive Lubricants for the Next Millennium", Industrial Lubrication and Tribology, 52, 5, 2000, pp.209-220
11. Li, D.F., Rohde, S.M. and Ezzat, H.A., "An Automotive Piston Lubrication Model", ASLE Trans., 26, 2, 1982, pp. 151-160
12. Knoll, G.D. and Peeken, H.J., "Hydrodynamic Lubrication of Piston Skirts", J. Lubrication Tech., Trans. ASME, 104, 1982, pp. 504-509
13. Oh, K.P., Li, C.H. and Goenka, P.K., "Elastohydrodynamic Lubrication of Piston Skirts", J. Tribology, Trans. ASME, 109, 1987, pp. 356-362
14. Wong, V.W., *et al.*, "A Numerical Model of Piston Secondary Motion and Piston Slap in Partially Flooded Elastohydrodynamic Skirt Lubrication", SAE Paper 940696, 1994
15. Furuhama, S., "A Dynamic Theory of Piston Ring Lubrication (1st report, calculation)", Bull. JSME, 2, 7, 1959, pp. 423-428
16. Rohde, S.M., "A Mixed Friction Model for Dynamically Loaded Contacts with Application to Piston Ring Lubrication", Surface Roughness Effects in Hydrodynamic and Mixed Lubrication, ed. S.M. Rohde and H.S. Cheng, ASME pub., 1980, pp. 19-50
17. Ruddy, B.L., Dowson, D. and Economou, P.N., "A Review of Studies of Piston Ring Lubrication", Proc. 9th Leeds-Lyon Symp. on Tribology: Tribology of Reciprocating Engines, Paper V(i), 1982, pp. 109-121
18. Jeng, Y.R., "Theoretical Analysis of Piston-Ring Lubrication Part I - Fully Flooded Lubrication", Tribology Transactions, STLE, 35, 4, 1992, pp.696-706
19. Jeng, Y.R., "Theoretical Analysis of Piston-Ring Lubrication Part II - Starved Lubrication and its Application to a Complete Ring Pack", Tribology Transactions, STLE, 35, 4, 1992, pp.707-714
20. Taylor, R.I., Brown, M.A., Thompson, D.M. and Bell, J.C., "The Influence of Lubricant Rheology on Friction in the Piston Ring-Pack", SAE Paper 941981, 1994
21. Mierbach, A., Hildyard, M.L., Parker, D.A. and Xu, H., "Piston Ring Performance Modelling", Leading Through Innovation, T&N Symposium 1995, Paper 15, 1995, pp.15.1-15.13
22. Tian, T., "Dynamic Behaviours of Piston Rings and their Practical Impact. Part 2: Oil Transport, Friction and Wear of Ring/Liner Interface and the Effects of Piston Ring Dynamics", Jour. Engineering Tribology, Proc. Instn. Mech. Engrs., Part J, 216, J4, 2002, pp.229-247
23. Barrell, D.J.W., Priest, M. and Taylor, C.M. "The Axial Motion of the Piston Ring in the Top Ring Groove of a Gasoline Engine", Tribology Research: From Model Experiment to Industrial Problem, Proc. 27th Leeds-Lyon Symposium on Tribology, Elsevier, 2001, pp.641-649
24. Barrell, D.J.W., Priest, M. and Taylor, C.M. "Impact and Sliding Wear in the Top Piston Ring Groove of a Gasoline Engine", Proc. ITC Nagasaki 2000, International Tribology Conference, Japanese Society of Tribologists, Nagasaki, 2001, pp. 2029-2034

25. McCallion, H., Lloyd, T. and Large, J., "The Formation of Piston Ring Profiles", Tribology Convention 1972, Instn. Mech. Engrs. conf. pub., Paper No.C93/72, 1972, pp. 55-61
26. Ting, L.L. and Mayer Jr, J.E., "Piston Ring Lubrication and Cylinder Bore Wear Analysis, Part I - Theory", J. Lubrication Tech., Trans. ASME, Paper 73-LUB-27, 1974, pp. 305-313
27. Ting, L.L. and Mayer Jr, J.E., "Piston Ring Lubrication and Cylinder Bore Wear Analysis, Part II - Theory Verification", J. Lubrication Tech., Trans. ASME, Paper 73-LUB-25, 1974, pp. 258-266
28. Nautiyal, P.C., Singhal, S. and Sharma, J.P., "Analysis of Wear Mechanism of Piston Ring", International Conf. on Wear of Materials: Wear of Materials 1979, ASME, April 16-18, 1979, pp. 124-131
29. Hannoschoeck, N., "Zur Tribologie des Kolbenring", MTZ Motortechnische Zeitschrift, December, 46, 12, 1985, pp. 503-506, 509-511
30. Priest, M. and Taylor, C.M., "Automobile Engine Tribology - Approaching the Surface", Wear, 241, No.2, 2000, pp.193-203
31. Saville, S.B., *et al.*, "A Study of Lubricant Condition in the Piston Ring Zone of Single-Cylinder Diesel Engines under Typical Operating Conditions", SAE Paper 881586, 1988
32. Picken, D.J., *et al.*, "Lubricating Oil - Extraction and Chemical Analysis", Experimental Methods in Engine Research and Development '91, I.Mech.E Seminar, HQ, December, 1991, pp.115-120
33. Taylor, R.I. and Bell, J.C., "The Influence of Lubricant Degradation on Friction in the Piston Ring Pack", Dissipative Processes in Tribology, Proc. 20th Leeds-Lyon Symposium on Tribology, 1993, pp.531-536
34. Burnett, P.J., "Relationship between Oil Consumption, Deposit Formation and Piston Ring Motion for Single Cylinder Diesel Engines", SAE Paper 920089, 1992
35. Gamble, R.J., Priest, M. and Taylor, C.M., "Detailed Analysis of Oil Transport in the Piston Assembly of a Gasoline Engine", Tribology Letters, *In Press*
36. Gamble, R.J., "The Influence of Lubricant Degradation on Piston Assembly Tribology", PhD thesis, The University of Leeds, UK, 2002

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Optimised Additive-Surface Material Combinations for Reduced Friction & Wear in Internal Combustion Engines

M. Voong[a], A. Neville[a], R. Castle[b]

[a] Corrosion and Surface Engineering Research Group, School of Engineering and Physical Sciences, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, Scotland

[b] Infineum UK Ltd, Milton Hill Technology Centre, Abingdon, OX13 6BB, England

There has been a general increase in the level of CO_2, NO_x and particulate emissions through the consumption of hydrocarbon fuels over the last 2-3 decades. These emissions have a detrimental effect on the environment and, as such, are undesirable. In an attempt to reduce emissions and to increase fuel efficiency of internal combustion engines there have been developments in two key areas: lubricant technology and materials technology. Improvements in lubricant formulation and/or improvements in the material technology performance of internal combustion engine components can potentially lead to an overall reduction in friction and wear. Because of the increased use of new materials and surface engineering in engines it is necessary to understand their lubrication requirements, which are anticipated to be different to those of conventional Fe-based materials. The work presented in this paper addresses aspects of lubrication of a traditional Cr-bearing steel and Al-Si alloy-based components. In particular, the interaction of surfaces with lubricant additives is investigated.

Tribological tests have been performed and are supported by post-test examination using X-ray Photoelectron Spectroscopy (XPS), and Environmental Scanning Electron Microscope (ESEM) with Energy Dispersive X-ray analysis (EDX). Preliminary results show that fully formulated lubricant oils are effective in the reduction of wear and friction for ferrous-based systems but not for aluminium systems. Initial XPS and EDX data suggests that a relatively thick anti-wear film is formed on ferrous materials but is thin (and therefore unstable) for steel on aluminium systems. In completely non-ferrous systems the wear film is absent. In this paper the fundamental aspects of the film formation as well as the practical aspects of the results will be discussed.

1. INTRODUCTION

In an attempt to halt the rise in CO_2 emissions recorded since 1979 [1] and to improve fuel economy there have been developments made in two key areas: lubricant technology and materials technology.

Lubricant technologies for ferrous materials are well documented [2-9]. In particular, the study of anti-wear additives to reduce the amount of wear of the material surface has been investigated. Studies have shown the formation of a relatively thick, stable anti-wear film when lubricated in formulations containing anti-wear P-based additives [2, 3]. The anti-wear film formed at the material surface, as a result of decomposition of anti-wear additives (typically zinc dialkyl dithiophosphate, ZDDP) acts to protect the bare surface against wear. The effectiveness of ZDDP as an anti-wear additive is strongly dependant on the test conditions. As shown by So *et al.* [2], under low loads the wear performance was improved by ZDDP. However, at higher loads, for steel on steel contacts, the wear rate can actually be increased through ZDDP additives [3].

There has been increased interest in the structure and nature of these films, since a better understanding of the formation process will inevitably lead to improvements in lubricant additive technology. Studies have been carried out to investigate these tribofilms and the studies of anti-wear film formation from simple ZDDP systems by Tonck *et al.* [4] have shown that ferrous substrates are covered with a bi-layered film consisting of a

thin sulphide/oxide base layer covered with a thicker phosphate layer. Conventional methods involve the use of solvents to wash the excess oil film off (Bec *et al.* [5]). However it was shown that using solvent not only removes the oil solution but, in addition, the alkyl phosphate precipitates. These are therefore not usually detected by conventional surface analysis.

Recent work conducted by Martin [6] has shown that zinc polyphosphates are important intermediate reaction products in the tribochemical reaction of ZDDP. ZDDP decomposition is thought to be through complex chemistry which can be either thermal, catalytic or oxidative. In this case Martin believes that ZDDP undergoes mainly thermo-oxidative decomposition prior to friction, when the lubricant is heated above 100°C.

The interactive effect of other additives on the behaviour and performance of ZDDP have been extensively investigated. A study conducted by Martin *et al.* [7] investigated the tribochemical interactions between ZDDP and *p*oly*i*so*b*utene *s*uccin*i*mide (PIBSI) dispersant in the mild wear regime, by the dual analysis approach of TEM and XPS. Traditionally an antagonistic interaction occurs between ZDDP and dispersant additives and this can be either through chemical interaction in the bulk lubricant or competition to adsorption. The binary additive system was compared directly to a simple system consisting of only base oil and ZDDP. The results from this study suggests the chemisorption of ZDDP on steel at 60°C was hindered by the presence of the PIBSI and this in turn interrupts the adhesion and abrasion protecting performance of ZDDP. Other work carried out [8, 9] shows a similar effect as that shown by Martin. It has been shown that the addition of a dispersant or detergent resulted in an improvement of the overall detergency of the lubricant but the anti-wear performance was reduced. This is most likely due to the interference of the formation of metaphosphate and it has been shown by several authors (Willermet *et al* [10] and Reyes and Neville [11]) that displacement of Zn by Ca when an overbased detergent is used can occur within the anti-wear film.

There has been an increased use of Al-Si alloys in internal combustion engine components such as pistons, since they exhibit excellent resistance to corrosion, good thermal conductivity, moderate costs, and ease of fabrication and machining. Despite these characteristics, which meet the requirements for such an application, they are vulnerable to the sliding process because of the oxide film formed resulting in high friction and wear.

The tribological performance of Al-Si alloy and in particular its resistance to seizure [12] can be improved in two ways. Firstly, improved lubrication of Al-Si alloy sliding against itself and other couples could be accomplished. Secondly, improvement in the tribological performance of the material itself can be achieved by reinforcement with ceramic phases such as SiC and Al_2O_3 in the form of fibres, whiskers and particles [13]. It has also been shown that the friction and wear properties of Al-Si alloys can be improved through embedding of solid lubricants such as MoS_2 into the aluminium oxide pores [14].

The tribological behaviour of Al-Si alloys under dry sliding has been extensively studied but there is little published work on the lubrication of these materials and in particular surface/lubricant additive interactions. Preliminary research on Al-Si alloys against steel [15, 16] has shown that commercially available engine oils are not as effective in lubricating this system as they are on steel against steel contacts. In boundary lubrication of aluminium, the chemical interaction, especially the chemical reaction of the lubricant with the rubbing surface is very important. Tests conducted by Wan *et al.* [15, 16] and Hah *et al.* [17] indicate that conventional P and S based anti-wear additives are not very effective in preventing seizure of the aluminium.

In this paper, the friction and wear characteristics of traditional ferrous materials and aluminium alloys lubricated with formulated crankcase lubricants have been evaluated and a link between the tribological performance and the nature of the tribochemical wear film is investigated. The tribological properties of different material couples with fully formulated lubricants were investigated using a reciprocating pin-on-plate apparatus. The behaviour of ferrous and non-ferrous material systems were compared to results in base oil free from additives.

2. EXPERIMENTAL DETAILS

This study has focused on material couples as shown in Table 1. Tribological properties of these materials in lubricating oils were examined using a reciprocating pin-on-plate system. The pin was loaded using a static load applied through a lever

Table 1
Material couples tested

Material Configuration	Pin Material	Plate Material
Ferrous System	BS EN 1452 Cast Iron	BS EN 31 Cr-bearing Steel
Intermediate System (1)	Al-Si Piston Alloy	BS EN 31 Cr-bearing Steel
Intermediate System (2)	BS EN 1452 Cast Iron	Al-Si Piston Alloy
Non-Ferrous System	Al-Si Piston Alloy	Al-Si Piston Alloy

arm and the plate was reciprocated underneath through a crank mechanism.

The pins used were 20mm in length, 6mm in diameter and the ends of the pins were machined to a 20mm radius of curvature. The rectangular plate measured 17 × 6 × 3mm^3. The test parameters are shown in Table 2. A bi-directional load cell measures the friction force on the static pin during the course of the experiment. The friction force was recorded every half-hour for 3 cycles.

Table 2
Test parameters and conditions

Test Parameters	Test Conditions
Test Stroke	10mm
Test Frequency	1 Hz
Load Applied	185 N
Corresponding Hertzian Pressure	0.5-0.9 GPa (depending on material couple)
Test Temperature	60°C and 100°C
Test Duration	8 hours

Using the friction force data, the friction coefficient (μ) was calculated by first calculating the average absolute friction force of the trace and then dividing this by the applied load to give μ. This is then plotted as a function of time over an 8-hour period. Three seconds of data was recorded which plots the friction force versus time as a square wave as shown schematically in Figure 1.

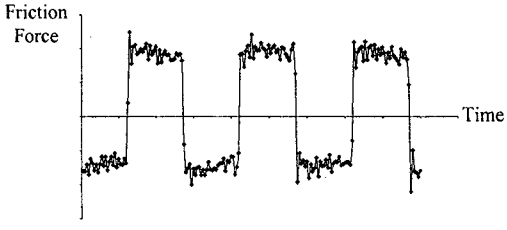


Figure 1. Square wave trace of friction force versus time

The typical friction coefficient values of 0.07 - 0.18 from preliminary results suggests that the contact regime is in the boundary lubrication regime and this is verified from the square wave traces. The square waves show the friction force (and thus the friction coefficient) is not sensitive to the variation in linear speed over the stroke. This is typical of boundary lubrication. The boundary lubrication regime is further verified by the calculation of the lambda ratio, Λ which are defined in equations 1 and 2. The lambda ratio calculations (based on the starting values of R_q) for all the materials is below unity ($0.055<\Lambda<0.176$).

$$\Lambda = \frac{h_{\min}}{\sqrt{R_{q1}^2 + R_{q2}^2}} \quad (1)$$

where R_{q1} and R_{q2} are the respective RMS surface roughness of the two solid surfaces and h_{min} is the minimum lubricant film thickness for EHD lubrication

$$h_{\min} = 3.63 \times R \times U_{\Sigma}^{0.68} \times G_{\Sigma}^{0.49} \times W_{\Sigma}^{-0.073} (1 - e^{-0.68 k_e}) \quad (2)$$

where $U_{\Sigma} = \frac{\eta_0 u_s}{E^* R}$; $G_{\Sigma} = \alpha_p E^*$; $W_{\Sigma} = \frac{P}{E^* R^2}$ and R is the radius of the pin (20mm), η_0 is the viscosity, u_s is the sliding velocity, α_p is the viscosity-pressure coefficient (2×10^{-8} Pa^{-1}), P is the normal load (185N), E^* is the effective modulus of elasticity and k_e is the elliptical parameter which is equal to 1.0339 for a sphere on flat type of contact [18]

At the start of each experiment, the pin and plate were cleaned using acetone in an ultrasonic bath to degrease and remove any impurities from the surface. They were then weighed using a Mettler AJ150 micro-balance with an accuracy of ± 0.1mg. Lubricating oils supplied by Infineum UK Ltd. were used (2.5ml of oil was used) in each test.

After each test the pins and plates were dipped in heptane for 1-2 seconds to remove excess oil from the surface and then weighed to calculate the volume

Table 3
Properties and composition of materials tested

Material	C (wt%)	O (wt%)	Si (wt%)	Cr (wt%)	Mo (wt%)	Mn (wt%)	Cu (wt%)	Al (wt%)	Mg (wt%)	Fe (wt%)	Hardness (Hv)	Density (g/cm^3)
BS EN31	0.64	-	0.55	1.57	-	0.49	-	-	-	96.75	185	7.75
Al-Si Alloy	4.44	2.29	22.62	-		-	1.12	68.25	1.27	-	120	2.62
BS EN1452	9.32	1.91	2.74	-	0.44	0.71	-	-	-	84.88	150	7.01

of material loss from the plate, pin and the pin/plate system. The plates were then analysed using optical microscopy, Environmental Scanning Electron Microscope (ESEM) with EDX attachment for chemical analysis and X-ray Photoelectron Spectroscopy (XPS). The Environmental Scanning Electron Microscope at the Centre of Environmental Scanning Electron Microscopy (CESEM), (Department of Petroleum Engineering-Heriot-Watt University) was used for SEM analysis. For this study, the ESCA300 photoelectron spectrometer situated at the Research Unit for Surfaces, Transforms and Interfaces (RUSTI) facility at Daresbury Laboratory was used for XPS analysis.

3. MATERIAL AND LUBRICANTS CHARACTERISTICS

This study focuses on the combined effects of material couple and the nature of the lubricant additives. Current material technologies used in internal combustion engines include aluminium alloys in conjunction with traditional ferrous materials. However, there is limited study and hence understanding of the lubrication of these systems. The lubricants tested are a base-oil and a fully formulated lubricant, which is designed to meet proposed changes in environmental legislation. The materials tested are current ferrous and non-ferrous materials. The materials and lubricants investigated in this study are described in sections 3.1 and 3.2.

3.1. Material

In this paper, current representative engine materials are tested. The materials tested include; BS EN31 Cr-bearing steel, BS EN1452 cast iron and an Al-Si alloy. The Al-Si alloy was supplied as cast pistons from Federal Mogul Ltd.

The Al-Si alloy shows distinct precipitates (<20μm), characteristic of Al-Si alloys of eutectic composition. The cast iron contain graphite flakes typical of grey cast iron, while no discernable microstructural features are present for the Cr-bearing steel. Table 3 shows the composition and properties of the materials.

3.2. Lubricants

A fully formulated lubricant (FF) and a base oil were used for the tests. The FF lubricant is typical of those used in Europe and the base oil used is the same as that used in the FF oil. It is a Group III synthetic base oil. The group designation relates to the amount of refining of the crude oil, Group I relates to mildly refined through to Group IV which is fully synthetic. The base oil has a kinematic viscosity of 4 cSt at 100°C.

4. RESULTS

In this section the results from tribological tests and post-test analysis using SEM and XPS are shown. These are then discussed in greater detail in section 5.

4.1. Friction Response

The experiments are separated into 4 different material couples as was described in Table 1. From friction and wear data, the results can be separated into two distinct categories, depending on the plate material: either EN31 chromium bearing steel or Al-Si alloy.

Figures 2 and 3 show the friction traces for the steel plate and the Al-Si base plate respectively in the two oils, with the 2 counter-face materials (a-cast iron pin and b-Al-Si alloy pin) and the 2 test temperatures. These will be described separately in the next sections and then the different response of the materials will be discussed.

4.1.1 EN31 Chromium Bearing Steel Plates

Figure 2 shows the influence of temperature, lubricant and pin material on the friction coefficient using EN31 plates. The results show there to be a steady trend for friction coefficient after an initial running-in period for all experiments conducted using the FF oil (i.e. there is little variation in μ over time). However, tests conducted using base oil show higher friction coefficient values and an increasing friction coefficient trend over time. There seems to be no effect of temperature on the friction coefficient for the Al-Si pins with the FF oil, but for the fully ferrous system with the FF oil there is a large reduction in friction coefficient with the increase in temperature.

In summary the main points relating to the friction performance using EN31 plates are:

- Results using the fully formulated oil show a steady trend for μ after an initial running-in period of about 3.5 hours
- Tests conducted using base oil show increasing μ over the test period
- Cast-iron on EN31 in the fully formulated lubricant shows large reduction in μ with increase in temperature

4.1.2 Al-Si Piston Alloy Plates

The friction coefficient results for Al-Si alloy plates are shown in Figure 3 where the effect of temperature, lubricant type and pin material are evident. The results show that for tests conducted using the FF oil, there is a steadily increasing trend for the test duration while tests conducted using base oil show a steady trend. For entirely non-ferrous systems (Al-Si alloy pins and plates) there is no effect of temperature on the friction coefficient while systems using cast iron pins show an increase in friction coefficient with temperature increase.

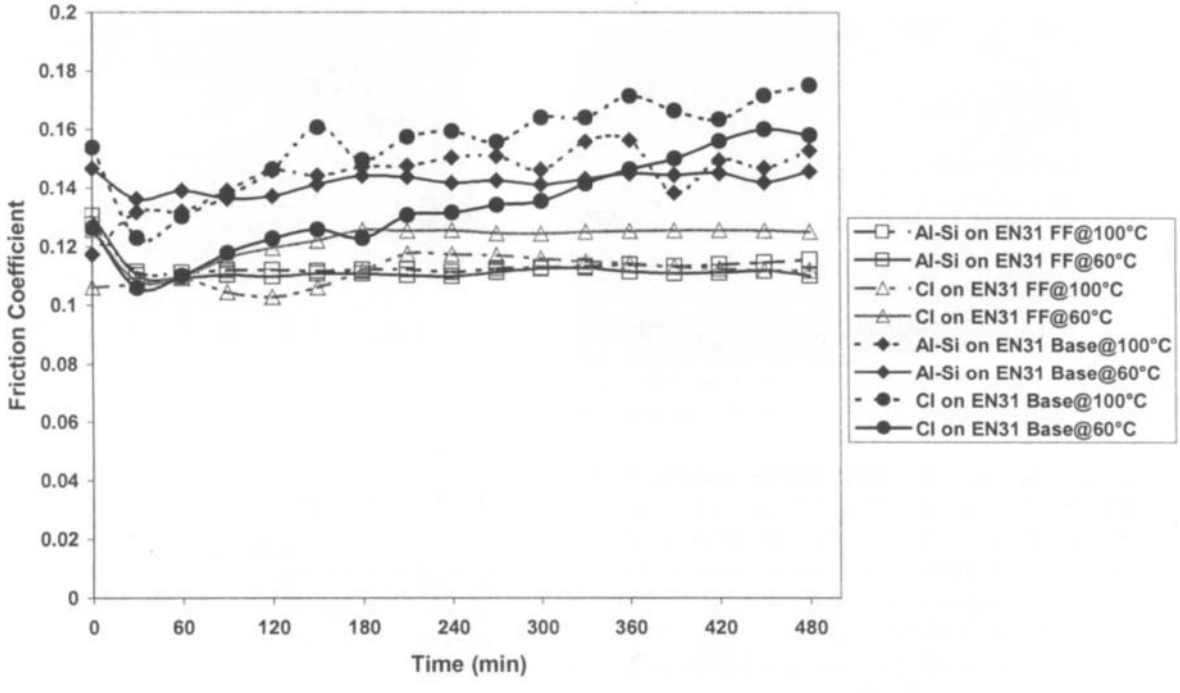


Figure 2. Friction coefficient versus time for EN31 plate systems

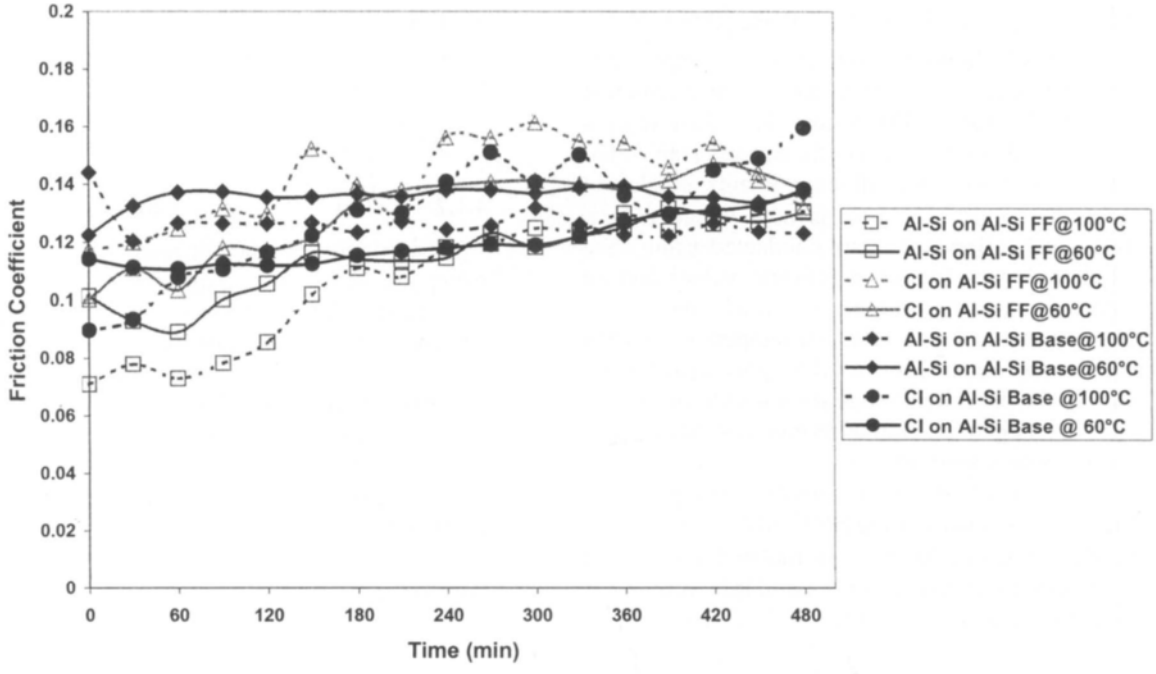
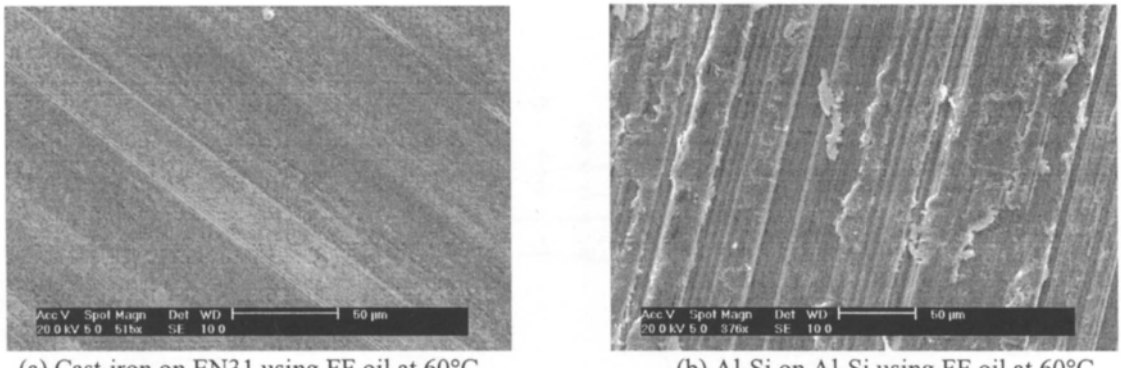


Figure 3. Friction coefficient versus time for Al-Si alloy plate systems



(a) Cast-iron on EN31 using FF oil at 60°C (b) Al-Si on Al-Si using FF oil at 60°C

Figure 4. SEM images of wear scar comparing ferrous and non-ferrous material couples

In summary the main points relating to the friction on the Al-Si alloy plates are:

- All results using the FF oil show a steadily increasing trend for μ up to around 5 hours which suggest a longer running-in period than for systems with EN31 plates
- Cast-iron on Al-Si shows an increase in μ with increasing temperature
- Al-Si on Al-Si using the FF oil shows no effect of temperature

4.2. Wear Results

SEM micrographs taken from the wear track compares the ferrous and non-ferrous systems as shown in Figure 4. A clear difference is evident after the test period of 8 hours at 60°C. The ferrous system remains relatively smooth with surface features at high magnification resembling a granular structure but with no gross deformation whereas on the non-ferrous system visible wear can be seen on the plate surfaces illustrating the lack of protection

from the lubricant additive. The wear coefficient (equation 3) results in Figure 5 show some interesting trends when comparing the same systems (lubricant and material couple) at different temperatures and using different lubricants. Fully ferrous systems using FF oil show that the higher temperature of 100°C has a beneficial effect on the wear characteristics of the system. However, with the Al-Si alloy plate systems using the FF oil, the higher temperature accelerates the wear rate of the system. As expected, tests conducted using base oil lubricant on EN31 plate systems show higher wear than systems using the FF oil. In contrast, systems using Al-Si alloy plates show less wear using base oil compared to tests conducted with the FF oil. It would appear that the overriding factor affecting wear for Al-Si alloy on Al-Si alloy systems is adhesive wear between the pin and the plate and there is little or no protection afforded by any wear film formed.

$$K = \frac{VolumeLoss \times Hardness}{Load \times Sliding\ Distance} \qquad (3)$$

4.3. EDX Analysis

Analyses carried out using EDX on wear plate samples tested at 60°C are shown in Table 4. The results show that tests conducted in the FF oil using EN31 plates (with either pin material) show evidence of an adhered tribofilm as demonstrated by the detection of elements such as S, P, Zn, Mo and Ca. However, analysis of the Al-Si alloy plate systems shows no evidence of a tribofilm although there is evidence of material transfer from the cast iron pin (Fe peak). This suggests that tribofilms are either not formed at all on the surface of Al-Si or that this film is very much thinner since the EDX has a penetration depth of a few microns. Therefore a very thin layer of a few nm may mean that the elements are present beyond the detection limit. XPS analysis was therefore carried out to probe the near surface region of the wear scar. EDX analysis also shows more evidence of Zn on the wear scar when cast iron was the counterface material, against EN31 than for Al-Si against EN31.

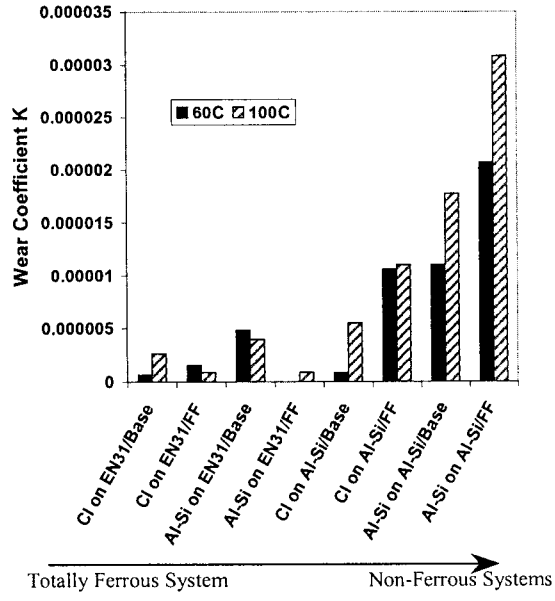


Figure 5. Wear Coefficient (K) from tests

4.4. XPS Analysis

The samples for XPS analyses (Table 5) were chosen to enable direct comparisons of the effect of (a) temperature, but, most importantly, (b) the material combination. Two types of XPS analyses were performed on the samples. An initial survey scan was carried out over the full range of Binding Energy (BE) values (0 – 1150eV). This was to identify the different species (elemental) present on the wear scar, such as C, O, Zn, P, S, Mo and Ca. A more comprehensive analysis involving multiple scans was then carried out on the energy regions corresponding to the individual elements to determine the quantity, structure and form of the elements. Full quantitative region scans were conducted after various etching periods to enable variations as a function of depth to be determined.

Table 4
EDX analysis carried out on plates

Test Specimen	C(at%)	O(at%)	S(at%)	P(at%)	Fe(at%)	Zn(at%)	Ca(at%)	Mo(at%)	Al(at%)	Si(at%)
CI/Al-Si 60°C FF	14.08	0.32	-	-	0.13	-	-	-	63.99	19.78
CI/EN31 60°C FF	22.94	11.69	0.48	0.28	60.27	0.67	0.47	0.16	-	0.49
Al-Si/Al-Si 60°C FF	9.44	3.66	-	-	-	-	-	-	64.56	20.56
Al-Si/EN31 60°C FF	15.46	16.38	0.77	0.49	61.28	0.51	0.65	0.17	0.86	0.67

Table 5
XPS analysis carried out on plates

Etch time	Experiment	C (at%)	O (at%)	S (at%)	P (at%)	Fe (at%)	Zn (at%)	Ca (at%)	Mo (at%)	Al (at%)	Si (at%)
45 min	1. Al-Si/Al-Si FF@100°C	67.4	11.3	-	-	-	-	-	-	18.5	2.7
60 min		60.2	15.1	-	-	-	-	-	-	22.2	2.5
45 min	2. CI/Al-Si FF@100°C	58.3	17.8	1.7	4.5	n/a	0.4	4.9	0.7	9.2	2.4
60 min		52.0	20.1	2.4	4.1	4.1	0.3	4.1	0.6	9.6	2.8
30 sec	3. CI/Al-Si FF@60°C	38.8	26.0	2.3	3.2	2.6	0.3	4.5	0.6	17.9	3.8
4 min		17.3	33.9	1.5	4.5	3.7	0.2	4.7	0.3	28.6	5.4
30 min		11.7	19.3	0.9	9.4	2.8	-	2.3	-	40.2	13.5
60 min		8.7	14.2	10.7	2.4	1.3	-	1.3	-	44.3	18.5
30 sec	4. Al-Si/EN31 FF@100°C	18.9	42.6	2.2	4.6	6.2	0.4	12.7	0.8	8.6	2.9
4 min		15.4	36.3	1.9	4.9	15.5	0.4	13.3	0.8	8.1	3.3
30 min		14.9	29.3	2.3	4.0	20.8	0.4	11.4	0.6	13.9	2.4
60 min		15.4	24.0	2.7	4.2	30.6	0.5	7.1	0.5	11.0	2.8
30 sec	5. CI/EN31 FF@100°C	32.5	38.3	3.6	5.8	4.6	0.9	13.3	1.1	-	-
4 min		18.6	33.7	3.2	6.7	19.7	0.8	13.7	0.8	-	2.7
30 min		19.2	28.7	3.8	4.5	30.5	0.5	8.6	1.0	-	1.6
60 min		20.2	19.1	2.7	2.8	44.2	0.7	4.2	0.9	-	2.8

4.4.1 Survey Scans

The survey scans are taken after 30-second Argon ion-etching and were conducted between 0 and 1150eV binding energy. A preliminary analysis of the survey scans shows that for a non-ferrous system (Al-Si alloy on Al-Si alloy), no traces of film elements such as P, S, Zn, Ca and Mo were detected. From the analysis of the peak position of aluminium and oxygen, the surface of the material was predominantly in the form of an oxide and hydroxide (Al_2O_3 74.3eV and $Al(OH)_3$ 74.4eV) evident from the broad peak around 74-75eV. However, for systems using cast iron pins traces of elements such as P, S, Mo, Zn and Ca as well as material transferred from the pin (Fe and C) were observed. As expected, systems using the steel plates (EN31) show large peaks of P, S, Mo, Zn and Ca (relative to the Al-Si alloy plate systems). However, to understand better the behaviour of the wear film on the surface of the materials, a more detailed long scan of the relevant regions was carried out.

4.4.2 Region Scans

For each analysis on tests 1-5 the Binding Energy range for elements in Table 5 was scanned.

4.4.2.1 Comparing the Effect of Material Couple

As with both friction and wear results, the results from XPS (in terms of material couple) can be distinguished by the plate material used (EN31 and Al-Si Alloy). Tests 1 and 2 compare the effect of pin material (cast iron and Al-Si alloy) on Al-Si alloy using the FF oil at 100°C. For the non-ferrous system tested (Al-Si on Al-Si) there is no evidence of S, Mo or Zn on the wear scar surface while trace amounts of P are evident and the P 2p position corresponds to that of phosphate (134eV). These are not large enough to facilitate quantification. For the system containing iron (cast iron on Al-Si alloy), the XPS results show evidence of a wear film in the form of phosphate (believed to be polyphosphate [4]-133eV) and sulphides (MoS_2-162.2eV and FeS_2-162.6eV for sulphur peaks) evident from the broad S peak around 162-163eV. Interestingly as the samples are etched the form of S is changed for the Al-Si alloy plate samples compared to that of steel samples and the peak position for S changes from the MoS_2 position (162.2ev) to the S 2p position of 165eV. In contrast, for the EN31 plate system there is no change in peak position for S as ion etching occurs.

Tests 4 and 5 compare the difference between pin materials on the EN31 plate using the FF oil at 100°C. From the results there is more evidence of elements such as S, P, Zn, Mo and Ca on the surface than systems using Al-Si alloy plates. From the traces taken from the survey scans, the only difference between different pin materials on EN31 plate is the transfer of material from the pin to the plate (evident from the aluminium peak). Scans at the BE for S, P and Mo show more defined peaks for the EN31 plates in comparison to systems using Al-Si alloy plates. Compositional analysis taken from the region scans shows a relatively large quantity of S, P and Ca. Therefore the wear film on the surface of the material is made up of sulphides, phosphates and carbonates ($CaCO_3$-347.1eV). Further etching of the sample still yields high amounts of these elements and therefore it can be assumed that the films formed on EN31 are much thicker than those formed on Al-Si alloy.

4.4.2.2 Comparing the Effect of Temperature

Tests 3 compared with test 2 enable the effect of temperature on the role of film formation from the FF oil to be compared. It is evident from the scans of the S 2p peak that there is a larger S peak for the higher temperature experiment. The large peak around 162eV corresponds to MoS_2 (162.2eV) and FeS_2 (162.6eV) while pure S should be evident at 164eV. There appears to be little difference in the phosphorous trace between 60°C and 100°C. Both traces show the same trend with a large peak evident around 134eV which is indicative of polyphosphate on the surface of the wear scar. Scans for molybdenum show only trace amounts (less than 1at%) at both the temperatures tested although there appears to be a larger peak for the 100°C-test specimen. This peak at 226eV corresponds to MoS_2 which confirms the sulphur trace. From these regional traces it is indicated that MoS_2 is more readily formed at the higher temperature of 100°C as compared to 60°C but that these traces are small in comparison to phosphate which is formed at both 60°C and 100°C. For both the 60°C and 100°C tests, there appears to be little change in the form of P as the sample is etched. The position for P remains around 134eV as the sample is ion etched. However in the case for sulphur, as the etching time increases the form of S changes from MoS_2 to the S2p form (around 165eV). This would suggest that some surface mechanism is required to instigate the formation of molybdenum disulphide since this is the layer on the uppermost of the anti-wear film and it is most likely to be sheared during the tribological tests.

5. DISCUSSION

Studies in the literature often focus on simple lubricant systems and do not take into account the complicated interaction of additives found in FF oils. In particular, proposed changes in environmental legislations will impact the type and amount of additives used in crankcase lubricants and this in turn will affect the tribological behaviour of these systems. This study goes some way to address these issues. The objective of this work was to investigate the behaviour of a formulated lubricant with four material couples. Measured friction and wear data shows that the performance is governed most significantly by the plate material, in this case either EN31 Cr-bearing steel or Al-Si alloy.

5.1. EN31 Plates

5.1.1 Tribological Analysis

Tribological tests conducted on EN31 plates show that the formulated oil reduces friction and wear relative to base oil. A comparison of the friction and the wear results at different test temperatures has shown that increasing temperature from 60°C to 100°C has a beneficial effect by reducing both friction and wear. In addition higher concentrations of surface elements were observed at higher temperatures suggesting that film formation is more effective at elevated temperatures. In some cases [3, 4] anti-wear additives have been shown to reduce wear but increase traction due to the presence of the thick surface films. Here the presence of a friction modifier offsets this. In this study it has been confirmed that the friction modifier (Mo-based) promotes the formation of molybdenum disulphide and the increase in temperature accelerates the promotion of MoS_2.

5.1.2 Post-Test Analysis (EDX and XPS)

The results from this paper confirm the formation of a relatively thick anti-wear film on the steel plate. This is evident from post-test analysis conducted by EDX (which has a penetration depth of several microns) where elements such as S, P, Zn,

Mo and Ca are detected on the wear scar surface. SEM micrographs show the smooth topographical nature of the wear scar and the lack of any gross deformation of the surface. This indicates a protective layer on the surface of the material between the contacting areas.

XPS results show the formation of phosphates, sulphides and oxides on the surface. These observations are similar to those made by Tonck *et al.* [5], whose studied anti-wear film formation from simple ZDDP systems. In his models, the ferrous substrates were covered with a thin sulphide/oxide layer. This in turn was covered with a thicker phosphate layer. In this paper the model of the film formation (determined from XPS depth profiling), as shown in Figure 6a, shows similar sulphides, oxides and phosphates.

However, in addition to these layers XPS peak position analysis also suggested evidence of additive interaction on the ferrous surface, from the presence of calcium carbonate ($CaCO_3$) on the surface at 347.1eV BE and molybdenum disulphide (MoS_2) at 162.2eV. As the samples are etched the $CaCO_3$ content is reduced, as is the sulphide layer (MoS_2), and phosphates and the substrate material beneath becomes more apparent. When using the Al-Si alloy pin there is evidence of material transfer on top of the iron sulphide/oxide layer.

There is some correlation between the work presented here and other work carried out to investigate the effect of calcium detergents on the anti-wear mechanism of ZDDP [9, 19, and 20]. The study by Wan *et al.* [20] has shown that there is an antagonistic effect between the detergent and the anti-wear additives which can reduce the effectiveness of the anti-wear properties of ZDDP. This effect occurs by the breakdown of calcium carbonate containing colloids and the deposition of Ca-soap molecules on the surface.

5.2. Al-Si Alloy Plates

5.2.1 Tribological Analysis

The tribological studies of aluminium alloys are not as extensively reported in the literature as their ferrous counterparts [10-15]. Hence these systems are much less well understood. In the work presented in this paper it has been shown that the measured friction and wear values for aluminium plate systems are significantly higher than those observed for steel plate systems. The FF lubricant show higher friction and wear than the base oil lubricant. Post-test analysis carried out using EDX and XPS confirm that the addition of additives to the base oil has a detrimental effect on friction and wear and this is explained in greater detail in the next section.

5.2.2 Post-Test Analysis (EDX and XPS)

No evidence of anti-wear film formation has been found for totally non-ferrous systems (Al-Si alloy pins on Al-Si alloy plates) from EDX and XPS analysis. This, however, is not the case for systems containing iron (cast-iron pins on Al-Si alloy plates), where traces of phosphates and sulphides are detected from XPS. EDX analysis of these systems does not however detect the P and S, which suggests that the anti-wear film formed was thinner, and hence less stable compared to films formed on ferrous systems (Figure 6b).

XPS analysis also detected calcium carbonate on the surface and it appears the Ca from the detergent exhibits a stronger affinity than S, P or Mo to the aluminium surface although the amount of calcium carbonate is less than that detected on EN31. Exactly why this is, is not fully understood but it is believed that the stable Al_2O_3 film on the aluminium alloy makes it unfavourable for the formation of the phosphate interlayer which protects the surface from wear. In the case of a ferrous system it has been shown [6] that Fe_2O_3 digestion is an important mechanism for film formation and this will not be present at all on Al-Si alloy. There is some protection afforded by the calcium carbonate but this appears to be easily removed from the surface and this can lead to the particles being reintroduced back into the bulk lubricant and can lead to abrasive wear of the surface. This explains the behaviour shown by tribological tests between the formulated oil and the base oil which shows that the addition of additives (calcium detergent) has a damaging effect on the material. This is evident when analysing the bulk oil after tribological tests where large amounts of wear debris are present from the Al-Si alloy plate systems.

Table 6 summarises the average friction coefficient and dimensionless wear coefficient performance of different lubricant and material couples using base oil as a reference. Results on EN31 plates show there is a significant reduction in the wear and friction coefficients than for the equivalent system using base oil. The exception is

cast iron on EN31 at 60°C using the FF oil where the values of wear are extremely small and real changes are difficult to determine accurately. This confirms that generally the addition of additives into the base oil forms an anti-wear film on ferrous surfaces, which not only helps to protect the surface against wear but helps reduces the friction as well. For the case using cast iron on EN31 at 60°C using the FF oil, it is believed at this temperature there is not enough activation energy for the Mo-based friction modifier to activate and form MoS_2. This wear film does not provide the same protection as that with a fully formed MoS_2 layer.

In contrast, wear results for Al-Si alloy plate systems show that there is more wear using the FF oil compared to the base oil and this ties in with our understanding of additive interactions on the material surface to form an abrasive medium. The friction data shows that for systems using cast iron pins there is an increase in the friction coefficient compared to the equivalent system using base oil whereas systems using Al-Si alloy on Al-Si alloy generally show a decrease in the friction coefficient. It is believed the increase in friction coefficient for the cast iron on Al-Si alloy system is due to the calcium carbonate layer (originating from the overbased detergent) formed on the material surface.

In Figure 7 the wear coefficient for the system is plotted against the average friction coefficient. The results are grouped according to the material couple type and it would appear there is a general trend of increasing friction coefficient versus wear for all the material couples except for Al-Si on Al-Si. This suggests that the influence of material type is greater than that of lubricant type and that greater attention needs to be paid on the material/lubricant combination.

6. CONCLUSION

From these preliminary tests it is apparent that there are differences between the surface and lubricant interaction and these appear to be dependent on the pin/plate combination. A summary of the main points from this paper is as follows:

- A relatively thick protective wear-film was formed on EN31 but not on Al-Si alloy
- XPS analysis showed there to be a thin (and therefore unstable) film formed on Al-Si alloy when rubbed with a cast-iron pin
- The fully formulated oil has a beneficial effect in terms of friction and wear for systems using steel plates
- For systems using aluminium plates the additives present in the oil have a negative effect

ACKNOWLEDGEMENTS

The author would like to thank Infineum UK Ltd for their support of this project and Jonathan Philby from Federal Mogul Ltd for supplying aluminium piston material.

Table 6
Comparison of wear and friction set against base oil as a reference

		***Wear Coef*/Friction Coef (using base oil as reference)**			
Lubricant	Temperature	Cast Iron on EN31	Al-Si on EN31	Cast Iron on Al-Si	Al-Si on Al-Si
FF Oil	**60°C**	*+133%/-9%*	*-100%/-21%*	*+1150%/+8%*	*+88%/-15%*
FF Oil	**100°C**	*-67%/-29%*	*-78%/-21%*	*+100%/+32%*	*+74%/-14%*

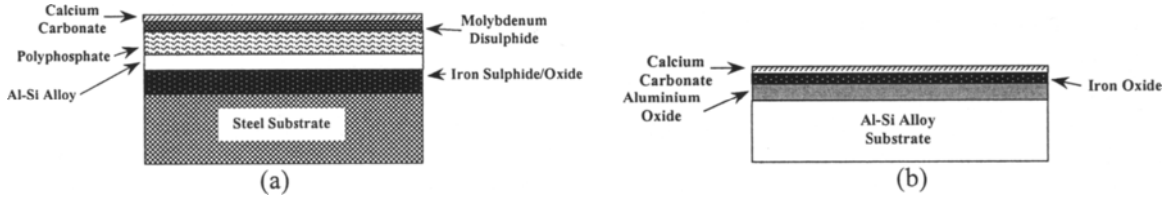


Figure 6. Schematic representation of film formation for (a) Cast Iron/Al-Si Alloy pin on EN31 plate and (b) Cast Iron pin on Al-Si Alloy Plate

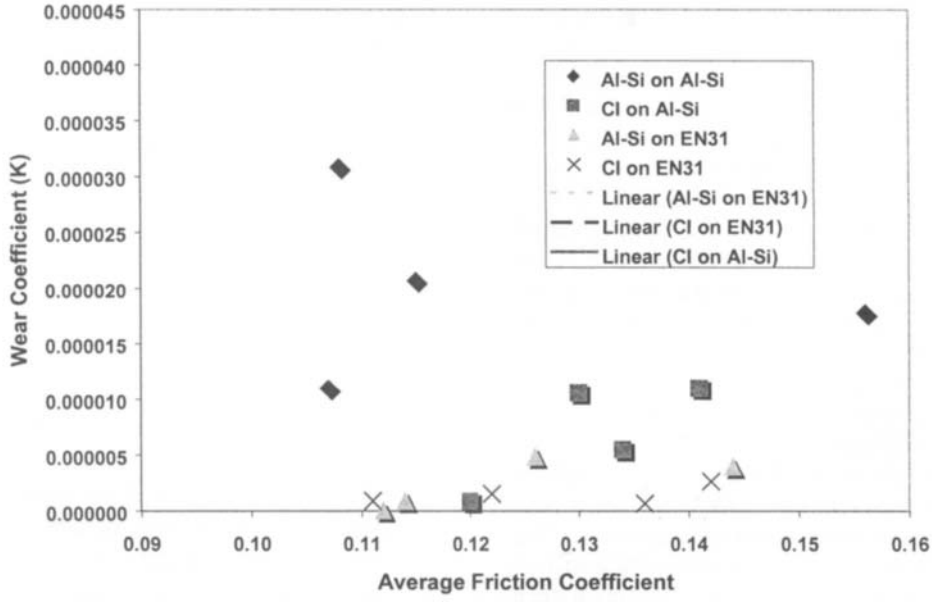


Figure 7. Wear Coefficient (K) versus average friction coefficient comparing material couples

REFERENCES

1. International Energy Annual 1999
2. H. So, Y.C. Lin, G.S. Huang, S.T. Chang, Wear 166 (1993) 17-26
3. S. Jahanmir, Journal of Tribology, Vol. 109 (1987) 577-586
4. A. Tonck, S. Bec, J.M. Georges, R.C. Coy, J.C. Bell, G.W. Roper, Lubrication at the Frontier/D. Dowson *et al.*, (1999) 39-47
5. S. Bec, A. Tonck, J.M. Georges, R.C. Coy, J.C. Bell, G.W. Roper, Proposed for publication in Proceedings of The Royal Society (1998)
6. J.M. Martin, Tribology Letters 6 (1999) 1-8
7. J.M. Martin, C. Grossiord, T. Le Mogne, J. Igarashi, Tribology International 33 (2000) 453-459
8. S.S.V. Ramakumar, N. Aggarwal, A. Madhusudhana Rao, A.S. Sarpal, S.P. Srivastava , A.K Bhatnagar, Lubrication Science 7-1 (1994) 29-38
9. P.A. Willermet, D.P. Daily, R.O. Carter, P.J. Schmitz, W. Zhu, Tribology International Vol. 28 No. 3 (1995) 177-187
10. P.A. Willermet, D.P. Daily, R.O. Carter, P.J. Schmitz, W. Zhu, J.C. Bell, D. Park, Tribology International Vol. 28 No. 3 (1995) 163-175
11. M. Reyes, A. Neville, 29th Leeds-Lyon Symposium on Tribology, 2002
12. L. Hu, J. Chen, W. Liu, Q. Xue, C. Kajdas, Wear 243 (2000) 60-67
13. S.Q. Wu, H.Z. Wang, S.C. Tjong, Composites Science and Technology 56 (1996) 1261-1270
14. X. Tao, C. Jianmin, Z. Jiazheng, D. Hongxin, Wear 196 (1996) 214-218
15. Y. Wan, L. Cao, Q. Xue, Tribology International Vol. 30 No. 10 (1998) 767-772
16. Y. Wan, Q. Xue, L. Cao, Wear 208 (1997) 57-60
17. S.R. Hah, T.E. Fischer, P.Gruffel, C. Carry, Wear 181-183 (1995) 165-177
18. B.J. Hamrock, D. Dowson, John Wiley and Son (1981)
19. P. Kapsa, J.M. Martin, C. Blanc, J.M. Georges, Journal of Lubrication Technology Vol. 103 (1981) 486-496
20. Y. Wan, M.L. Suominen Fuller, M. Kasrai, G.M. Bancroft, K. Fyfe, J.R. Torkelson, Y.F. Hu , K.H. Tan, Tribology International Vol. 31 No. 10 (1998) 627-644

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Prediction of total engine friction power loss from detailed component models

Q. Zhou, I. Shilling and S.H. Richardson

Jaguar Cars Ltd, Engineering Centre, Abbey Road, Whitley, Coventry, UK, CV3 4LF

This paper presents an analytical approach to estimate the total engine friction power loss from the detailed tribological analysis of major components such as piston assembly, cam/tappet interfaces and bearings. For the piston assembly (rings and skirt), it uses the starved and mixed lubrication model, coded as RPM99. For cam/tappet interfaces, it uses the CTM99, an in-house developed code that considers the surface roughness and EHL effects. For the bearings, including mains, big ends and camshaft, it uses the mobility method. The advantages of the above approach include: (a) the reflection of major component design details in the model, and (b) the consideration of lubricant properties at specific engine operating conditions.

1. INTRODUCTION

Engine friction losses directly affect the vehicle fuel economy performance. To meet the ever increasing demand on fuel economy, new technologies have been used increasingly in the modern engines, such as variable valve timing (VVT), variable valve lift (VVL), variable compression ratio (VCR), and variable displacement (VD), etc. These techniques, however, are mostly effective in reducing the pumping losses in most engine operating conditions. The pure mechanical friction power losses, on the other hand, may even be increased due to more rubbing components and bigger oil pumps.

In the engine community, there is always a desire to estimate the total engine friction power loss at the design stage [1-4]. Many commercial and in-house engine performance predictive tools adopt the regression curve, which calculates the friction solely as a function of engine speed. This clearly is not satisfactory, as we know the oil properties, oil temperature, surface finish and component design parameters all have big impact on the mechanical friction loss.

In this paper, we will focus on the pure mechanical friction power losses only. In general, the pure mechanical losses can be further broken down to contributions from the piston assembly, crankshaft bearings (including big end bearings), valvetrain, oil pump and auxiliary components including water pump, power assisted steering (PAS) pump, fan, generator, idle pulleys and so on. In the past decades, many detailed component friction prediction models have been reported [5-10]. Examples include the piston assembly's starved and mixed lubrication model [5-6], the cam/tappet's elastohydrodynamic lubrication (EHL) model [7-8], the oil pump's computational fluid dynamics (CFD) model [9], and various bearing lubrication and friction models [10]. In this work, the in-house code, RPM99, was adopted for piston assembly; the in-house code, CTM99, for cam/tappet interface; and the mobility method for bearings.

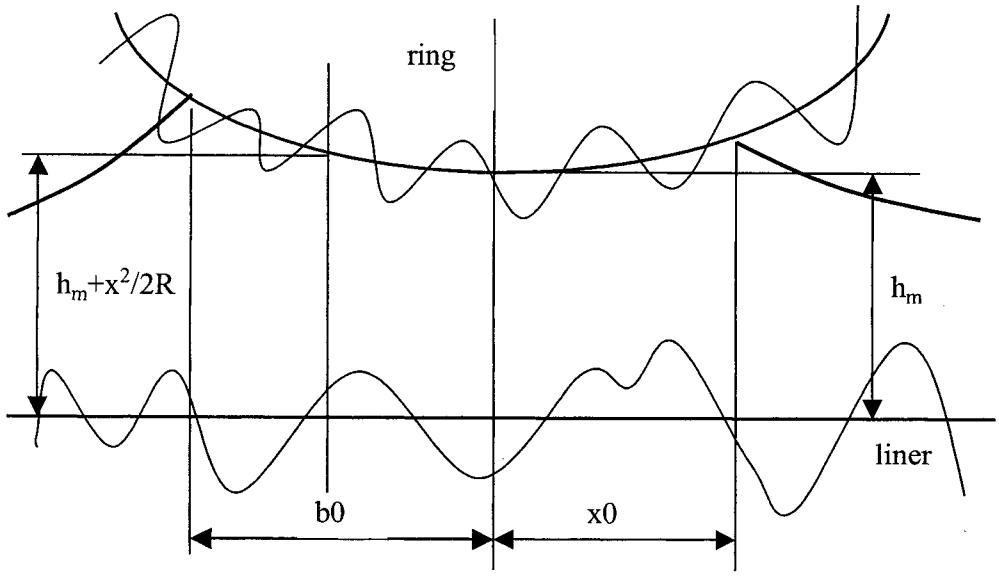


Figure 1. Piston ring and liner interface.

2. FRICTION MODELS

2.1 Piston assembly model

Figure 1 shows a typical gasoline engine piston/liner interface. The most simple and commonly used tool to simulate this lubrication system is by one-dimensional Reynolds equations. Although there are some 2D/3D tools available, they are generally either too complicated or lack the required accuracy.

For one-dimensional analysis, the general Reynolds equation for rough surfaces is

$$\frac{\partial}{\partial x}(\varphi_x \frac{h^3}{12\mu}\frac{d\bar{p}_h}{dx}) = \frac{U}{2}\frac{d\bar{h}_T}{dx} + \frac{U\sigma}{2}\frac{d\varphi_s}{dx} + \frac{d\bar{h}_T}{dt} \qquad (1)$$

Where h ($=h_m+x^2/2R$) is the oil film thickness, U the entrainment velocity, $\bar{h}_T$ the averaged oil film thickness, $\bar{p}_h$ the hydrodynamic pressure, φ_x the pressure flow factor, and φ_s the shear flow factor.

A detailed solution procedure, including the boundary conditions, load and friction contributed by the asperity contact, has been documented by Zhou [6]. In general, this model can reasonably predict the effect of oil properties, surface texture (includes roughness and roughness orientation), ring pack design parameters and engine operating conditions on the oil film thickness and friction power loss.

Zhou and Richardson [11] compared two other codes: EnFas (developed by Leeds University) and CASE (developed by Michigan State University), with the in-house code RPM99, and found the latter has advantages in overall performance including stability, sensitivity and accuracy.

2.2 Cam/tappet friction model

The equation (1) is also applicable to the cam/tappet interface. Zhou and Richardson

[8] showed that, by including the squeeze term, surface roughness and elasticity, it is possible to predict the friction coefficient from the first principle without setting any artificial limiting value. This was achieved by adopting an appropriate surface roughness contact model [8]. The asperity contact could share a large percentage of normal load under a thin film situation, so the actual pressure on the oil film and therefore the pressure induced viscosity increase would be lower when compared to the pure smooth surface model. Obviously, the model is most sensitive to the surface roughness. The lubricant properties and cam profile (or entrainment velocity) also have profound impacts on the shape of the friction coefficient curve and the maximum value. For example, when the lubricant (15W40) temperature changes from 60 to 130 °C, the maximum friction coefficient increases from 0.053 to 0.159.

Figure 2 shows the predicted friction coefficient of an intake cam of the engine studied in this paper (see section 3), for a smooth surface (Ra=0) and a rough surface (Ra=0.25 micron), both with 15W40 oil at 90 °C and 3000 rpm camshaft speed. The friction coefficient spikes (for Ra=0.25 micron) at the cam shoulders correlated well with the polishing marks often seen on the running engines.

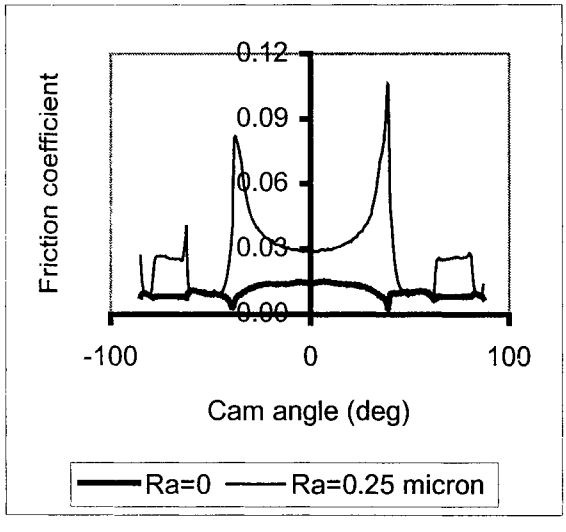


Figure 2. Predicted friction coefficient.

2.3 Bearing friction model

Engine bearings can contribute more than 20% of total friction. Nowadays, the most advanced bearing models have been developed with the EHL theory. These models, however, are so special and complex that they are not suitable for quick engine friction analysis. Instead, they are often used individually to study the influence of specific design parameters on the bearing performance.

In the automotive industry, including major bearing OEM's, the day-to-day support analysis still employs the traditional mobility method. For this reason, we have integrated the mobility method, as described by Booker [12], into our total engine friction analysis tool. The bearing loads are directly calculated from the cylinder pressure and mass inertia. The output includes the minimum oil film thickness and location, the maximum oil film pressure and location, the oil leakage and the friction torque.

With bearing dimensions becoming smaller, loads increasing, oil viscosity getting less, and temperature going up, the mobility method may under-estimate the friction. Developing a quick and more accurate method may be required in the near future.

2.4 Oil pump model

Ford and CFD Research Corporation (CFDRC) have developed a gerotor oil pump model to simulate the oil flow, pressure pulse, friction torque and cavitation within the pump. Jiang et al [9] has published the methodology. Although the primary function of this code is to predict the cavitation, it automatically calculates the instantaneous friction torques at each time step. According to Jiang [9], the code under-predicts friction when the pressure relief valve is closed. When the pressure relief valve is open, it will overestimate the power consumption due to the absence of a relief valve model.

Although (CFD-ACE+) is used as our oil pump analysis tool, for a quick friction only analysis, we often use a much-simplified equation as given below:

$$F_{op} = N * V * P / \eta_m \text{ (kW)} \qquad (2)$$

Where N is the pump speed in rev/sec, V the pump displacement in cc/rev, P the discharge oil pressure in MPa and η_m the mechanical efficiency.

3. RESULTS

An AJV6 3.0L engine has been selected for the case and benchmarking study. The basic engine details are listed below.

Bore diameter	89 mm
Stroke	79.5 mm
Compression ratio	10.5
Main bearing (4) Diameter	63 mm
Width	17.3 mm
Nominal radial clearance	16 microns
Big end bearing (6) Diameter	50 mm
Width	17.3 mm
Nominal radial clearance	25 microns
Cam bearing (14) Diameter	27 mm
Width	15 mm
Nominal radial clearance	24 microns
Valvetrain Type	DAMB
Intake valve maximum lift	9.00 mm
Exhaust valve maximum lift	9.06 mm
Base circle radius	18 mm
Cam width	11 mm
Spring pre-load	175 N
Viscosity-pressure coefficient	1.3e-8 m^2/N
Oil pump capacity	16.4 cc/rev

The base engine has been measured on a friction break down motored dyno so the friction contribution of each component (assembly) can be assessed. In the test, the coolant and oil temperatures were maintained at 90 °C to minimise their influence on the clearance and oil viscosity. Before the motoring test, total engine friction was recorded, which was used as the baseline.

Figure 3 shows the measured break-down friction data in percentage. At low speed. The valvetrain accounts for more than 40% friction but it reduces to about 15% at 6000 rpm. Whereas, the friction contribution due to the piston and main and big end bearings goes from 40% at low speed to about 60% at 6000 rpm. The friction losses are also plotted on the absolute scale, as shown in Figure 4.

Figures 3 and 4 also show the predicted friction results from 1000 to 6000 rpm. The plots show good levels of agreement between the measured and predicted results, validating the approach and methods employed. This method is planned as the main tool for firstly identifying the source of significant friction loss, secondly, to evaluate design changes and their effect on overall friction; and finally to support new programs at concept stage to ensure the best practice of friction reduction be followed.

4. FUTURE WORK AND CONCLUSIONS

This work demonstrates that it is possible to predict the engine friction using detailed component models. Although auxiliary components are not included at present, a regression curve is considered sufficient, as their effect is largely speed dependent. More work is needed to measure the engine friction under different oil temperatures so the code can be further evaluated. It is also desirable to expand the code to enable the engine warm-up process to be included. The code in current status can accept different oil temperature (viscosity) for different components; this is input rather than generated from transient heat management.

ACKNOWLEDGEMENT

Thanks are given to Jaguar Cars Ltd for permitting the publication of this paper.

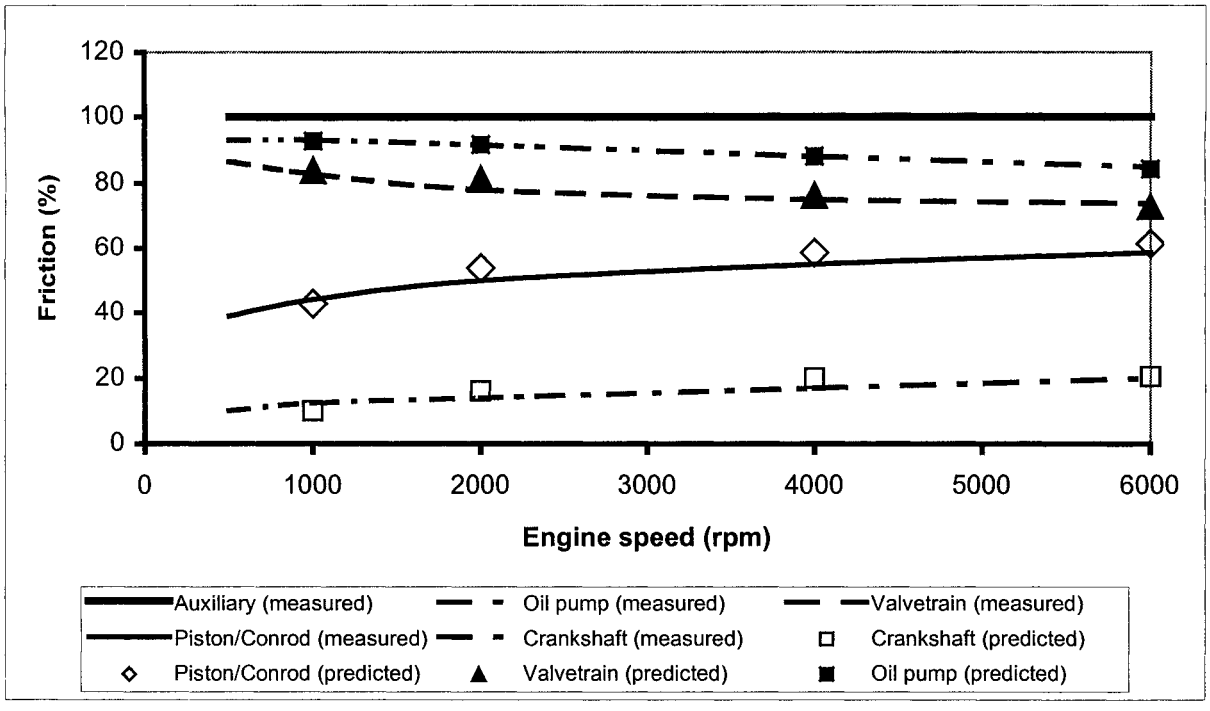


Figure 3. Measured and predicted pure mechanical friction in percentage, 15W40 at 90 °C.

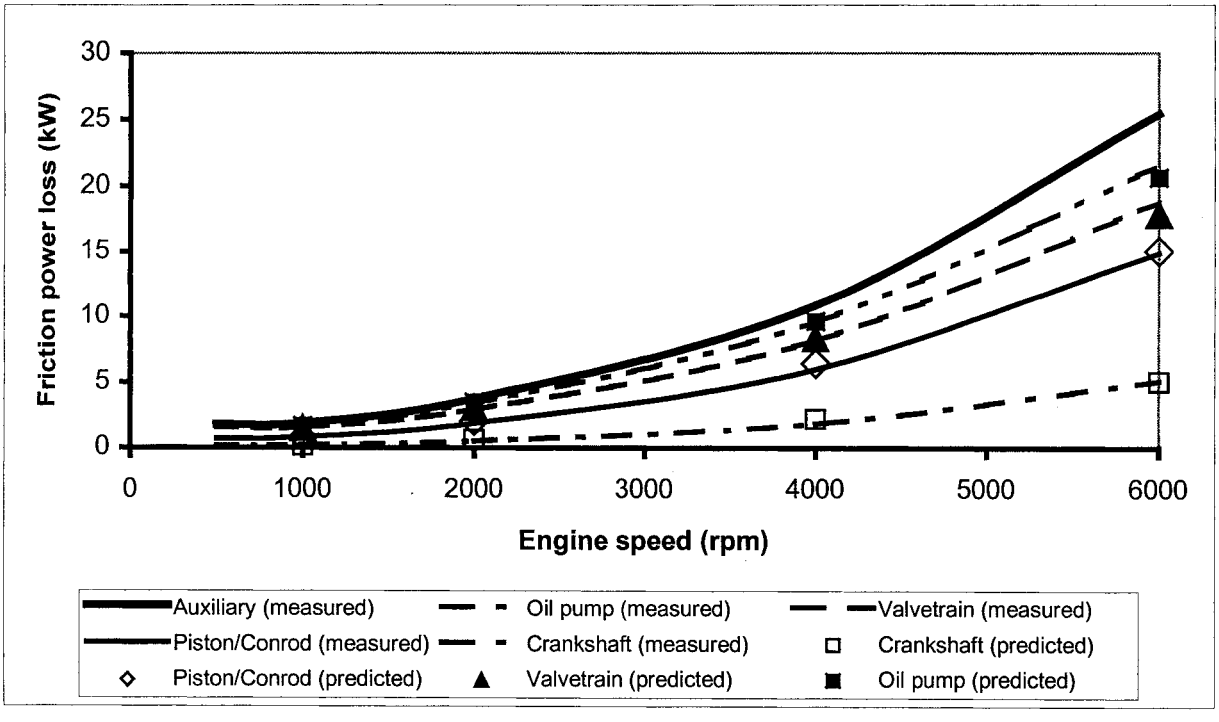


Figure 4. Measured and predicted pure mechanical friction in absolute value, 15W40 at 90 °C.

REFERENCES

1. Rosenberg, R.C.: General friction considerations for engine design. SAE 821576, 1986.
2. Patton, K.J., Nitschke, R.G and Heywood, J.B.: Development and evaluation of a friction model for spark0ignition engines. SAE 890836.
3. Taraza, D., Henein, N. and Bryzik, W.: Friction losses in multi-cylinder diesel engines. SAE 2000-01-0921.
4. Nakagawa, I.: Developing a high-precision engine friction tester. SAE 2001-08-0114.
5. Priest, M.: Factors influencing boundary friction and wear of piston rings. Thinning films and tribological interfaces, Dowson, D. (ed.), Tribology Series, 38, Elsevier, 2000, pp409-416.
6. Zhou, Q.: The mixed lubrication model of the piston ring pack in internal combustion engine and its application. Trans. CSICE, 9 (1991), pp83-90.
7. Dowson, D., Taylor, C.M. and Zhu, G.: A transient elastohydrodynamic lubrication analysis of a cam and follower. J. Phys. D: Appl. Phys. 25 (1992), pp A313-A320.
8. Zhou, Q. and Richardson, S.H.: Friction prediction of a cam/tappet interface using a EHL model with consideration of surface texture. 28th Leeds-Lyon Tribology Symposium, 2001.
9. Jiang, Y. and Perng, C-Y.: An efficient 3D transient computational model for vane oil pump and gerotór oil pump simulations. SAE 970841.
10. Katagiri, T., Hanahashi, M. and Okamoto, Y.: Parametric study for design factors on engine bearings by using TEHL analysis. SAE 2002-01-0298.
11. Zhou, Q. and Richardson, S.H.: Piston ring tribological analysis of Jaguar AJ27 4.0L V8 engine. Topics in Engine Lubrication and Engine Controls, Bell, S.R. (ed.), ASME, Vol.3, 2001, pp1-13.
12. Booker, J.: Design of dynamically loaded journal bearings. Fundamental of the design of fluid film bearings. Rohde, S.M. et al (ed.), ASME 1979, pp31-44.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Experimental Evaluation of Engine Valve Train Friction under Motored and Fired Conditions

R A Mufti and M Priest

Institute of Tribology, School of Mechanical Engineering, The University of Leeds, UK.

One way of improving fuel economy whilst maintaining acceptable exhaust emissions is to reduce engine friction. In the last three decades there has been a resurgence of interest in reducing friction in the engine valve train. However, further reduction in friction requires better understanding and measurement in the valve train friction on a real firing engine.

A new method of directly measuring camshaft friction that offers exceptional accuracy is described in this paper. The technique uses strain gauges to measure the torque exerted upon the camshaft drive pulley sprocket. The advantage of this method over previous techniques is the instantaneous valve train friction as a function of crank angle can be measured, simultaneously from both inlet and exhaust camshafts, under both motored and fired conditions. Experiments have been carried out for both motored and fired conditions on a single cylinder gasoline engine. Results are presented for instantaneous and mean valve train friction as a function of engine speed and oil temperature, indicating a decrease in friction with increase in engine speed and an increase with increasing temperature. The fuel economy benefit of friction modified lubricants was also observed. A significant difference in valve train friction between motored and fired condition was found, especially for the exhaust camshaft.

1. INTRODUCTION

Internal combustion engine designers and lubricant developers rely heavily on computer based analytical tools. The sophistication and functionality of these tools is growing rapidly. It is therefore important that the models, on which these techniques are based, are validated and continually improved by very reliable and accurate experimental techniques.

A number of experimental studies have been carried out in the last three decades to study the friction losses at the cam/follower interface and the effect of lubricant properties and additives with regard to the failure of cams and followers. Dyson and Naylor [1] were possibly the first to report upon attempts to measure the friction at the cam/follower interface. Their apparatus incorporated a single flat-faced follower and cam arrangement. Armstrong and Buck [2] measured the drive torque and power requirements of various types of valve train using a motored cylinder head with an in-line torque transducer. Ito, Yang and Negishi [3] used a push rod cam/follower rig to simulate cam/follower operation. The cam/follower interface friction force was measured using an instrumented shaft that held the follower. Dowson et al [4] measured the friction and power loss of a single cam and bucket follower and associated components in a motored rig. Instantaneous and average friction torque was measured using an in-line torque transducer.

Most of the experiments reported were carried out on test rigs under motored conditions. A new experimental method is described here allowing detailed analysis of a single cylinder engine valve train friction under both motored and fired conditions. As the technique was implemented on a single cylinder engine, this permitted in-depth study of the influence of cylinder gas pressure on valve train friction during valve opening and closing. This technique allows instantaneous as well as

mean friction measurements to be carried out at any engine speed.

The authors applied this technique successfully on the same engine under motored conditions. The results were used to validate an engine valve-train friction model [5]. The experiments were carried out at different engine speeds and lubricant temperatures, studying the behaviour of 0W20 lubricant with and without an organic friction modifier. Both mean and instantaneous friction was measured and found to correlate very well with the Stribeck curve.

2. INSTRUMENTATION

2.1. Pulley Torque Transducer

The specially designed pulley torque transducer consists of three parts: the instrumented central hub; the slip rings and the belt drive gear, figure 1. The central hub consists of four strain gauged beams. The strain gauges are installed in such a way that the transducer is insensitive to belt tension and only registers applied torque on the belt gear. The central hub is manufactured independent of the drive gear and is attached via four bolts, allowing the transducer to be fitted to any automotive cam drive. A female adapter is machined in the central hub so that the installation procedure of the torque transducer is exactly the same as the original cam pulleys. Thus the transducer is integrated seamlessly into the original engine drive train arrangement. The signals from both the pulley torque transducer for inlet and exhaust camshaft are transmitted via the slip rings to a Fylde high performance transducer amplifier. The complete detail of the transducer can be found in [5].

The experiments were carried out on a single cylinder Ricardo Hydra gasoline engine the design of which is based on a standard 2.0 litre GM engine. It is a four-valve engine with both the cams on each camshaft in phase. Also the camshafts are short and rigid giving the benefit of low torsional vibration in the drive train system. The engine drive train layout is shown in figure 2.

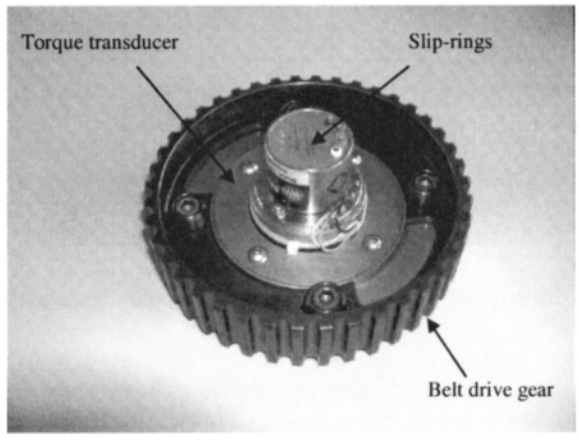


Figure 1. Specially designed camshaft pulley torque transducer

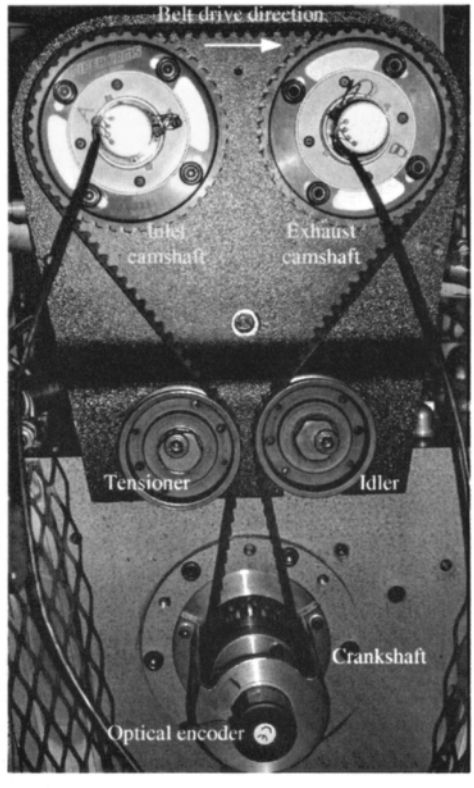


Figure 2. Ricardo engine drive train layout.

2.2. Speed Measurement

As the experiments were carried out on a single cylinder engine it was important to monitor the speed variation of the camshafts. A 720 pulse per revolution optical encoder was fitted via a bellows flexible coupling to the belt drive end of the crankshaft to measure crankshaft and thus the camshaft speed as the speed ratio is 2:1. The encoder index signal was used to synchronise the reading of camshaft torque with crank angle.

2.3. Lubricant System

A completely independent oil circuit with an external oil sump was used to supply lubricant to the engine valvetrain. Lubricant temperature was measured in the oil sump, at the inlet to the camshaft bearings and at the centre of the tip of the oil feed nozzle to the cam/follower interface. All temperatures were measured using K-type thermocouples.

2.4. Cylinder Pressure

Cylinder pressure was measured under both motored and fired condition to study the effects on engine valve train friction. For this a 6067B Kistler water cooled piezo-electric pressure transducer was used along with its signal conditioning unit, a Kistler 5011 charge amplifier. The piezo-electric pressure transducer measures the relative pressure and to relate this pressure to the absolute value an absolute resistive pressure transducer was installed in the cylinder liner at 120 degrees crank angle after top dead centre from top of the piston.

2.5. Data Acquisition System

The two torque transducers, cylinder pressure transducer, liner absolute pressure transducer and two encoder signal were connected to a 12-bit high speed simultaneous sampling data acquisition system. Two National Instruments 6110E boards were used each having four analogue input channels and two 20 MHz timer/counter channels. Unlike most of the data acquisition boards the 6110E does not work in multiplex mode as each channel has its own analogue to digital converter. Thus there is no phase lag between the channels. The operation of both the boards was synchronised using an external switch as a trigger, connected in-line to the encoder signal. The incremental 720 pulse per revolution signal from the encoder was used as an external clock for sampling the analogue channels, thus recording the data every 0.25 degree of cam angle. The same encoder signal was also used to monitor camshaft speed variation by measuring the time period of each encoder pulse. By numerical differentiation of this time data, the variation of cam speed, and thus the acceleration of the camshaft was calculated.

3. CAMSHAFTS

The Ricardo Hydra engine has direct-acting overhead camshafts acting on flat-faced, hydraulic lash adjusted bucket followers driven by a toothed timing belt. Each camshaft has three journal bearings. A solid modelling package was used to calculate the inertia of the complete camshaft assembly for both the inlet and exhaust camshafts, figure 3. This was required to determine the inertial torque effects.

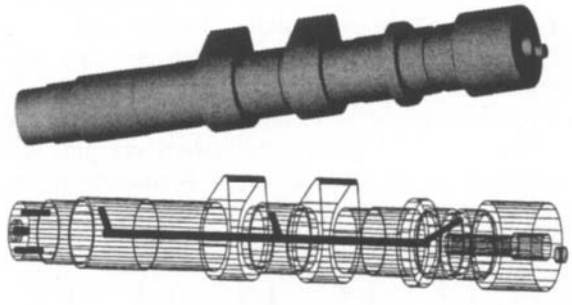


Figure 3. Single cylinder gasoline engine camshaft.

4. CALIBRATION

Calibration of both the camshaft torque transducers was carried out to establish accurate transducers correlation between the torque and output voltage over the range expected in the Hydra engine and to ensure that external factors such as

drive belt tension did not influence the output signal during running.

The transducers were statically calibrated in both clockwise and anti-clockwise directions over a range of ±20Nm. The response of the transducers was linear and highly repeatable. The details of the calibration method can be found in [5]. The influence of belt loading on the transducer was evaluated by applying a simulated belt tension. The results clearly showed no significant effect of belt loading on the transducer output reading, with errors less than 0.06 Nm and 0.03 Nm for instantaneous and mean drive torque respectively [5].

5. VALVETRAIN FRICTION SIGNAL

The data collected from the camshaft pulley torque transducers has a characteristic waveform with a frequency proportional to engine speed. As the experiments were carried out under both motored and fired conditions the shape of the trace remained constant, demonstrating high transducer repeatability. Due to this high repeatability, the experimental results given in this paper are given as average of 25-engine cycles. The vibration of the camshaft system was studied by carrying out a frequency spectrum analysis and the data was digitally filtered using a zero phase Butterworth filter. Due to the short length and rigidity of the camshafts the vibration in the system was quite low at engine speeds up to 2500rpm. A snapshot of the transducer output voltage is shown in figure 4.

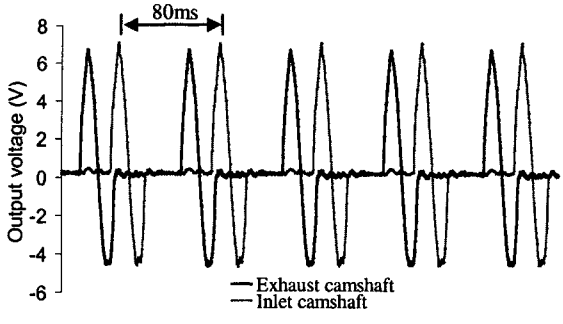


Figure 4. Typical torque transducer output voltages, engine speed 1500rpm.

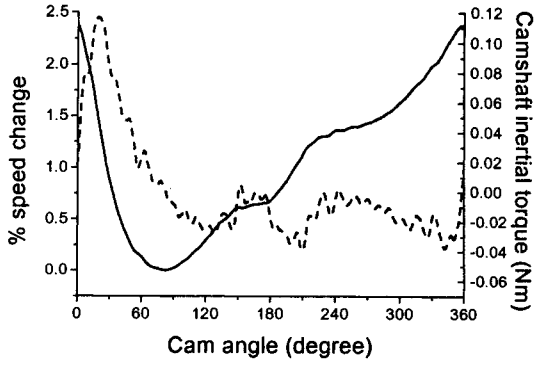


Figure 5. Camshaft speed variation and predicted assembly inertial effect.
Engine speed 2000rpm, fired.
——— Speed, -------- Inertia

Under motored conditions the camshaft torque transducer measures the sum of geometric and friction torque, which is a function of tribological performance. Geometric torque arises from the force required to open the valve against the resistance of the valve spring, the energy from which is stored in the valve spring and returned to the camshaft as torque in the opposite rotational direction as the valve closes. In the conventions used in this paper, during valve opening the geometric torque is positive and when closing it is negative. Friction torque is positive throughout so that during valve opening, more friction give longer positive values of total torque. Conversely, during valve closing more friction diminishes, make more positive, the negative torque. Taking the mean of the drive torque over cam event effectively removes the geometric torque contribution, which ideally has equal positive and negative parts and yields an average cyclic friction torque, assuming there is no energy loss in the geometric torque during valve opening and closing. Under fired conditions, the camshaft torque transducer measures two additional torques. One is the inertial torque of the complete camshaft assembly due to speed variation throughout the complete engine cycle, figure 5, and the other is the torque required by the exhaust camshaft to open the valve against the cylinder pressure gas force, and is referred to as gas torque

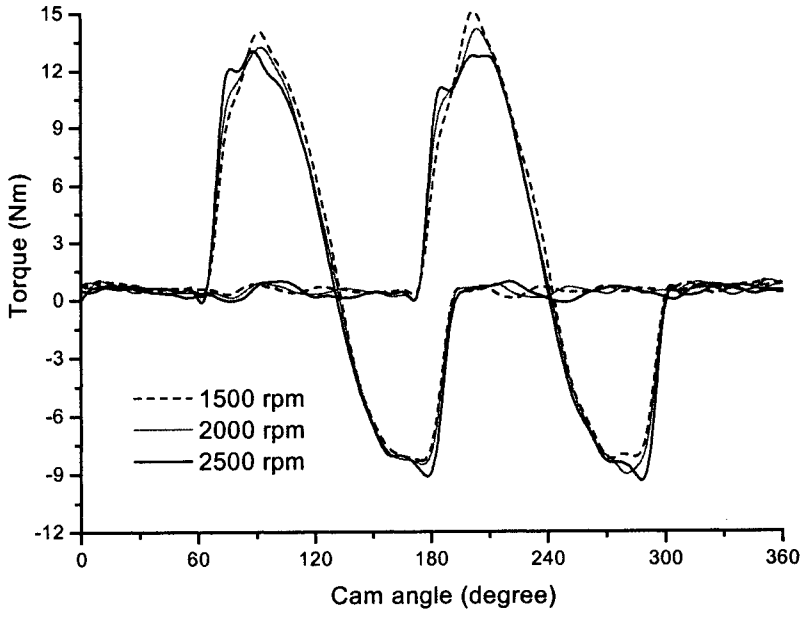


Figure 6. Instantaneous valve train drive torque under fired conditions, 10W40, 95°C oil inlet temperature

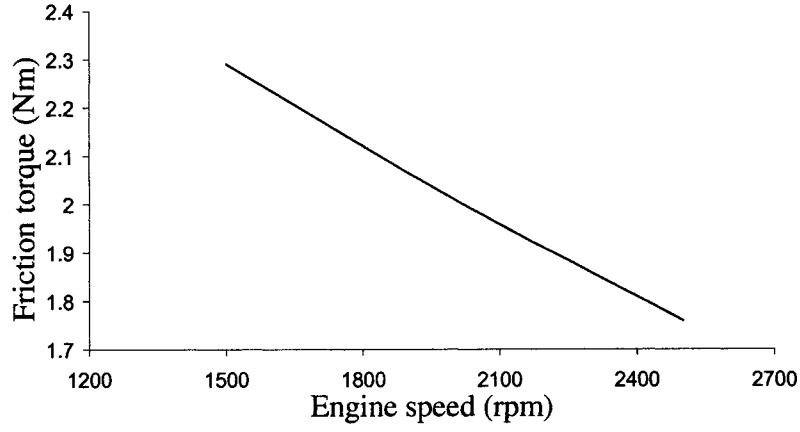


Figure 7. Averaged valve train friction torque under fired conditions, 10W40, 95°C oil inlet temperature

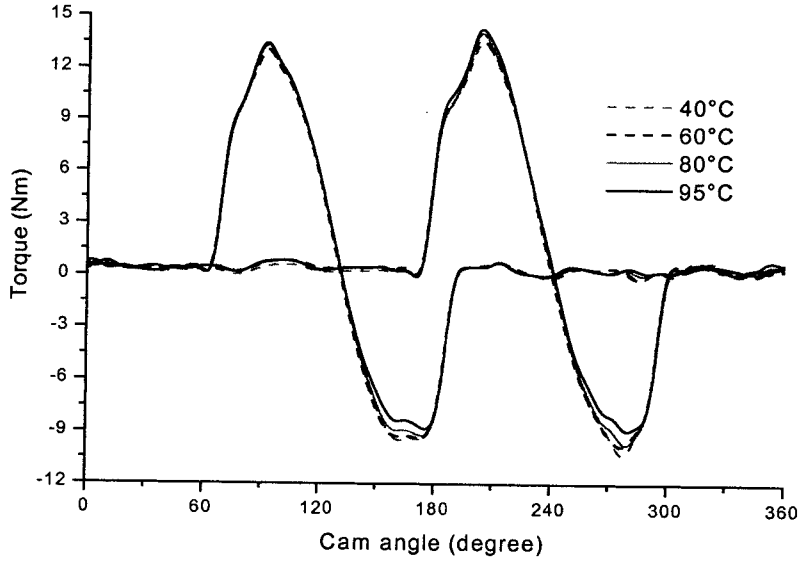


Figure 8. Instantaneous camshaft drive torque, motored 2000rpm, 10W40.

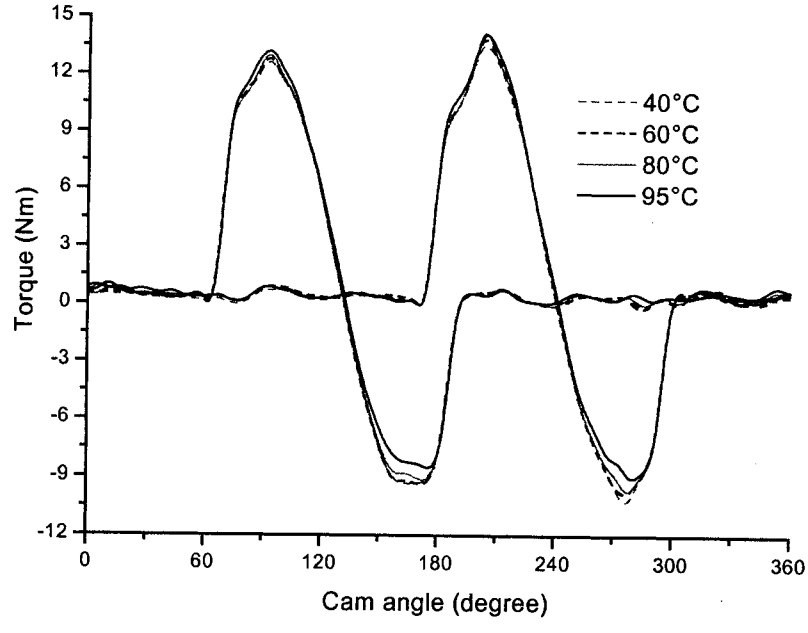


Figure 9. Instantaneous camshaft drive torque, fired 2000rpm 10W40.

in this paper. This inertia torque is the product of the mass inertia of the camshaft assembly and the angular acceleration. The encoder connected to the crankshaft measures the angular acceleration of the camshaft. The gas torque is the product of gas force acting on the exhaust valve (valve area × cylinder pressure) and the eccentricity of the point of contact between the cam and follower from the follower centre axis. Thus before averaging the drive torque measured under fired conditions, the effect of the inertia torque and the gas torque was removed from the output data.

6. VALVETRAIN FRICTION WITH ENGINE SPEED VARIATION UNDER FIRED CONDITION

The experiments were carried out using two fully formulated lubricants, 10W40 and 5W30 with friction modifier. The variation of instantaneous and averaged valve train friction under fired conditions with engine speed is presented in figures 6 and 7 respectively for the 10W40 oil at a lubricant inlet temperature of 95°C. The inertia and gas torque has been removed from the data in both the figures. It is a relatively small contribution for the given conditions, as can be seen by comparing figures 5 and 6.

It can be seen in figure 7 that as the camshaft rotational frequency increases the average friction torque decreases markedly due to improved lubrication conditions at higher frequencies arising from increased lubricant entrainment and reduced load at the cam nose. At relatively low engine speeds, the cam load is dominated by valve spring compression resulting in high friction around the cam nose. However at high speeds, inertia plays a major role as the acceleration of the tappet has a negative value around the cam nose and it becomes more negative as engine speed increases, resulting in a decrease in load around the cam nose. The opposite effect can be seen in the instantaneous friction data of figure 6 for a short period during the start of valve opening and end of closing as the acceleration and thus the inertia of the tappet increases with speed in this region.

7. EFFECT OF OIL TEMPERATURE UNDER MOTORED AND FIRED CONDITIONS

Figures 8 and 9 show the variation of instantaneous valve train friction with lubricant temperature for both motored and fired conditions respectively. It can be seen that as the oil temperature increases the valve train friction increases. This is because as the temperature increases the viscosity decreases exponentially and so the thickness of the film between the surfaces decreases. This results in an increase in the boundary component of the friction and a decrease in the hydrodynamic component. The effect is very dominant at the peaks where the load is high. This can also be seen in the average data of figure 10 at any engine speed for the 5W30 lubricant, a slight decrease and then relatively high increase in average valve train friction for the exhaust camshaft, reflecting a gradual decrease in hydrodynamic action and an increase in boundary lubrication.

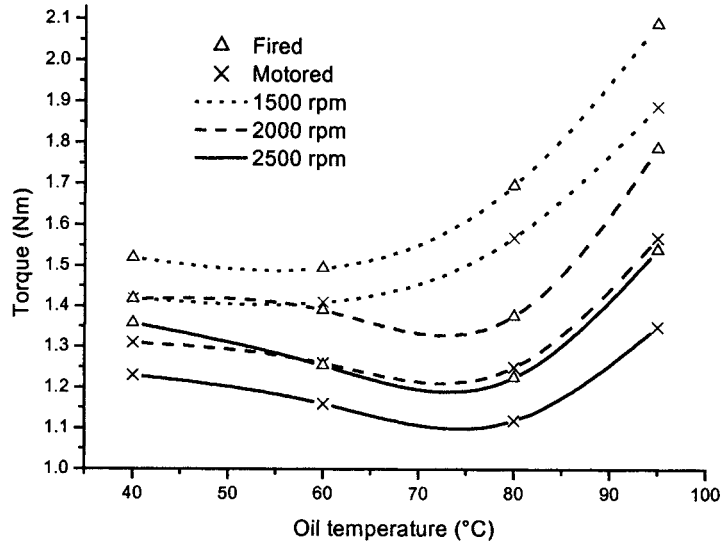


Figure 10. Average exhaust camshaft friction torque for cam profile period only, 5W30

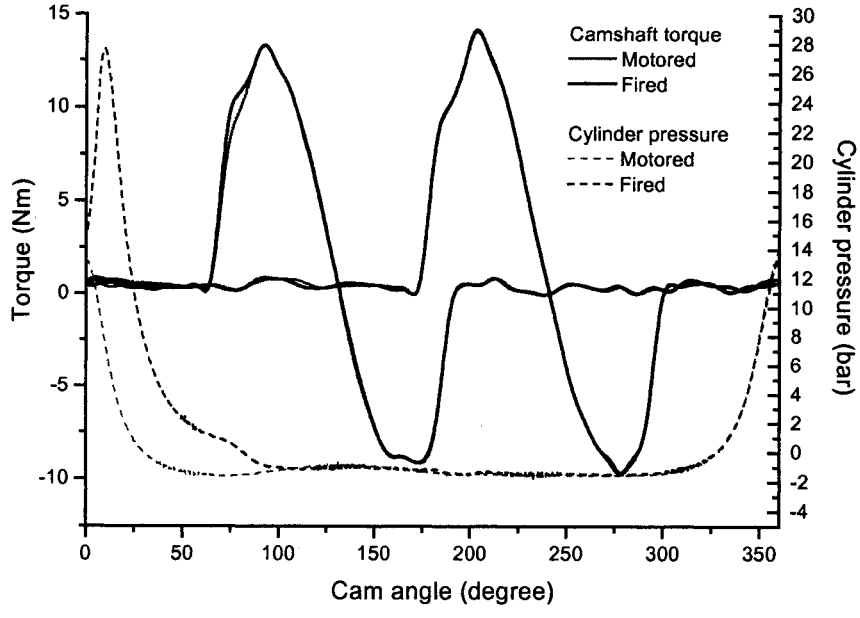


Figure 11. Drive torque under motored and fired conditions, 2000rpm, oil 80°C, 10W40.

8. DIFFERENCE OF VALVETRAIN FRICTION BETWEEN MOTORED AND FIRED CONDITIONS

By comparing figures 8 and 9 for motored and fired conditions it can be seen that a significant difference in the shape of the exhaust camshaft data whereas the inlet camshaft data remains nearly the same. A clearer picture can be seen in figure 11, comparing the motored and fired valve train torques for an engine speed of 2000rpm and oil temperature of 95°C. There is hardly any difference in the inlet camshaft drive torque but a sharp rise in torque for the exhaust camshaft during valve opening can be seen under fired conditions.

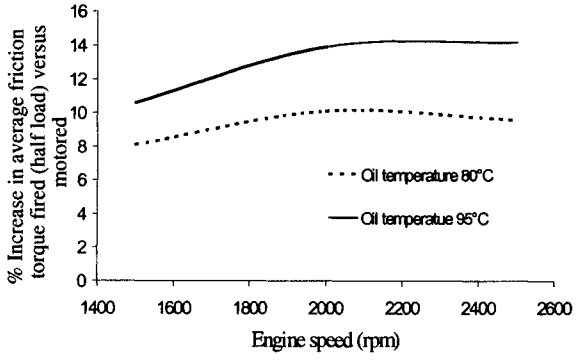


Figure 12. % increase in exhaust camshaft friction torque, fired (half load) versus motored conditions, 5W30.

When comparing the results with the cylinder pressures in figure 11 it can be seen that this increase in friction torque is due to the extra loading acting at the cam/follower interface as a result of additional work carried out by the cam to open the valve against the cylinder pressure acting on the valve face. The difference between the motored and fired exhaust camshaft average friction torque remains nearly constant for any given engine speed, figure 10.

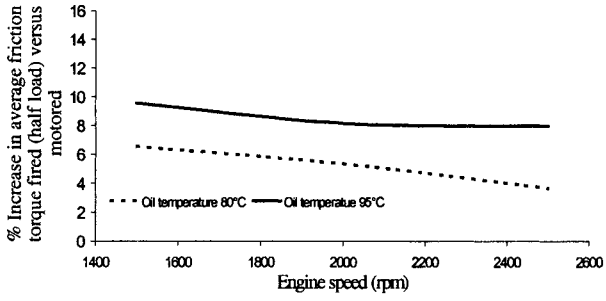


Figure 13. %increase in total valve train friction torque, fired (half load) versus motored conditions, 5W30.

The exhaust camshaft friction torque difference between fired and motored increases as the engine speed increases, figure 12. It increases roughly from 8% to 10% and 10% to 14% for engine speeds of 1500rpm to 2500rpm for oil temperatures of 80°C and 95°C respectively. The percentage increase in total average engine valve train friction torque, inlet and exhaust camshafts, remains fairly constant as shown in figure 13, at around 4% to 7% and 9% to 10% for lubricant temperature 80°C and 95°C respectively.

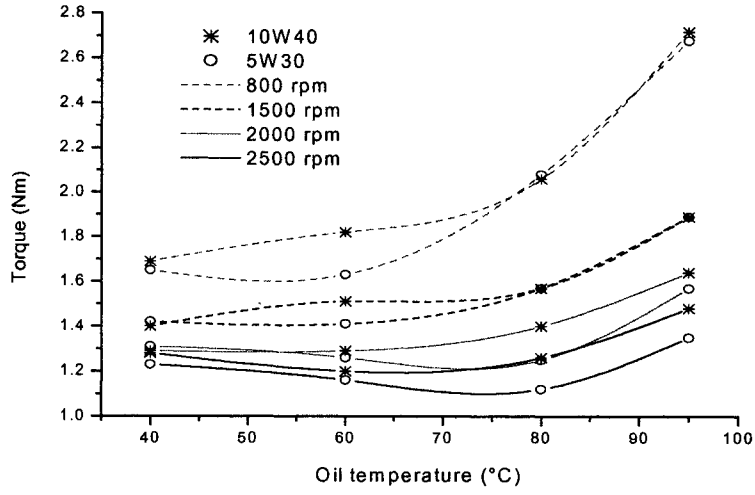


Figure 14. Averaged exhaust camshaft friction torque for cam profile period only. 10W40 and 5W30

9. BENEFITS OF FRICTION MODIFIER

The effect of both the lubricants 10W40 and 5W30 on the engine valve train was also studied under different operating conditions. Comparisons at oil temperatures of 40°C, 60°C, 80°C, and 95°C and engine speeds of 800rpm, 1500rpm, 2000rpm and 2500rpm are shown in figure 14.

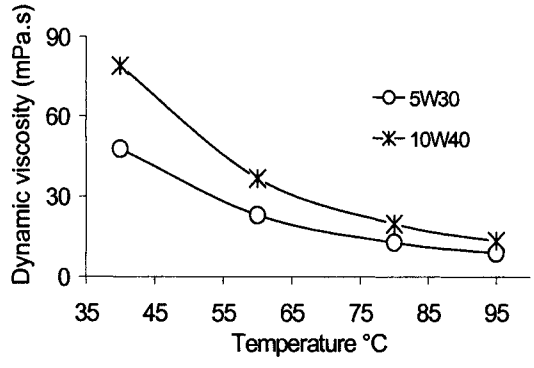


Figure 15. 10W40 and 5W30, viscosity-temperature relationship.

It can be seen that with 10W40, the valve train friction rises with oil temperature but with 5W30 it can be seen that the friction decreases and then increases. In order for valve train friction to decrease with rising temperature the lubrication regime must be mostly hydrodynamic. The benefit of 5W30 can be clearly seen at higher lubricant temperatures due to the presence of the temperature sensitive additive, friction modifier.

Figure 15, shows the viscosity-temperature relationship for both the oils for reference.

10. CONCLUSION

- A new experimental technique to measure instantaneous and mean camshaft drive torque under both motored and fired engine conditions, without modification of the drive train has been developed and demonstrated to be a powerful tool to investigate all aspects of valve-train friction. Instantaneous friction data provides detailed insights into the mechanisms of valve train friction whereas mean friction data is of considerable interest to camshaft designers, engine developers and lubricant formulators.

- Engine valve train friction decreased with increasing engine speed but increased as the oil temperature increased and the viscosity reduced.
- Total valve train friction under fired conditions was measured and found to be greater than motored friction. The difference varies as the speed increases. There is a variation of 4% to 10% average friction torque under fired condition as compared to motored for an engine speed range from 1500 to 2500 rpm.
- There is hardly any difference in the inlet camshaft friction torque between motored and fired but there is significant difference for the exhaust camshaft friction. The percentage increase in exhaust valve train friction under fired conditions increases from 10% to nearly 14% as speed increases from 1500 to 2500 engine rpm, compared to motored.
- The fuel economy benefits of low viscosity and friction modified lubricant is demonstrated by comparing the results of two fully formulated engine oils 10W40 and 5W30. The effect was significant at high oil temperatures.

ACKNOWLEDGEMENTS

The authors would like to thank the Engineering and Physical Sciences Research Council (UK), Federal Mogul Corporation, Ford Motor Company, Jaguar Cars Limited and Shell Global Solutions for their financial and technical support of this research.

REFERENCES

[1] Dyson, A., and Naylor, H., 1960, "Application of the Flash Temperature Concept to Cam and Tappet Wear Problems," Proc. Instn. Mech. Engrs., Automotive Div., No. 8, pp. 255-280.

[2] Amstrong, W. B., Buck, B. A., 1981, "Valve Gear Energy Consumption: Effect of Design and Operation Parameters", SAE Paper 810787.

[3] Ito, A., Yang, L., and Negishi, H., 1998, "Study on Cam Wear Mechanism with a Newly Developed Friction Measurement Apparatus", SAE Paper 982663.

[4] Dowson, D., Taylor, C. M., and Zhu, G., 1989, "An Experimental Study of the Tribology of a Cam and Flat-faced follower", 2nd Int. Conf., Combustion Engines - Reduction of Friction and Wear, Instn. Mech. Engrs. Conf. Pub., London, Paper C375/025, pp. 97-108.

[5] Mufti. R. A., and Priest, M., 2003, "Experimental and Theoretical Study of Instantaneous Engine Valve Train Friction", American Society of Mechanical Engineers, Journal of Tribology, Vol. 125, pp. 1-10.

APPENDIX

ENGINE AND VALVE TRAIN DATA FOR THE RICARDO HYDRA GASOLINE ENGINE.

Engine Bore	86 mm
Engine stroke	86 mm
No. of cylinders	1
No. of valves/cylinder	2 Inlet, 2 Exhaust.

LUBRICANTS

	SAE 10W40	SAE 5W30
Dynamic viscosity at 40°C	79 mPa.s	48 mPa.s
Dynamic viscosity at 95°C	13 mPa.s	9 mPa.s

CAMSHAFTS

Oil feed pressure	4.0 bar		
Valve timing	Inlet valve	open: 20°BTDC	close: 60° ABDC
	Exhaust valve	open: 50°BBDC	close: 32° ATDC
Maximum lift	Inlet: 9.49 mm	Exhaust: 9.48 mm	
Spring rate	35.8 kN/m		
Cam width	14 mm		
Base circle radius	17 mm		
Load over cam base circle	65 N		
No. of bearings	3 for each camshaft		
Cam journal diameter	28 mm		
Each cam journal width	21 mm, 16 mm, 27 mm		

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Inertia, shear-thinning and thermal effects on connecting rod bearing behaviour

B. Fantino and B. Bou-Saïd

INSA – Laboratoire de Mécanique des Contacts – UMR 5514, Bât. : Jean d'Alembert
18, Rue des Sciences 69621 Villeurbanne Cédex – France

Abstract

In earlier works Fantino et al [5-6] and Martin [8] give a summary of the studies performed in connecting rod bearings where elasticity and inertia effects from solids were included In these theoretical and experimental works, it has been shown the strong influence of the housing deformations as well as the shape defects in the connecting rod bearings performances. More recently, the study of bearings subjected to dynamic loads has showed that inertial effects and surface accelerations play an important role in the bearing response (Bou-Saïd et al [3]). Although the lubricant was considered Newtonian, this assumption is no longer valid with modern lubricant. In industrial applications, long soluble polymer chains are added to mineral lubricants to optimise properties under different operating conditions. The addition of these polymers results in a viscosity drop for shear rate values ranging from 10^{+6} to 10^{+8} s^{-1}. A previous study from Bou-Saïd et al [3] has examined the influence of fluid inertia, and effects of shear thinning in a connecting rod bearing behaviour under dynamic loading. In the present study we include the thermal effects which usually play an important role in bearing performances. Using all these assumptions (shear thinning and inertial properties of the lubricant, surface accelerations and thermal effects), the results show important differences in shaft responses compared to the classical cases.

Nomenclature

c (m), radial clearance = R_0-R_1
F_r,F_t (N), bearing radial and tangential forces
F_r*,F_t* (N), dimensionless bearing forces
h (m), film thickness
I_{xx},I_{xz},I_{zz} (m^3s^{-2}) convective inertia variables
K (Ns^nm^{-2}), viscosity parameter
m (Kg), rotor mass
M (-), dimensionless rotor mass
P,P_a (Nm^{-2}), pressure, ambient pressure
p* (-), dimensionless pressure
= $8(p-p_a)^2 c/\mu \omega L^2$
q_x,q_z (m^2s^{-1}), flow rate variables
R,R_1,R_0 (m), bearing radius ; inner, outer radius, R=R_1=R_0
Re (-), Couette Reynolds number
Re* (-), modified Reynolds number
u,v,w (ms^{-2}), fluid velocity components in x,y,z
U (ms^{-2}), surface sliding velocity
U_1,W_1 (ms^{-2}), average parabolic velocities in x,z.
W (N), externally applied load
W* (-), dimensionless externally applied load
W_{rel}* (-), relative impulsive load = W_d/W_s
W_s,W_d (N), steady and dynamic load components
x,y,z (m), spatial independant variables
α, β (-), extent of positive pressure film
ε (-), eccentricity ratio = e/c
θ (-), bearing angular co-ordinate
μ (Pa.s), dynamic viscosity
μ_a (Pa.s), kinematic viscosity
n (-), power low index
ρ (kgm^{-3}), density
τ (-), dimensionless time
ϕ (-), attitude angle
ψ (-), angle of applied force
ω (s^{-2}), shaft rotation speed
T temperature
$\dot{\gamma}$ shear rate

1. INTRODUCTION

Tichy and Bou-Saïd [15] and Bou-Saïd and Ehret [3] have presented theoretical responses of the fluid in a journal bearing with impulsive loads taking inertia and shear thinning effects into account. This led the authors to conclude that these effects are relevant factors in such operating conditions. Impulsive loads are not rare. They may be caused, for example, by the sudden unbalanced load due to the loss of a blade in a turbine shaft bearing, or by the rapid change of force on crankshaft bearings at the firing time. Interest of such analyses is high, from a safety and reliability point of view. For more than 100 years, plain bearing hydrodynamic lubrication has been the subject of experimental and theoretical studies. Nevertheless, the performance of dynamically loaded engine bearings is still a major concern for both automotive and petroleum industries. Thanks to the many engine tests run by car manufacturers and lubricant/additive producers during the development of a new engine or new oil, field performances are generally satisfactory and mechanical troubles are rare. A better theoretical understanding of the engine bearing hydrodynamic operation will probably never replace engine testing which includes boundary lubrication effects. But it should help to describe the working conditions of the lubricant and to better define the desirable rheological characteristics of the oil, in relation to the oil film thickness, the friction torque and the oil flow. In spite of the simplicity of their design, dynamically loaded journal bearings are not easily calculated or instrumented. From the rheological point of view, the main simplifying assumption is certainly the isoviscosity of the lubricant in the bearing. Even so, it is often necessary to simplify the equations (short bearing or long bearing theories) and introduce assumptions concerning the extent of the cavitation zone. Fortunately, if the dynamic loading and variable surface speed relative to the load direction increase the complexity of the problem, the rotation of the load distributes the heat generated in the oil around the bearing shell. Under a static load, the temperature difference around the bearing can reach 40°C, in a dynamically loaded bearing the maximum temperature variation is about 4°C, corresponding to a maximum variation of viscosity of only 7% as described by Fantino et al [5]. In others words the assumption of the isothermal operation of the bearing is more acceptable in the case of engine bearings than in the case of statically loaded bearing rigs under the assumption of Newtonian flow. However, the assumption of isoviscosity cannot be reduced to the assumption of isothermal operation, because it is known that pressure and shear rate (or shear stress) affects the viscosity of multigrade engine oils. In the case of a statically and heavily loaded bearing it may be calculated that the increase in the minimum film oil thickness due to pressure is only about 6%. Accordingly, it seems that piezoviscosity will not have a significant effect on the operation of an engine bearing. The importance of rheological characteristics of polymer oils has been periodically mentioned since the early work by Ockrent [9]. Even if normal stress effects and extensional viscosity are receiving increasing attention, the theoretical models and the experimental methods of evaluation of such rheological parameters are far from being universally accepted (ASTM) [1]. This study presents improvements on Tichy and Bou-Saïd's theory and takes into account surface accelerations, inertia effects as well as shear-thinning and thermal effects brought about by compounded lubricants. Additives, which can be described as long polymer chain, were originally used to limit the viscosity decrease of temperature. However, because these polymers tend to line up in the direction of the flow, viscosity drops when the shear rate reaches 10^6 - 10^8 s^{-1}. These values are common in engine bearings. In the abundant literature on this subject, the authors will concentrate on the Paranjpe's paper [10]. Using the constitutive equation proposed by Gecim [7], and Elrod's algorithm to include the Jacobson, Floberge and Olson cavitation boundary conditions, Paranjpe solved the case of dynamically loaded finite journal bearings. Earlier, Rastogy and Gupta [13], obtained

dynamic responses in short journal bearing using a power law model. Assuming the velocity profiles and an infinitely long bearing, Safar [14] computed shaft trajectories for dynamic loads and power law fluids. In comparison, steady state characteristics of journal bearings have been studied with different constitutive laws; Reiner-Philippof's law, used by Rajalingham et al [12], Rabinowitsch's relation used by Williams and Symmons, [16] and Gecim's law [7].

2. ANALYSIS

Isothermal, incompressible flow and thin film operating conditions are assumed. The viscosity is presented as a function of shear rate $\overset{o}{\gamma}$ Body forces are neglected. The equation of continuity and the equations of motion according to the configuration shown in fig. 1 a and 1 b read :

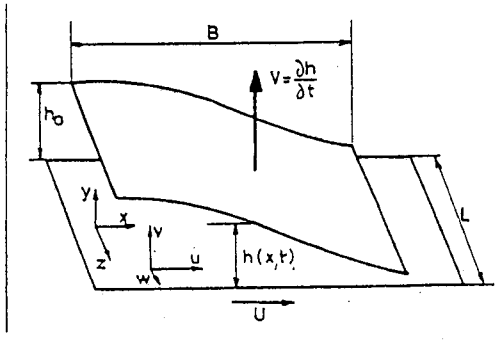


Figure 1a : Contact configuration

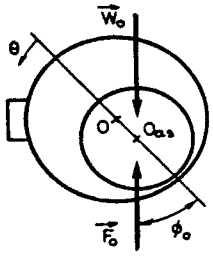


Figure 1b : Contact configuration

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \quad (1)$$

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}\right)=-\frac{\partial P}{\partial x}+\frac{\partial}{\partial y}\left(\mu\frac{\partial u}{\partial y}\right) \quad (2)$$

$$0=\frac{\partial P}{\partial y} \quad (3)$$

$$\rho\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}\right)=-\frac{\partial P}{\partial z}+\frac{\partial}{\partial y}\left(\mu\frac{\partial v}{\partial y}\right) \quad (4)$$

With boundary conditions:

$$y=0\ u=U\ v=w=0$$

$$y=h(x,t)\ u=0\ v=\frac{\partial h}{\partial t}\ w=0 \quad (5)$$

Following the method already applied by Tichy and Bou-Saïd [15], and Bou-Saïd and Ehret [3], the previous equations are integrated across the film. The following variables are defined :

$$q_x=\int_0^h u\,dy\ \ q_z=\int_0^h w\,dy\ \ I_{xx}=\int_0^h u^2\,dy$$

$$I_{xz}=\int_0^h uw\,dy\ \ I_{zz}=\int_0^h w^2\,dy \quad (6)$$

The averaged forms for the equation of motion and for the equation of continuity read :

$$\frac{\partial q_x}{\partial x}+\frac{\partial q_z}{\partial z}+\frac{\partial h}{\partial t}=0 \quad (7)$$

$$\rho\left(\frac{\partial q_x}{\partial t}+\frac{\partial I_{xx}}{\partial x}+\frac{\partial I_{xz}}{\partial z}\right)=-h\frac{\partial P}{\partial x}+\mu_h\left.\frac{\partial u}{\partial y}\right|_h-\mu_0\left.\frac{\partial u}{\partial y}\right|_0 \quad (8)$$

$$\rho\left(\frac{\partial q_z}{\partial t}+\frac{\partial I_{xz}}{\partial x}+\frac{\partial I_{zz}}{\partial z}\right)=-h\frac{\partial P}{\partial z}+\mu_h\left.\frac{\partial w}{\partial y}\right|_h-\mu_0\left.\frac{\partial u}{\partial y}\right|_0 \quad (9)$$

The parameter n represents the index of the power law. U_1 and W_1 are the averaged velocities due to the Poiseuille flow. For n=1, the Newtonian cases are obtained. The integrations are performed and substituted in the equation of motion and the equation of continuity. Hence, three non-linear partial differential equations are found. By replacing the terms $\frac{\partial}{\partial t}(hU_1)$ and $\frac{\partial}{\partial t}(hW_1)$ in the equation of continuity by the expressions issued from the equation of motion (see Bou-Said and Ehret [3]) , the following generalised Reynolds' equation is obtained :

$$\frac{\partial}{\partial x}\left(h^3\frac{\partial P}{\partial x}\right)+\frac{\partial}{\partial z}\left(h^3\frac{\partial P}{\partial z}\right)=\frac{2(2n+1)}{n}$$

$$(\mu_h+\mu_0)\left[\frac{\partial h}{\partial t}+\frac{1}{2}\frac{\partial h}{\partial x}U\right]$$
$$-\rho\left[2h\frac{\partial h}{\partial x}\left(\frac{\partial q_x}{\partial t}+\frac{\partial I_{xx}}{\partial x}+\frac{\partial I_{xz}}{\partial z}\right)\right.$$
$$\left.+h^2\left(-\frac{\partial^2 h}{\partial t^2}+\frac{\partial^2 I_{xx}}{\partial x^2}+2\frac{\partial^2 I_{xz}}{\partial x\partial z}+\frac{\partial^2 I_{zz}}{\partial z^2}\right)\right]$$
$$-2\frac{(2n+1)}{n}h\left[U_1\frac{\partial}{\partial x}(\mu_h+\mu_0)+W_1\frac{\partial}{\partial z}(\mu_h+\mu_0)\right]$$
$$-\frac{\partial}{\partial x}(Uh(\mu_h+\mu_0)) \qquad (11)$$

The differences with the Reynolds' equation are mainly due to the appearance of the second term in parentheses on the right hand side. Only this term contains accelerations, it characterises averaged fluid inertia across the film. The first term is unchanged; its factor accounts both variations of viscosity across the film and change in velocity profiles. The two last terms on the right hand side take into account the variation of viscosity along the film.

Short Journal Bearing Case

Realistic short journal bearing assumptions simplify. Short journal bearing theory is commonly used to model the behaviour of crankshaft bearing. It rests on the following conditions:

$$\frac{\partial p}{\partial x}<<\frac{\partial p}{\partial z}\quad U_1=0, U=\text{constant}\quad h=h(x,t)$$
$$\text{and } W_1=z\Omega(x,t) \qquad (12)$$

As we are in the thin film regime, it is reasonable to assume $\partial w/\partial y << \partial u/\partial y$ because of the thin film thickness. The shear rate in the sliding direction ($\sim 10^5 s^{-1}$) is more important than the shear rate in axial direction ($\sim 10^3 s^{-1}$) due to the axial pressure gradient. So we neglect in this study the term $\partial w/\partial y$.

The constitutive equation is taken as the power law. Although this relation is not able to describe thoroughly both Newtonian areas at low and high shear rate, it represents one of the simplest non-linear forms. Only two parameters need to be defined, Barnes, Hutton, Walters [2].

$$\mu_a=K\left|\dot{\gamma}\right|^{n-1} \qquad (13)$$

For a short journal bearing we set $\dot{\gamma}=R\omega/h$.

New kinematics variables are introduced in the journal bearing to locate the shaft centre positions and to facilitate the resolution of the generalised Reynolds' equation, under dynamic loading conditions.

$$\theta=\frac{x}{R}-\phi(t)-\pi/2,\ \tau=\omega t,$$
$$U=R\omega, \text{ and } h=c(1+\varepsilon(t)\cos\theta) \qquad (14)$$

The developments are similar to those performed by Tichy and Bou-Saïd [15]. The radial and tangential hydrodynamic forces are defined as below:

$$Fr=2\int_0^{L/2}dz\int_\alpha^\beta(p-p_a)R\cos\theta\,d\theta$$
$$Ft=2\int_0^{L/2}dz\int_\alpha^\beta(p-p_a)R\sin\theta\,d\theta \qquad (15)$$

As shown by Ehret [4] in our case the term in right hand side in (11) is independent of the z variable. Hence, we can express $\frac{\partial^2 p}{\partial z^2}$ as:

$$p-p_a=\frac{1}{2}\frac{\partial^2 P}{\partial z^2}\left(z^2-\frac{L^2}{4}\right) \qquad (16)$$

Substituting the new variables in (11) yields the following expression of the generalised Reynolds' equation.

$$\frac{\partial^2 P}{\partial z^2}=$$
$$\left[4\frac{(2n+1)}{n}\dot{\gamma}^{n-1}\left[\left\{\varepsilon\frac{d\phi}{d\tau}-\frac{1}{2}\varepsilon\right\}\frac{\sin\theta}{h^3}+\frac{d\varepsilon}{d\tau}\frac{\cos\theta}{h^3}\right]\right.$$
$$\left.-Re^*\left[\left(-\varepsilon\frac{d\phi}{d\tau}+\frac{1}{6}\varepsilon-\frac{d^2\varepsilon}{d\tau^2}+\varepsilon\left(\frac{d\phi}{d\tau}\right)^2\right)\frac{\cos\theta}{h}\right.\right.$$

$$+\left(\frac{d\varepsilon}{d\tau}-2\frac{d\varepsilon}{d\tau}\frac{d\phi}{d\tau}-\varepsilon\frac{d^2\phi}{d\tau^2}\right)\frac{\sin\theta}{h}$$

$$-4\left(\frac{(2n+1)}{3n+2}\varepsilon^2\frac{d\phi}{d\tau}+\frac{15n+8}{6(3n+2)}\varepsilon^2+\frac{4(2n+1)}{3n+2}\varepsilon^2\left(\frac{d\phi}{d\tau}\right)^2\right)$$

$$\frac{\sin^2\theta}{h^2}$$

$$+\left(-4\frac{(2n+1)}{3n+2}\varepsilon\frac{d\varepsilon}{d\tau}+\frac{8(2n+1)}{(3n+2)}\varepsilon+\frac{d\phi}{d\tau}\frac{d\varepsilon}{d\tau}\right)$$

$$\frac{\sin\theta\cos\theta}{h^2}$$

$$-4\frac{(2n+1)}{3n+2}\left(\frac{d\varepsilon}{d\tau}\right)^2\frac{\cos^2\theta}{h^2}\Bigg]\Bigg]\frac{\mu\omega}{c^2} \qquad (17)$$

As previously introduced by Tichy and Bou-Saïd [15] and in Pinkus and Sternlicht's work [11], the importance of fluid inertia is represented by the modified Reynolds number.

$$Re^*=\frac{\rho\omega c^2}{\mu} \qquad (18)$$

The usual Reynolds number is defined by the ratio below. And it indicates the nature of the flow, laminar or turbulent.

$$Re=\frac{\rho\omega Rc}{\mu} \qquad (19)$$

The limits of the cavitation area are marked by the points of angular locations α and β where (16) goes through zero. In the usual theory, this represents the Gumbel's condition. As it is already noted in Tichy and Bou-Saïd's paper [15], for inertia case, β does not necessary equal $\alpha+\pi$.

Thermal Effects

Taking into account the thermal effects needs to solve simultaneously the generalised Reynolds equation, the energy equation for thin film flow as well as the heat transfer to the solids (Bou-Saïd [17]. The basic equations in the case of non-Newtonian flow established in Najji et al [18] do not account for surface accelerations and inertia effects. A simple way to obtain an estimation of the temperature mean value in the contact is possible by the means of energy equilibrium relationship. We can write:

$$(1-\alpha)\,\rho\,C_p\;Q\,\Delta T=C_a\omega \qquad (20)$$

where :

α : percentage of energy discharged in the solids ($\alpha\cong 0.2$)
ρ : fluid density
ω : rotational speed
ΔT : temperature variation
Q : fluid flow
Ca : friction torque
Cp : specific heat at constant pressure
ΔT : increase of temperature. ΔT represents the increase of temperature, i.e. the difference between the running temperature and the feeding temperature of the fluid.
The viscosity variation is explained as

$$\mu=\mu_0\exp\left(\beta\left(\frac{1}{T}-\frac{1}{T_0}\right)\right) \qquad (21)$$

where β is the coefficient of thermoviscosity and μ_0 the viscosity at temperature T_0

Shaft Centre Trajectory with Dynamic Load

The calculation of the shaft centre trajectory results in the integration of the fundamental equations of mechanics. In dimensionless form, these equations are :

$$M\frac{d^2x^*}{d\tau^2}=Fr^*\sin\phi+Ft^*\cos\phi \qquad (22)$$
$$-W^*\cos\psi$$

$$M\frac{d^2y^*}{d\tau^2}=-Fr^*\cos\phi+Ft^*\sin\phi \qquad (23)$$
$$-W^*\sin\psi$$

where

$$x^*=\frac{x}{c},\ y^*=\frac{y}{c},\ M=\frac{12m\omega\,c^3}{\mu\omega\,RL^3}$$

$$Fr^*=\frac{12c^2Fr}{\mu\omega RL^3},Ft^*=\frac{12c^2Ft}{\mu\omega RL^3},$$

$$W^*=\frac{12c^2W}{\mu\omega RL^3} \qquad (24)$$

m is the rotor mass, W is the applied external load at angle ψ. Bearing geometries and running conditions are given in figure 1 and table 1. These data correspond to an actual connecting-rod bearing

used in a commercial petrol engine. The load diagram is given in figure 2. It is interpolated using cubic splines for each two degrees of crankshaft rotation to faithfully represent the load peak. :

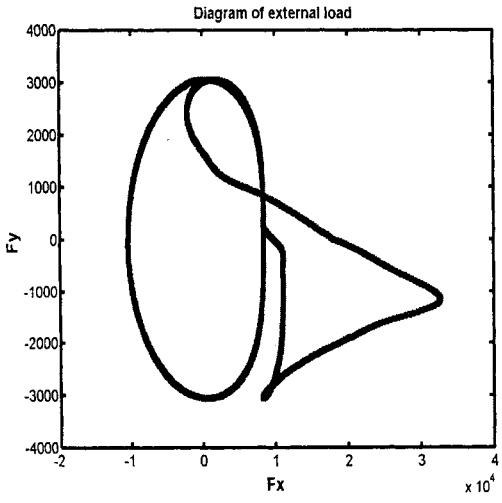


Figure 2 : Diagram of external load

Although this explicit procedure is not difficult to implement, it requires small steps of time to converge.

3. RESULTS

In order to visualise clearly the changes due to the inertia and shear thinning effects different configurations have been tested. The Reynolds number is equal to 356, the flow is thus laminar. The modified Reynolds number Re^* is equal to 4.6. In figures 3 and 4 we note a high modification of the axial flow magnitude when the thermal effects are taken into account for both index power law n=0.8 and n=1.0 .The remperature effect increases the flow. In figure 5 one can notice that the friction torque is widely overestimate in the isothermal case. Few differences between isothermal and thermohydrodynalic regimes are observed for the minimal film thickness and the shaft trajectories (figures 6 and 7). The running temperature is presented figure 8. On average the increase is of 5 °C .This is in perfect concordance with Fantino et al [6] results.

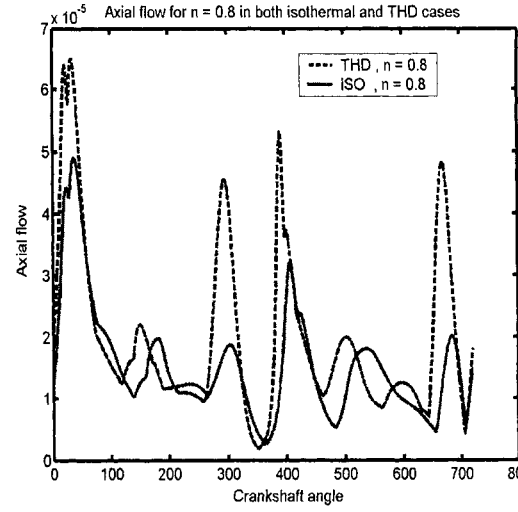


Figure 3 : axial flow

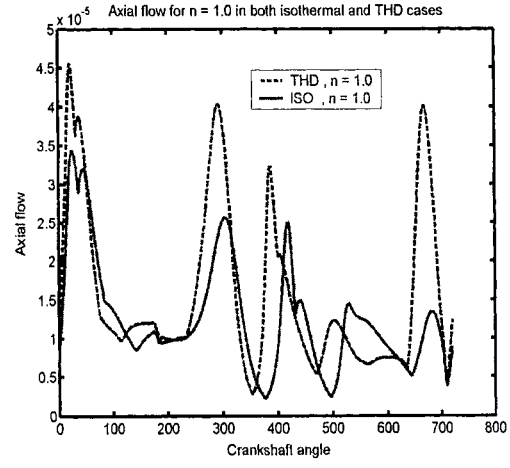
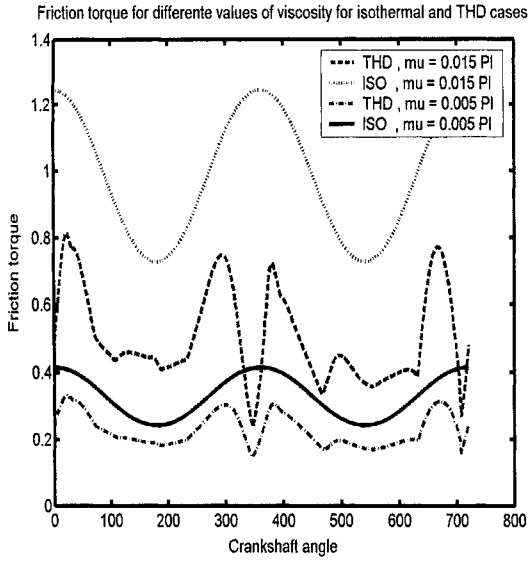


Figure 4 : axial flow n=1



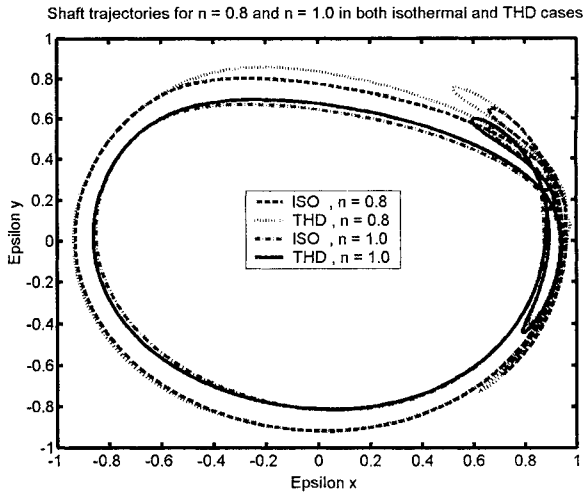


Figure 7 : Shaft trajectories

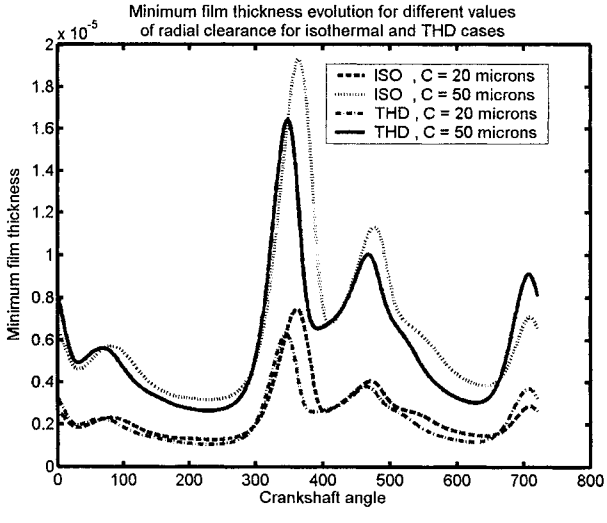


Figure 6 : Minimum film thickness

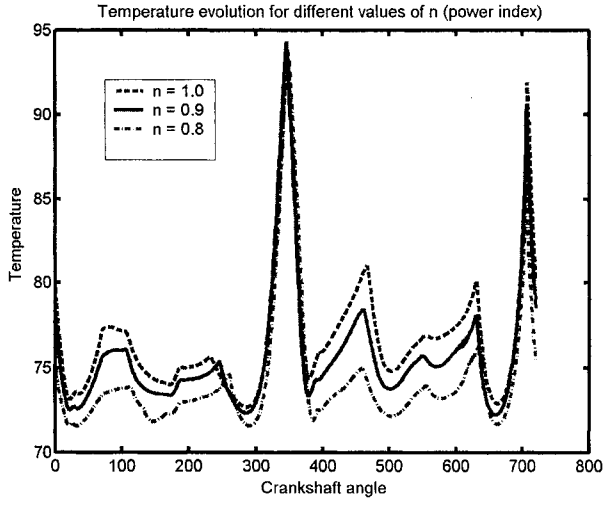


Figure 8 : Temperature evolution

4. CONCLUSIONS

This work is a continuation of the Tichy and Bou-Saïd's theory, which considers convective inertia terms and surface accelerations, and in which shear thinning is also taken into account. The lubricant is described as a power law fluid. One of the main assumptions concerns the use of the velocity

profiles deduced from the power law squeeze film problem. A comparison is performed between the results obtained by the classical theories and running conditions and those computed for lubricants of which power law indexes are in the range of 0.8-1 and inertia as well as thermal effects are present. These represent usual characteristics of additived engine lubricants. Significant discrepancies are observed between different theories. The essential consequence of the thermal effect is the increase of the axial flow and the friction torque while the minimum film thickness, the shaft tracjectories stay still the same compared to the isothermal situation.

References

[1] ASTM Status Report "The relation ship between hight-temperature oil rheology and engine operation"

[2] Barnes, H.A., Hutton, J.F., Walters, K, Rheology series, 3. An Introduction to Rheology, Elsevier, 1989.

[3] Bou-Said B., Ehret P.,"Inertia and shear thinning effects on bearing behavior with impulsive loads", Journal of Tribology-Vol 116, pp535-540, 1994

[4] Ehret P., "Contribution à l'étude du comportement de mécanismes lubrifiés sous chargement transitoire", Thèse Doctorat en Mécanique, soutenue à l'INSA de Lyon le 14.04.1993

[5] Fantino B., Frene J., Du Parquet J.,"Viscosity effects on the dynamic characteristics of an elastic Engine Bearing", S.A.E Technical Paper Series. Printed from SP. 640 Engine Lubrication, 1985.

[6] Fantino B., Godet M., and Frene J., "Dynamic behaviour of an elastic connecting rod bearing. Theoritical study". Proceeding of the Society of Automotive Engines; Studies of Engine Bearing and Lubrication SP. 539 n°830307, 1983, pp. 23-32

[7] Gecim, B.A.,"Non-Newtonian Effects of Multigrade Oils on Journal Bearing Performance". S.T.L.E- Vol. 23 N° 3, 1990, pp384-394

[8] Martin F.A. Developments in Engine Bearings "9th Leeds-Lyon Symposium on Tribology", Leeds England, Sept.,1982, Edited by D. Dowson, C.M. Taylor, M. Godet and D. Berthe. Tribology of Reciprocating Engines, Butherworks Westbury House, 1983, pp. 928.

[9] Ockrent E.H,"The effect of lubricant viscosity and composition on engine friction and bearing wear" ASLE.Trans.vol.4, (97), 1961.

[10] Paranjpe, R.S., 1991, "Analysis of Non-Newtonian Effects in Dynamically Loaded Finite Journal Bearings Including Mass Conserving Cavitation" ASME, Journal of Tribology, Vol. 114, N°4, 1992,pp. 736-745.

[11] Pinkus, O., Sternlicht, B., "Theory of Hydrodynamic Lubrication", MacGraw-Hill book Company, Inc., 465 p., 1961.

[12] Rajalingham, C., Prabhu, B.S., Rao, B.V.A, "The Steady State Performance of a Hydrodynamic Journal Bearing with a Non-Newtonian Lubricant", WEAR, Vol. 55, p. 107-120, 1979.

[13] Rastogi, A., Gupta, R.K., "Accounting for Lubricant Shear-Thinning in the Design of Short Journal Bearing" Journal of Rheology, Vol. 35, n°4, p. 589-603, 1991.

[14] Safar, Z.S., "Dynamically Loaded Bearings Operating with Non-Newtonian Lubricant Films", WEAR, Vol. 55, p. 295-304, 1979.

[15] Tichy, J., Bou-Saïd, B., , "Hydro-dynamic Lubrication and Bearing Behaviour", Tribology Transactions, Vol. 34, n° 4, p.505-512, 1991.

[16] Williams, P.D., Symmons, G.R., "Analysis of Hydrodynamic Journal Bearing Lubricated with Non-Newtonian Fluids", Tribology International, Vol.20, n° 3, p.119-124, 1987.

[17] Bou-Said B. "A global method for thermohydrodynamic problems by finite element method". Symposium on flow thin fluids films, ASME Fluids engineering Spring conference, Cincinatti, Ohio, June 14-18, vol. 48, p. 1-6, 1987.

[18] Najji B., Bou-Said B., Berthe D. "New formulation for lubrication with non-newtonian fluids". ASME-STLE Tribology conference, 16-19 octobre 1988, Baltimore. Journal of Tribology, vol. 111, n° 1, p. 29, janvier 1989.

TABLE 1 Standard geometry and operating conditions

Internal bearing radius R_i	0.027 m
Bearing length L	0.0186 m
Radial clearance C	$35*10^{-6}$ m
Shaft speed N	4500 rev/mn
or ω	471.24 rd/s
Oil dynamic viscosity μ	0.010Pa.s
Oil density ρ	800 kg/m^3
Connecting rod mass M	3.5 kg
Crankshaft throw $L2$	0.03425 m
Connecting-rod length $L3$	0.13055 m
Shear thinning effect n	0.8 to 1
Temperature supply	70°C

This Page Intentionally Left Blank

SESSION XXI
PHYSICAL AND CHEMICAL ACTIONS IN TRIBOLOGY

Chairman : Professor R C Coy

Paper XXI (i) **Experimental Study on the Effects of Sliding Velocity and Roughness Orientation of Metal Surfaces on the Function of Lubricant Additives in Controlling Friction in a Boundary Lubrication Regime**

Paper XXI (ii) **Extraordinarily Low Friction of the Diamond/ H- and Ag- Terminated Si(111) Sliding System**

Paper XXI (iii) **A Novel Experimental Analysis of the Rheology of ZDDP Tribofilms**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Experimental Study on the Effects of Sliding Velocity and Roughness Orientation of Metal Surfaces on the Function of Lubricant Additives in Controlling Friction in a Boundary Lubrication Regime

Masabumi Masuko, Keiji Tomizawa, Saiko Aoki and Akihito Suzuki

Department of Chemical Engineering, Graduate School of Science and Engineering, Tokyo Institute of Technology,
12-1 O-okayama 2-chome, Meguro-ku, Tokyo 152-8552, Japan

We developed a new type of cylinder-on-disk (nominal line contact) tribometer that uses a unique friction data acquisition system to precisely measure friction under boundary lubrication conditions ranging from ultra-low (5μm/s) to moderate (17cm/s) sliding velocity. As the tribometer was designed to measure friction at specific positions on a disk surface, fine distinctions in friction characteristics can be evaluated in relation to the surface condition. Two disk specimens were prepared, one with a polished surface and the other with a ground surface. The polished surface had isotropic low roughness (Ra≈0.01μm) while the ground surface had anisotropic high roughness (Ra≈0.2μm). When the cylindrical specimen slid on the ground surface with anisotropic roughness, the sliding direction changed between transverse and longitudinal as the roughness orientation changed during each disk revolution. Several different base oils and additives were used to examine their performance. When fatty acid was used as an additive, low friction was measured with all the specimens, but remarkably less friction was measured in the transverse direction than in the longitudinal direction or with the isotropic low roughness specimens at the small Λ value region. In addition, friction measured in the transverse direction gradually decreased as sliding velocity decreased throughout the entire range. On the other hand, in the case of longitudinal or isotropic surface roughness, friction increased remarkably as sliding velocity decreased at the very low velocity region. It was also demonstrated that the surface roughness orientation affects friction even at ultra-low sliding velocities, where the hydrodynamic action of the lubricating oils is considered ineffective.

1. INTRODUCTION

The technology of friction control has become a key to reducing the amount of CO_2 evolution—that is, to minimize undesirable energy consumption. Recent trends to reduce the size or enhance the performance of machinery have brought about more-severe operating conditions, such as higher contact pressures, higher operating temperatures, and use of lower viscosity oils to reduce fluid film shear stress. Since these changes reduce film thickness in liquid lubrication, machines often operate in mixed or boundary lubrication regimes. In such conditions, friction control depends not only the nature of material surfaces but also on the performance of lubricating oils (both base oils and additives). Under these mixed or boundary conditions, however, neither lubrication performance nor the lubricant technologies are fully understood.

In many applications, such as automotive engines, oil is required to minimize friction. It has been pointed out that the use of low-viscosity oil is very effective in reducing fuel consumption, and at the same time, application of a friction modifier such as MoDTC becomes important in reducing friction under boundary lubrication conditions. On the other hand, automatic transmission fluid (ATF) or fluids for belt-type continuously variable transmissions (b-CVTF) are often used under mixed or boundary lubrication regimes with appreciable solid contact, and they are required to have a high static friction coefficient, along with no friction reduction with increasing sliding velocity. This unique friction-velocity characteristic can be developed through

advances in material and lubricant technologies. There has, however, been a lack of systematic research on the friction-velocity characteristics of lubricants. To address this need, we present a newly developed tribometer that attains a wide range of sliding velocity change, ranging from ultra-low (quasi-static) to moderate.

Apart from the desire to improve application performance, we are also interested from an academic point of view in the surface effects and lubrication mechanisms of lubricants under mixed or boundary lubrication conditions, especially in the relationship between friction and sliding velocity. In the case of mixed lubrication, surface roughness is a key factor because the solid contact can be determined by the oil film parameter Λ, which is the calculated film thickness divided by the composite surface roughness. The less rough the surface is, the easier it is to attain full film lubrication and thereby reduce solid contact. Therefore, super-finishing to obtain a very smooth surface is sometimes effective for reducing friction.

Many investigators have studied the effects of the solid-surface texture on friction, mainly in full-film EHL or mixed lubrication where the hydrodynamic action of the fluid plays a role [1-3]. The consensus on the effect of surface roughness orientation on friction is that transversely oriented roughness results in a much thicker hydrodynamic film than does longitudinally oriented roughness. However, the effect of surface texture in the performance of lubricants in a boundary lubrication regime is much less understood, and the difference between mixed and boundary lubrication is basically unclear. The reason for this insufficient understanding is hydrodynamic action, including micro-EHL, is not well understood and we cannot yet predict where it should appear. More research on mixed lubrication is necessary [4].

This paper also discusses the effects of hydrodynamic action and surface roughness orientation on the action of additives.

2. EXPERIMENTAL

2.1. Tribometer

Figures 1 and 2 respectively show schematic illustrations of the cylinder-on-disk type tribometer developed for this study and the configuration of the test specimens.

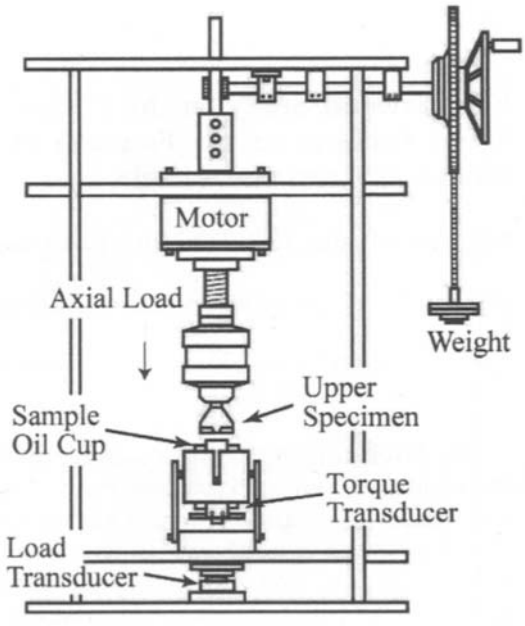


Fig. 1 Schematic illustration of Tribometer

An upper cylinder is attached to a direct-drive pulse-motor through a specimen holder, and a lower stationary disk is mounted in an oil cup. This sample oil cup is fixed in a cup holder suspended with a gimbal mechanism. As the drive motor is allowed to move its position toward the vertical direction, an axial load is applied by pressing the upper specimen to the lower specimen by dead weight through a rack-and-pinion jack. A heating element with a temperature controller is installed on the sample oil cup to keep the sample oil temperature constant.

Friction torque was measured with a piezo transducer installed just under the cupholder. The axial load was measured by a load transducer mounted under the gimbal support. The friction coefficient was calculated using simultaneously obtained values for friction torque and axial load. The pulse-motor and the two transducers were connected via a computer through a pulse-motor driver and A/D converters.

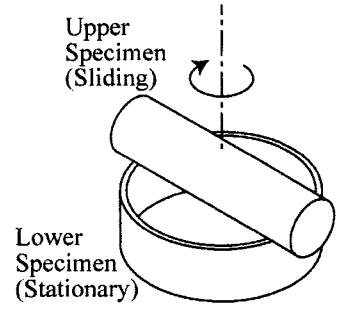


Fig. 2 Test specimens

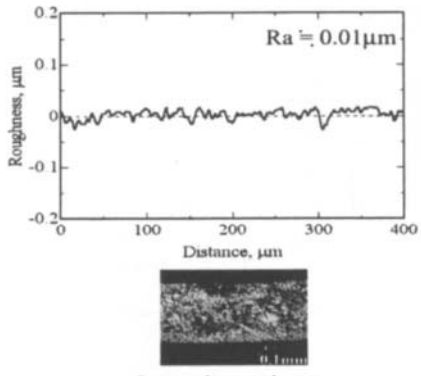


Fig. 3 Isotropic roughness specimen

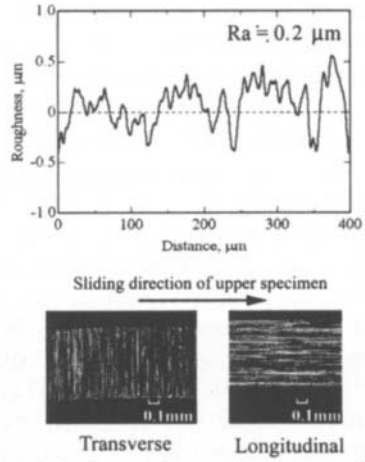


Fig. 4 Anisotropic roughness specimen

2.2. Test Specimens

The upper cylinder specimen, 10mm in diameter and 40mm long, was a roller for commercially available roller bearings. The lower stationary disk specimen was an outer race of a commercially available conical roller bearing. The centerline circle diameter of the friction surface of the disk was 33.5mm, and the width of the contact part was 0.85mm. Both specimens were made of heat-treated high carbon chromium bearing steel. Two kinds of surface finish were applied to the disk specimens. One disk was finished by a surface polisher to 'isotropic roughness' of approximately Ra≈0.01μm, and the other disk was finished by a surface grinder to 'anisotropic roughness', a texture in one direction with a roughness of approximately Ra≈0.2μm. Microphotographs of surface profiles are shown in Figs. 3 and 4. Just before the experiment, both specimens were ultrasonically cleaned in a toluene bath followed by drying and UV-ozone cleaning.

2.3. Data Acquisition System

Table 1 shows the experimental conditions. The tribometer and data acquisition functions were operated by computer software we developed. The operation procedure consisted of three parts: a running-in, low-speed range measurement and high-speed range measurement. For the running-in, the upper specimen was rotated at 29.5 rpm (5.18cm/s) for 100 revolutions under an experimental load and temperature with a testing sample oil.

Four angles, at 90° intervals on the disk specimen (0°, 90°, 180°, and 270°) were chosen for the measuring positions (Fig. 5). Two pairs of opposite positions (0° and 180°) or (90° and 270°) were identical because those positions contacted

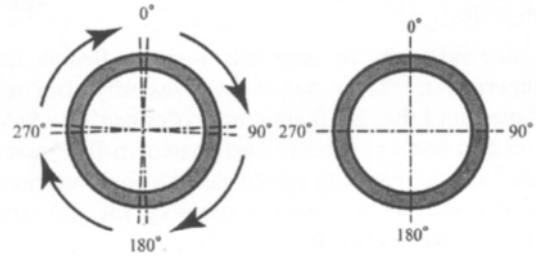


Fig. 5 Friction measuring positions

Table 1 Test specimens and experimental conditions

Test specimen (High carbon chromium bearing steel)	Upper rotating cylinder	d=10 mm, L=40 mm
	Lower stationary disk	d=33.5 mm W=0.85 mm
Axial load (Hertzian mean contact pressure)		143 N (615 MPa)
Sliding velocity		5.18μm/s ~ 17.0 cm/s
Temperature		40 °C

the cylinder specimen at the same time.

For the measurement at low-speed range (5.18μm~2.65mm/s), friction was measured within a range of 5° each of the four positions: 0° (-2°~3°), 90° (88°~93°), 180° (178°~183°), and 270° (268°~273°). At each position, upper specimen slides at a specific velocity for the 5° range and 50 friction data (i.e., one datum at each 0.1°) were collected. The data were averaged and the mean value was used as a representative value of the position at the measured velocity. To shorten the total experimental time, the upper specimen was moved from position to position at 2.95 rpm (5.18mm/s). After two rotations, the measuring velocity was changed and the data acquisition procedures were repeated. The measurement was carried out in two sets, one of velocity 'increment' (5.18μm→2.65mm/s) and the other of velocity 'decrement' (2.65mm/s →5.18μm).

For the measurements at the high-speed range (5.30cm/s ~ 17.0cm/s), the upper specimen was rotated continuously at a specific velocity, and friction data were acquired only when the upper specimen slid to one of the four positions (0°, 90°, 180°, 270°). The measurement was continued for ten rotations and the data were averaged. Then the measuring velocity was changed to the next value. The measurement was carried out with two sets of the velocity 'increment' (5.30cm/s→17.0cm/s) mode.

2.4. Sample Oils

Several kinds of base oils were studied in the experiment. Their names and viscosity data are shown in Table 2. Stearic acid (designated StA) and tri n-butyl phosphate (designated n-BuP) were used as load-carrying additives. The concentrations of the sample oils were 5mmol/kg for StA and 50mmol/kg for n-BuP.

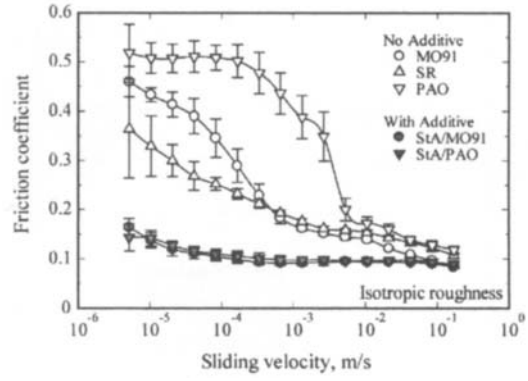


Fig. 6 Experimental results with isotropic roughness specimen

3. RESULTS AND DISCUSSION

3.1. Isotropic roughness specimen

As mentioned above, data acquisition was carried out at four positions. It was found, however, that the four sets of data were almost identical to each other when the disk surface had isotropic low roughness. This might be due to the usage of carefully finished commercially available bearing parts made of well-defined materials. Therefore, the measured data at the four positions were averaged for each sliding velocity and were then compared as a function of the velocity. Figure 6 shows the results. For additive-free base oils, the friction coefficient gradually increased with decreasing sliding velocity. The friction coefficient increased drastically below a certain velocity, which varied depending on the base oil. Oils formulated with stearic acid showed stable low-friction coefficients. Although friction characteristics are not clearly shown in Fig. 6, three regions are recognized with the stearic acid-formulated oils. In the velocity region higher than 2 x 10^{-2} m/s, friction coefficients increased slightly

Table 2 Sample base oils

Oil	CODE	Kinematic viscosity @40°C, mm^2/s	Viscosity @40°C, Pa·s
Hydrorefined paraffinic mineral oil	MO91	91.0	78.6
	MO36	36.4	30.9
	MO18	18.4	15.6
	MO10	9.6	8.2
Solvent refined paraffinic mineral oil	SR	24.4	20.7
Polyalphaolefin	PAO	16.8	13.4

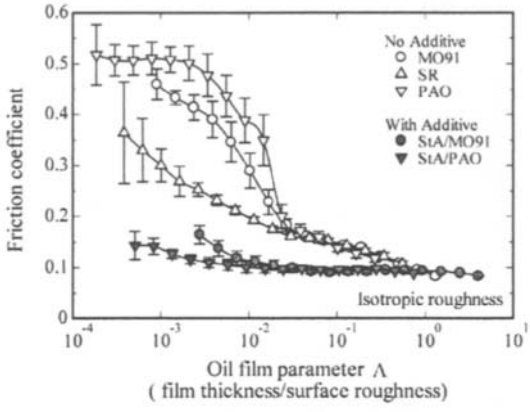


Fig. 7 Results as a function of oil film patameter

with decreasing sliding velocity. From 2 x 10^{-2} m/s to approximately 2 x 10^{-4} m/s, friction coefficients decreased gradually with decreasing velocity. Beyond that point, friction coefficients again increased.

Since the sample base oils had different viscosities, we cannot compare their friction characteristics directly. In the mixed lubrication regime, contact severity was often judged by oil film parameter Λ (calculated EHD film thickness divided by the composite surface roughness after the experiments). Figure 7 shows the results as a function of oil film parameter Λ. The oil film parameter ranged over four orders of magnitude and seemed meaningless when it became less than 10^{-1}. For the three additive-free base oils with different viscosities, however, almost identical results are clearly shown in the right half region, where values are greater than approximately 2 x 10^{-2}. In addition, for the additive-formulated oils, the higher velocity part of the friction characteristic seems to extrapolate the data for the additive-free base oils.

This friction characteristic tells us that the hydrodynamic lubrication effect (macro-EHL effect) is still dominant in controlling friction at the region over 10^{-2} order in Λ value. Compared with both MO91 and PAO, SR had the smallest friction coefficient in the region of small Λ values. This may be attributed to the oiliness effect of polar compounds or aromatics, whose content is higher in SR than in the other oils. A slight difference between StA/MO91 and StA/PAO was observed in Fig. 6 but does not appear in Fig. 7. This also means that the hydrodynamic lubrication effect is valid when stearic acid-formulated oils are used, though the contribution of this effect may be small.

3.2. Anisotropic roughness specimen

As the surface texture of the striped ground scar was set along the axis of 0° to 180°, the sliding direction of the upper specimen on the disk surface at both 0° and 180° was transverse, and that at both 90° and 270° was longitudinal. The results obtained with the anisotropic roughness specimen are shown in Fig. 8. These data differ remarkably from the isotropic data for stearic acid-formulated oil (StA/MO91). The friction coefficient data for the isotropic, transverse, and longitudinal roughness were similar above 10^{-3} m/s, but large differences were observed below 10^{-3} m/s.

The friction coefficient obtained at the transverse roughness position gradually decreased with decreasing sliding velocity, across the entire velocity range measured. At the longitudinal roughness position, however, the friction coefficient increased drastically at low velocity, approaching the data for additive-free oil. For the isotropic roughness, the data were mostly intermediate. The friction characteristics of the additive-free base oils did not differ remarkably among three positions. Since the mean surface roughness differed between the isotropic and anisotropic surfaces, the friction characteristics cannot be compared directly. In addition, since surface damage was apparent when additive-free oil was used, particularly in the lower velocity region, it became difficult to assess the effects of roughness orientation in detail. Therefore, further discussion on the roughness orientation refers to only the additive-formulated oils.

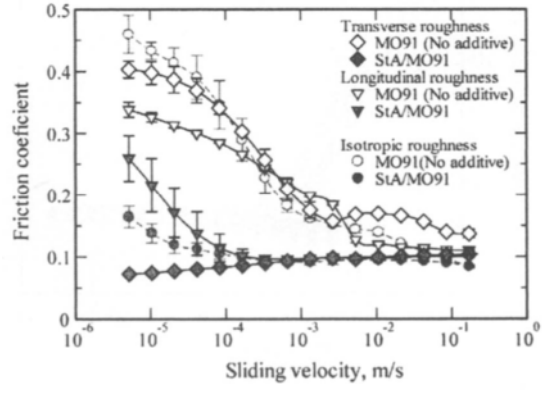


Fig. 8 Results with anisotropic roughness

To examine the effects of viscosity, four kinds of hydro-refined mineral oil with different viscosities were used as base oils. The results with transverse roughness and those with longitudinal roughness are compared in Figs. 9 and 10, as a function of velocity and oil film parameters, respectively. From these figures, it can be accepted that the friction characteristic is basically the same for all of the base oils despite their different viscosities. It is also recognized that transverse roughness produced small friction coefficients below the Λ value of 1×10^{-2}, particularly below 10^{-3}.

It has been well accepted that stearic acid forms an adsorbed molecular layer on its surface, and this organic layer can support part of an axial load. Long straight alkyl chains are very effective for forming oriented molecular films to support large loads and thereby reduce metal-metal contact. In addition, the low shear strength at the contacts between alkyl chains reduces friction and wear.

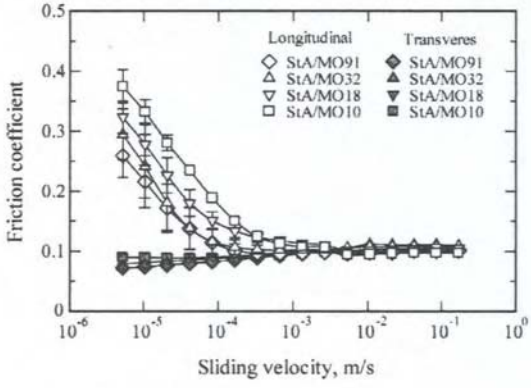


Fig. 9 Results with stearic acid

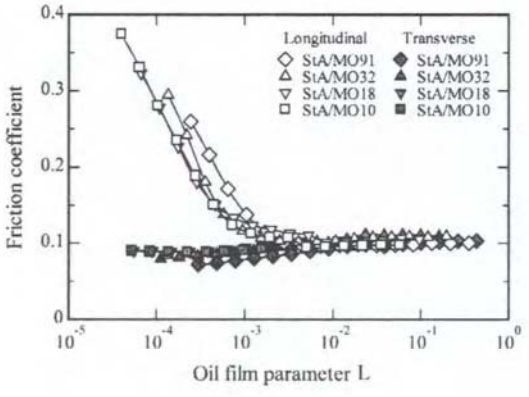


Fig. 10 Comparison with transverse and longitudinal

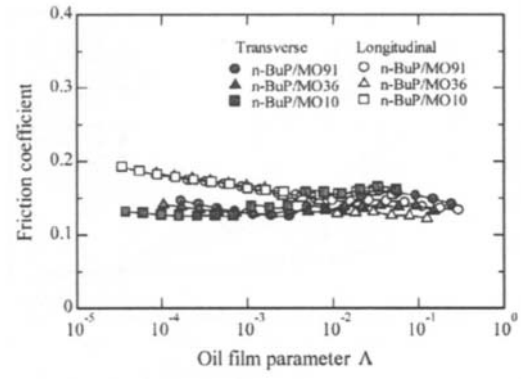


Fig. 11 Results with n-BuP on anisotropic roughness

Another type of load-carrying additives is EP additives, which can react with metal surfaces to produce inorganic surface films. These films protect sub-surface metal and thereby reduce metal-metal contact. In this study, tri n-butylphosphate (n-BuP) was used as a representative of this type of additive. n-BuP forms a metal phosphate layer on the surface tribochemically. It is also believed that the contribution of the organic portion as an adsorbed molecular layer cannot be expected from the n-butyl group, because the chain length is too short. Therefore, the major contribution of n-BuP to lubrication is the formation of an inorganic surface layer.

The n-BuP results at the transverse roughness position and the longitudinal roughness position are compared in Fig. 11. The results for both StA and n-BuP are compared as a function of oil film parameters in Fig. 12 (transverse roughness) and Fig. 13 (longitudinal roughness).

In the case of transverse roughness, the friction coefficients were smaller with StA than with n-BuP across the entire velocity region. In the case of longitudinal roughness, however, StA showed smaller friction coefficients than n-BuP in the large Λ value region, but the inverse was found in the smaller Λ value region.

In addition, it is shown that a positive slope of friction-velocity characteristics is obtained only with a combination of StA and transverse roughness. The mechanism to show a positive slope of friction-velocity characteristics with StA was proposed by the authors in previous studies [5, 6].

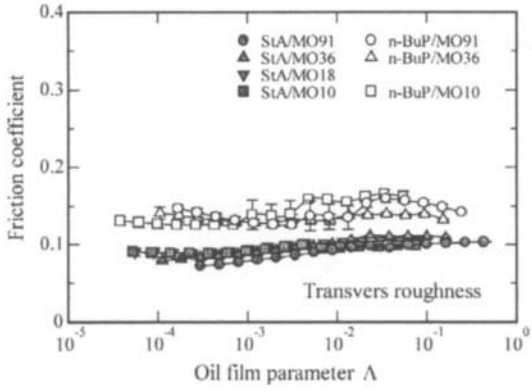


Fig. 12 Comparison with StA and n-BuP on transverse roughness

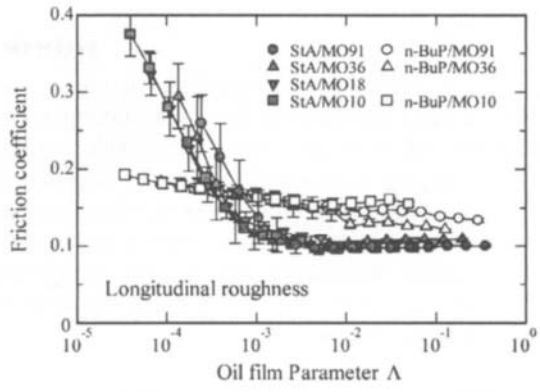


Fig. 13 Comparison of result on longitudinal roughness

3.3. Macro-EHL effect and micro-EHL effect in boundary lubrication conditions

EHL film thickness, used for the calculation of the oil film parameter, is a function of the lubricant viscosity at the inlet of the contact, the pressure-viscosity coefficient of the oil, and the sliding velocity. Therefore, the oil film parameter can be regarded as a parameter of the macroscopic EHL effect. Figure 7 clearly shows the presence of this effect, because, as explained above, all of the results can be summarized by the oil film parameter.

As shown in Fig. 10, it is very interesting that the adsorbed molecular layer lost its load-carrying capability at some small Λ value with longitudinal roughness, even though it could effectively support the axial load with transverse roughness, where the macroscopic EHL effect seemed the same. The authors consider that the intrinsic ability of an adsorbed molecular layer for lubrication appeared in longitudinal roughness and that some supporting mechanism was present in transverse roughness. When the lubricant was stearic acid, as soon as the adsorbed molecular layer lost its load-carrying capability, metal-metal contacts and some seizure-like phenomenon occurred, similar to the case with the additive-free oil.

As shown in Fig. 11, n-BuP as an additive formed an inorganic layer that could support more load than the adsorbed molecular layer, thus protecting the metal surface and preventing local seizures. Therefore, no appreciable friction increase occurred. The gradual increase in friction with decreasing velocity might be due to the reduction of the macroscopic EHL effect of the oil. With n-BuP, transverse roughness also showed a lower friction coefficient than longitudinal roughness.

In the case of transverse roughness, lubricating oil trapped between asperities in the contacting zone hardly flowed out. This was because the sliding direction was perpendicular to the direction of the striped ground scar, thus generating some additional pressure to support the load or to reduce the contact severity. This is a sort of microscopic EHL effect. On the other hand, in the case of longitudinal roughness, the striped ground scar ran in the sliding direction, allowing trapped oil to flow out easily and thus preventing the microscopic EHL effect.

The question is whether or not 'micro-EHL' plays an important role under such ultra-low sliding velocities. This point is still unclear, but the following simple calculation supports a possibility. When axial load is constant, EHL film thickness is determined by the oil viscosity at inlet, the pressure-viscosity coefficient, and the sliding velocity. In this paper, the hydrodynamic capability of fluid is judged by viscosity times sliding velocity ($\eta \times u$) for simplification. For macro-EHL, inlet viscosity is that occurring under atmospheric pressure. However, for micro-EHL, it is regarded as high-pressure viscosity in a Hertzian contact zone. Therefore, if we apply a simple Barus equation to predict high-pressure viscosity, the value of sliding velocity u times viscosity η under contact becomes 85 when u = 5μm/s, and this is almost equivalent to u = 1m/s for macro-EHL (η for atmospheric pressure).

4. CONCLUSIONS

A new type of cylinder-on-disk tribometer was developed that can measure friction accurately under conditions of boundary lubrication with varying velocity ranging from ultra-low (5μm/s) to moderate (17cm/s). An isotropic roughness specimen and an anisotropic roughness specimen were used to study the boundary lubrication performance of several lubricants. The primary conclusions at this stage of the study are as follows:

With the isotropic roughness specimen;

1) Friction characteristics of oils having different viscosities can be summarized into a single line in the larger oil film parameter region.
2) The macroscopic hydrodynamic effect of lubricating oil plays a role in friction, even in the small oil film parameter region.

With the anisotropic roughness specimen;

1) Friction with load-carrying additives at the transverse roughness position was smaller than that at the longitudinal roughness position in the small oil film parameter region.
2) At the longitudinal roughness position, stearic acid lost its load-carrying capability in the small oil film parameter region, resulting in a drastic friction increase.
3) At the longitudinal roughness position, tri n-butyl phosphate kept its load-carrying capability across the entire region tested, and its friction coefficient increased gradually with decreasing sliding velocity due to the reduction of the macroscopic hydrodynamic lubrication effect of the oil.
4) At the transverse roughness position, stearic acid maintained its load-carrying capability, and its friction coefficient decreased with decreasing sliding velocity across the entire velocity region tested.
5) At the transverse roughness position, tri n-butyl phosphate showed a lower friction coefficient than at the longitudinal roughness position, but its coefficient was higher than that of stearic acid.
6) The friction characteristic with transverse roughness is considered to be affected by the micro-EHL effect of the oil.

REFERENCES

1. N. Patir and H. S. Cheng, Proc. 5th Leeds-Lyon Symp., (1979) 15.
2. Y-R. Jeng, Tribology Trans., 33, (1990) 402.
3. D. Zhu and Y-Z Hu, Tribology Trans., 44 (2001) 391.
4. H. A. Spikes & A. V. Oliver, Proc. 13th International Colloquium on Tribology, Esslingen, (2002) 19.
5. H. Okabe, M. Masuko and K. Sakurai, ASLE Trans., 24 (1981) 467.
6. M. Masuko, K. Moriki and H. Okabe, Journal of Japan Soc. Lub. Eng. "Junkatsu", 30 (1985) 422.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Extraordinarily Low Friction of the Diamond/H- and Ag-Terminated Si(111) Sliding System

H. Masuda and F. Honda

Toyota Technological Institute
2 Hisakata, Tempaku, Nagoya 468-8511, Japan

Hydrogen-terminated Si(111) surfaces have been studied for tribological performance in the sliding of a diamond slider to obtain fundamental information on contact problems between a Si single crystal and a diamond slider. We focused on the tribological role of the atoms on the contact surfaces, and observed an extra-ordinarily low coefficient of friction 0.003, of the relevant sliding surfaces, a coefficient as low as that of the sliding system of a diamond/Ag monolayer/Si(111). The sliding tests were carried out under an ultra-high vacuum, lower than 1.5×10^{-7} Pa to avoid contamination, with a load of 250 mN and 0.1 mm/sec of speed. The Si(111) surface was H-terminated by a hydrofluoric acid solution, and by hydrogen gas adsorption before the sliding tests. By sliding both of the H-terminated Si surfaces, a very low and stable friction coefficient was observed, while the desorbed surface, which had been heated to 600℃, showed a very high friction of 0.4 or higher. The observed friction reduction of 2 orders was attributed to a monolayer of hydrogen atoms terminating the Si (111) dangling bonds.

The results are discussed in terms of the sliding mechanism, surface source of friction, and design of the sliding system in a vacuum.

1. INTRODUCTION

The extra ordinarily low coefficient of the friction of the diamond / Ag monolayer / Si(111) sliding system has been reported under vacuum conditions [1-3]. This phenomenon is nearly comparable to the super lubricity phenomenon of the atomic layers of the traditional solid lubricant MoS_2 [4], but the as-deposited Ag layer is not order-structured and the average amount of Ag is about 1-3 monolayers. The sliding mechanism seems different from the crystal plane sliding of MoS_2 [4]. Details of the atomic level interactions and the role of surface atoms on the macroscopic sliding interface are, however, poorly understood due to the lack of information regarding the individual atoms on the contact surfaces. To understand how atomic adhesive force affects the total friction force, and how atomic level surface structures are involved in macroscopic sliding performance, we have deposited 1-3 monolayers of Ag on Si(111), and slid a diamond sphere (3 mm in radius) under vacuum conditions [2]. The results show a drastic decrease in the friction force (2 orders lower) compared to a bare Si(111) surface without Ag. We tentatively estimated that the atomic level interaction force between contact surfaces strongly affects the friction force, i.e., the very low interactive force between C (diamond or H terminating the diamond surface) and Ag atoms worked to reduce the macroscopic friction force, and formed a sliding plane between them. The 1-3 monolayers of Ag were sufficient to induce sliding with a very low friction force over 1000 cycles of reciprocal sliding. In this sliding system, the adhesive force described by Buckley [5], Yoshizawa and Israelachivili [6], and Gardos and Ravi [7] was not negligibly small under our experimental conditions. In the sliding model proposed from the Ag/Si system, an H-terminated Si(111) surface could be estimated to exhibit a similar friction performance, because H-terminated Si may have a very low bonding force with a diamond C or H-terminated diamond surface. This paper describes the experimental results of a diamond slider on

hydrogen-terminated Si(111) surfaces and gas-adsorption experiments, and reports that the system showed as low a friction as estimated from the sliding model of the Ag monolayer/Si(111) system. The results may provide some tips for understanding a fundamental function of surface atoms on the sliding surface, designing low friction sliding surfaces of solids, micro machine sliding systems using an Si surface, and sliding systems used in vacuum space laboratories.

2. EXPERIMENT

2.1 Procedure

The systematic diagram in Fig.1 depicts the experimental apparatus used for sliding tests in an ultra-high vacuum condition. It basically consists of a Si wafer specimen through which a direct current was supplied to heat the Si wafer, an Ag evaporating source by electron beam bombardment, quartz-oscillation film-thickness monitor, W-filament to dissociate hydrogen gas to adsorb the atomized hydrogen, and a slider arm with a diamond spherical pin which moves reciprocally.

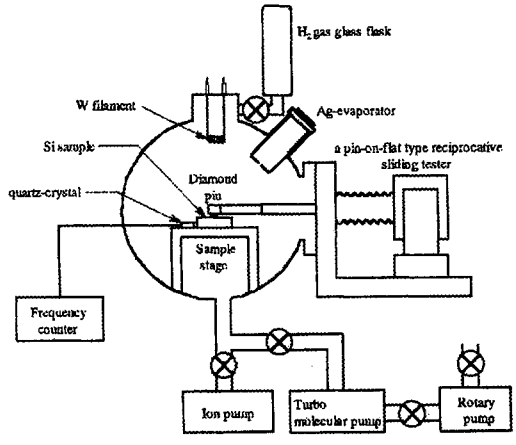


Fig. 1. Schematic diagram of the experimental apparatus.

The entire system was evacuated using oil-free pumps to 10^{-8}Pa. The ultra-high vacuum (UHV) condition was necessary to obtain a well-defined clean Si surface and to maintain the clean state during the sliding experiments. A normal load was applied to the diamond/Si contact point and was adjusted to handle a constant load from outside the vacuum chamber. The temperature of the sample specimen was measured using an optical thermometer through a silica view-port on the vacuum chamber and calibrated to an accuracy of within 10℃.

A cleaned Si(111) surface was obtained in UHV by degassing at 600℃ for 20 hr., mainly to desorb water and hydrocarbon molecules on the specimen surface, and finally by flashing several times at approximately 1200℃ to evaporate oxides of Si. A hydrogen-terminated Si(111) surface was prepared using two independent methods. One surface was prepared by immersing the specimen in a 50% HF + 40% NH_4F (1:4) aqueous solution (buffered hydrofluoric acid: BHF) [8,9] for one minute, followed by repeated rinsings at room temperature. Another hydrogen-terminated surface was prepared by dissociating H_2 molecules using a thermal W-filament and adsorbing the atomized hydrogen on the cleaned Si surface in a vacuum. The hydrogen gas of 99.9999% purity was dissociated and adsorbed on the Si, by an exposure of 2000 L (Langmuir). By each procedure, H atoms are known to terminate the dangling bonds of Si(111) and to passivate the Si surfaces.

Sliding experiments were carried out at room temperature in a UHV condition of less than $1x10^{-7}$ Pa throughout the experiment. A normal load of 250 mN was applied on the Si wafer by a spherical diamond slider (3 mm curvature radius), and the sliding speed was 0.1 mm/sec controlled by a pulse-drive motor. The slider was driven at 100 pulses per second. In this contact condition, approximately 350 MPa of Hertzian pressure was applied at maximum on a circular contact area 32 μm in diameter, as determined by our calculations. In the observed tracks, the width was less than 40 μm, which was close to the calculated width after 1000 cycles of reciprocal sliding, so that the lateral fluctuation of the contact point was stable enough to assume that the iterated sliding was in contact approximately on the same track.

The slider arm was stiff enough not to buckle or twist during sliding. A double-leaf spring was used

and the calibration result of the load/output voltage of the strain gauge is shown in Fig. 2. A straight line was observed down to a lowest load of 6 μN. We determined that the lowest reliable and friction coefficient of our tester was 0.003, a value traditionally acceptable for calibration using load/output-voltage measurements.

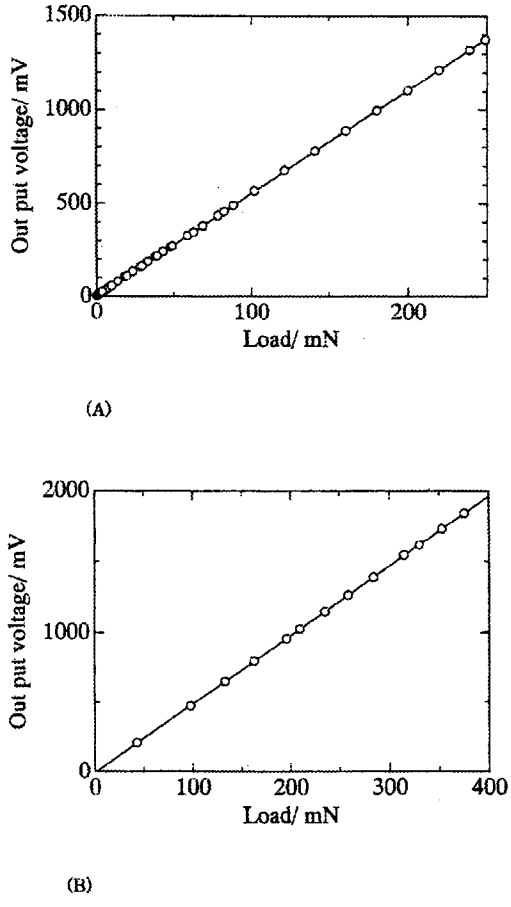


Fig. 2. (A) Calibration curves for the friction force and (B) for the normal load measurements.

In the very low friction range of friction measurements, however, vibration of the slider arm/pin due to the pulse-drive movement makes the apparent friction force lower as a function of the response time constant and minimum sensitivity of the strain gauge to the applied stress. The apparent friction force needed to be corrected subtly due to the vibration effects to obtain the absolute value of the friction force in the very low friction force range. However, further discussion is needed on how to express a coefficient of friction as the most reliable value in very low friction range, which would prove more accurate than the traditional expression of coefficient of friction. The observed results of friction force are presented without correction in this report, following the conventional definition of the coefficient of friction.

2.2 Test specimen

The Si(111) was used in the sliding direction <-2,1,1> of a spherically polished, natural diamond (111) edge. The contact point of the diamond edge was verified by the Laue diffraction pattern to within 1 degree of error from the exact crystal orientation. The surface asperity of the diamond sphere was 0.05 μmRa, as a result of the non-contact high-accuracy roughness measurements. The surface of the diamond was cleaned with dilute nitric acid solution and distilled water, but an atomically well-defined clean surface was difficult to obtain in our vacuum system. During some initial cycles of the sliding experiments, therefore, a possible effect of surface contaminants was unavoidable, but we assumed the limited amount of contaminant on the small contact area of the diamond (circular area of about 32 μm in diameter) to be negligibly small compared to the Si surface in contact (32 μm x 5 mm).

3. RESULTS AND DISCUSSION

3.1 Friction force of H-terminated Si(111)

To understand the effect of a single layer of H atoms on Si, clean Si was used to determine the basic sliding performance of the diamond slider. After cleaning the Si wafer in a vacuum chamber, we observed the coefficient of friction as shown in Fig. 3, upper line (Si substrate). Results were shown to be reproducible by repeated runs of the experiment and were substantially equivalent to previous reports [2,3]. From the beginning of sliding, the coefficient of friction was as high as 0.5 on average, and debris totally covered the track after 100 cycles of sliding due to the strong shearing

stress. The cleaned Si(111) surfaces and also oxygen-covered surfaces are shown to have rather high coefficients of friction by our experimental results. In Fig. 3, the lower line shows the friction force of Ag 1.6 ML/Si, which has very low friction near our detection limit.

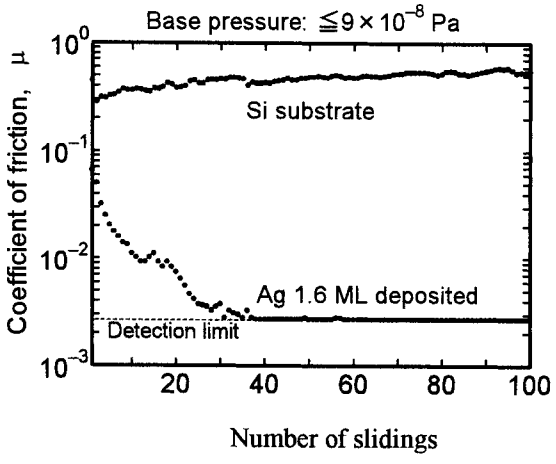


Fig. 3. The coefficient of friction of the clean Si(111) substrate surface (upper line) and of the Ag 1.6 monolayer film deposited on the clean Si(111) surface (lower line).

Figure 4 shows the coefficient of friction of an H-terminated Si(111) surface prepared by the BHF solution. To eliminate the probable effect of the water and adsorbed hydrocarbon molecules, the specimen was heated to 400℃ for 20 hrs to desorb these adsorbates prior to the sliding tests. From the initial stage of the sliding experiments, excluding the first several cycles, the friction force was as low as the level of Ag films. The lowest coefficient of friction virtually corresponded to our lowest detection limit. The low friction of sliding lasted at least 1000 reciprocal cycles. An apparent difference between the results of Ag-deposited and H-terminated surfaces was that at the beginning of some cycles of the Ag as-deposited surface of over 1 monolayer had a rather high coefficient of friction that gradually decreased with the increasing number of cycles. The as-deposited Ag-film is known to form 3-dimensional islands when it is deposited over 1 monolayer, and some fraction of the deposited surface is still found as clean Si (7 x 7) structure domains. The surface hydrogen atoms terminating the Si surface do not adsorb over one monolayer, although mono-, di-, and tri-hydrated Si coexist on the BHF treated surfaces [9]. The effects of the initial 3-dimensional structures are probably reflected in the initial cycles of sliding, during which the microscopic islands of Ag atoms are expanded and displaced along the Si surface. After the Ag atoms were uniformly smoothed, the final coefficients of friction were nearly equivalent in both systems.

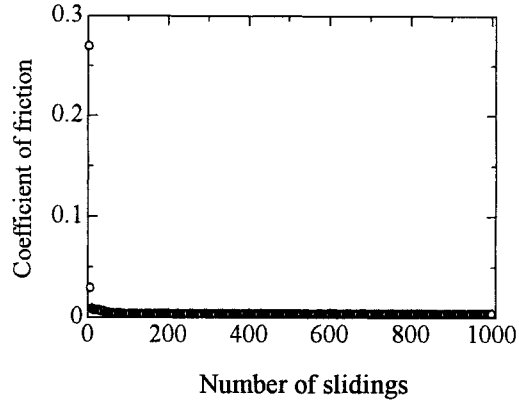


Fig. 4. The coefficient of friction of an H-terminated Si(111) surface by buffered hydrofluoric acid solution treatment. The abscissa indicates number of reciprocal sliding cycles.

One of the procedures to confirm the role of H atoms is the desorption effect of hydrogen on the sliding performance of the BHF-treated Si specimen. Since hydrogen atoms are reported to desorb between 500 – 550°C [10], the specimen was heated to 600°C in a UHV chamber. The results of the sliding experiment indicate a very high friction force, as indicated in Fig. 5. The drastic change of the coefficient of friction by heating was, therefore, attributed to the hydrogen desorption.

To verify the phenomenon by which the friction force is reduced on the H-terminated surfaces, another process of H-termination was used to prepare test-specimens before the low friction phenomenon could be attributed to hydrogen atoms

on the surface, because our apparatus could not observe exact 2-dimensional H/Si atomic confirmation or determine H abundance on the surface quantitatively. Hydrogen gas molecules are not known to adsorb to Si in the molecular state, and W-assisted dissociation of hydrogen molecules to ions is allowed to adsorb on Si [11]. By exposure to 2000 L of hydrogen, the Si surface was estimated to have been covered by hydrogen, according to the results [12]. The friction tests on the adsorbed surface proved that sliding exhibited a very low friction force, as shown in Fig. 6.

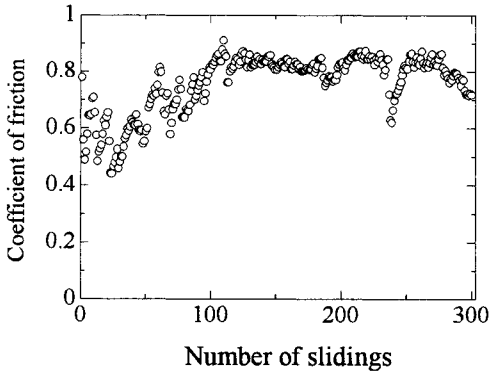


Fig. 5. The coefficient of friction of the BHF-treated Si(111) surface and desorbed the adsorbed H-atoms, by heating to 600℃ under ultra-high vacuum.

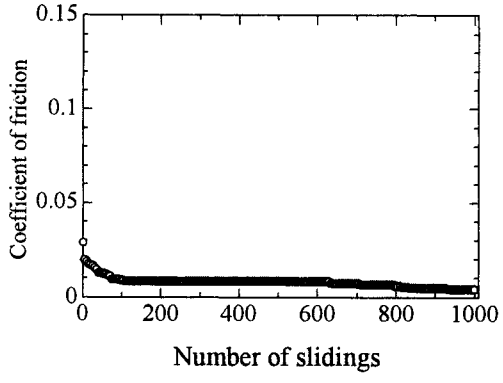


Fig. 6. The coefficient of friction of the H-terminated Si(111) surface prepared by H_2 molecule chemisorption. W-filament assisted the dissociation of H_2 in the vacuum chamber.

These results were approximately equivalent with those obtained from the BFH-treated surfaces as shown in Fig. 4. After the sliding tests, hydrogen atoms were desorbed again at 600℃, following the process of BHF treatment. Very high friction again appeared after desorbing the hydrogen atoms, as shown in Fig. 7. Due to the high friction forces, the contact surfaces were worn and wear debris covered the entire contact surface area. These surface features were totally different from the sliding surfaces with very low friction of which a clear contrast in SEM images is observed, but no worn particles were found in or out of the tracks. This indicates that the microscopic surface morphology was obviously deformed plastically and was detectable enough to be observed by SEM, but the bulk contact pattern of diamond/Si was macroscopically an elastic contact condition. In fact, the slide tracks were not detectably deformed in the SEM images but the trails were simply seen in contrast by the dark area of the contacted part. As a result of these sliding experiments on a H-covered Si surface, the surface hydrogen atoms were found to contribute dramatically in reducing the friction force at least within the tested sliding conditions.

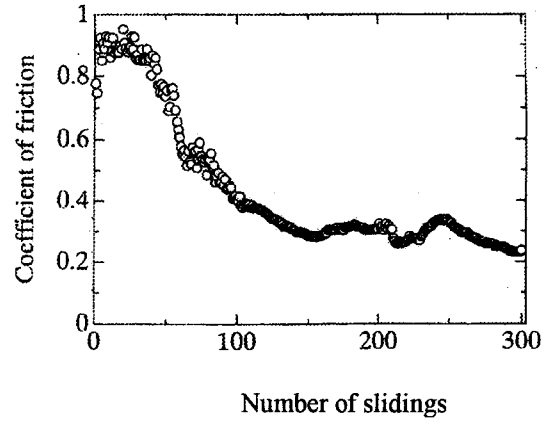


Fig. 7. The coefficient of friction of the H-chemisorbed Si(111) surface and desorbed the hydrogen by heating under ultra-high vacuum.

3.2 Effects of chemical state on surface atoms

From the above results, we confirmed that surface H atoms strongly contribute to the friction reduction of H-Si. Surface oxygen atoms on Si worked to oppose the contribution, i.e., the friction coefficient of an O-covered surface of Si had very high friction, and sliding experiments of Ag/Si in an O_2-containing atmosphere produced a fatal effect of sliding under low friction, as shown in Fig. 8. If the partial pressure of molecular oxygen exceeds 10^{-6}Pa in the atmosphere, oxygen molecules exert a strong effect to increase the friction force, according to the experimental results. This is due to an increase of oxygen adsorption by Si atoms exposed to the vacuum space when the surface Ag atoms are removed during sliding.

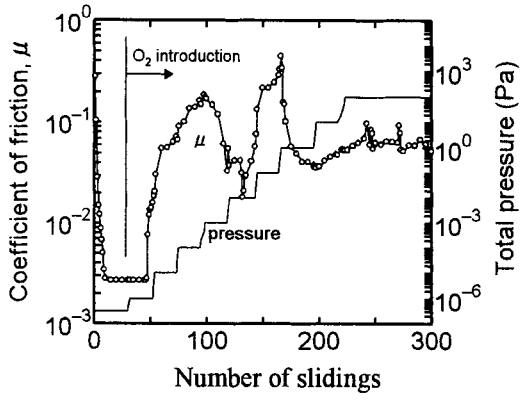


Fig. 8. The coefficient of friction of Ag-deposited Si(111) as a function of the O_2 gas pressure in the atmosphere. Oxygen pressure increased stepwise during sliding.

When water vapor was introduced over the Ag/Si sliding surface stepwise up to 10^{-4}Pa, the observed results indicate that the adsorbed water molecules did not affect low friction sliding performance, at least within the tested pressure region (Fig. 9). These results are comparable to those from the oxygen-containing atmosphere. We could not determine which atom in the H_2O molecule is coordinated with Si atoms, but assumed that O is coordinated to Si and H atoms are opposed to the sliding diamond surface, because the O atom has a lone-pair electron available to bind with the Si dangling bond. The oxygen atoms beneath the topmost atoms were not considered to effect the increase of the friction force.

The chemical state of the adsorbed molecules of -H, -O, and H_2O must be considered an important point among friction-determining factors.

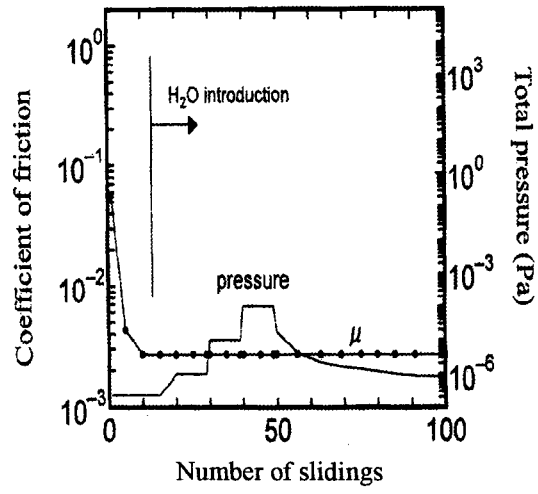


Fig. 9. The coefficient of friction of the Ag-deposited Si(111) surface as a function of water vapor pressure in the atmosphere.

3.3. Ag on H-terminated surface

H-termination and Ag deposition procedures were successful in reducing the friction forces in the use of H-terminated Si surfaces to a sliding plane, but each had characteristic shortcomings. H-termination by BHF is easily prepared by aqueous solution under atmospheric conditions, but it is hard to recover the desorbed defects on the surface if the H atoms were desorbed during sliding. The Ag atomic layers must be formed on a perfectly clean Si surface under UHV conditions. Sliding experiments were carried out for the Ag film deposited on the H-terminated Si, with the expectation of a low friction sliding of Ag layer on the easily prepared H-terminated surface. Results of the sliding tests indicate that the friction force was low compared to H-terminated Si, and the

low-friction sliding lasted over 1000 cycles. We conclude that BHF treatment is a useful process for the preparation of clean substrates in atmosphere for Ag thin films to slide with very low friction force in a vacuum. This result provides a useful tip to prepare a sliding surface using a much simpler method than the flash-heat cleaning of a Si(111) surface under UHV, although other merits such as longer duration life have not been examined yet.

3.4 Model of the sliding mechanism on H- and Ag-Si

By summarizing the observed results, we can discuss a plausible mechanism for the extraordinarily low friction force of the relevant sliding system.

(1) Ag film of 1-3 monolayers on a cleaned Si(111) surface reduced friction drastically.

(2) Ag on an O-covered Si surface did not work as a lubricant layer due to a non-uniform deposition layer of Ag on Si.

(3) Oxygen molecules strongly disturbed the low-friction sliding in the atmosphere over the sliding Ag–deposited surface.

(4) Hydrogen-terminated Si(111) surfaces showed very low friction, as low as Ag-deposited surfaces.

(5) Direct contact of diamond and cleaned-Si has very high coefficient of friction: 0.6.

(6) Contact of diamond and oxidized-Si also has rather high friction: 0.4

(7) N_2 and H_2O molecules in the atmosphere of the sliding chamber do not disturb low-friction sliding.

Using a hypothetical approach, a mechanism for the observed low friction performance in our limited experimental results can be proposed. Obviously, the surface atoms of Ag and/or H have occupied the Si dangling bonds and reduced the chemical reactivity of the bonds. These surface atoms have a low bonding energy with the diamond slider surface atoms, according to the traditional bonding force between C(diamond) – H on Si, or C(diamond) – Ag on Si. Recent studies describe the diamond surface as covered with H atoms to fill the dangling bonds, so that we may suppose H(diamond) – H, and H(diamond) – Ag bonds between the contact surfaces, although we could not determine the H concentration on the diamond. The very low friction forces were previously attributed to the atomic layers that form the sliding plane of a very low shearing force. This mechanism is plausible, and renders the adhesive force term as the atomic-level friction source to be very small, which should be considered in addition to the term proportional to normal load to the contacted area, as expressed by the following equation, after reference [6],

$F = Fa + \mu Fn$, $Fa = cA$,

where, F represents the total friction force, Fa the fraction of the friction force due to the adhesive force, Fn the normal load, μ the constants, A the contact area, and c a constant characteristic to the atomic pair on the contact surfaces.

This model explains well all the experimental results. Result (2) the result is explained, since Ag does not deposit uniform 2D Ag thin layers on O-covered Si surface. Oxygen atoms are mostly exposed to the vacuum space atmosphere and most have one unoccupied bond. Another bond of the O atom is used to bond to the adjacent surface Si atom so that the low shear force plane does not remain. If the Ag atoms dislocate during sliding, the clean Si atoms may react with O_2 molecules, and oxygen atoms may be trapped during sliding by the dangling bonds of the Si. When the sliding tests were carried out in a high vacuum, we predicted the Si atoms on the contacted surface would be recovered by Ag atoms again due to a surface diffusion process because the low friction sliding lasts at least through 1000 cycles of reciprocal movement. Without this self-recovering process, reciprocal sliding of over 1000 cycles would have to make contact at a virgin contact point at all times. A surface recovering mechanism is suggested by experimental result (3), since the surface Ag atoms are displaced during sliding with a low shearing force.

Another mechanism is that the surface adatoms are mobile for some distance along the surface due to very low energy. This restriction is satisfied by the Ag/Si system because the as-deposited Ag atoms are mobile even at room temperature, and the energy barrier to dislocate the surface Ag atoms to the neighbouring adsorption site is very low [13],

compared to the bonding energy of Ag-Si. If moving surface atoms along the surface is acceptable, then the surface layer performs as a liquid layer forming a low shearing plane at the interface. This sliding mechanism has not yet been disproved. In the case of an H-terminated surface, the necessary condition for low friction is satisfied: H atoms stabilize the reactivity by occupying the surface dangling bonds. Atomic mobility of H is difficult to estimate from the paucity of observed data concerning the atomic migration energy of H atoms. In this case, hydrogen molecules are included even in the high vacuum atmosphere, and hydrogen can be supplied to the desorbed points during sliding. O and Si atoms on surfaces are, on the other hand, chemically very active and retain a free bond on –O-Si and –Si-Si.

The results indicate the atomic level interaction force is reflected directly in the macroscopic sliding system adopted in this experiment. These results provide information to help understand the microscopic phenomenon on the sliding surface, give insight into fundamental restrictions of design for a sliding surface with low friction force, as well as the design of useful solid lubricant surfaces in vacuum space and memory disks.

4. CONCLUSIONS

The friction force of diamond sliding on a H-terminated Si(111) surface was as low as that of a diamond/Ag/Si system. We have reached tentative conclusions to explain this extraordinarily low friction force. A mechanism of sliding has been presented which includes very low chemical interactions between opponent surface atoms that reduced the friction force, and can be summarized as follows.

1) The diamond carbon atoms or the H atoms on a diamond surface decreased the friction with the H-terminated Si(111) surface due to the very low interaction of C(H) – H(on Si).

2) Diamond(H) – Ag has very low energy bonds, so that the atomic pairs have a very low friction force.

3) The diamond(H) atoms have very strong interactions with Si and O-covered Si showing a very high friction force between them.

5. ACKNOWLEDGEMENT

The authors express their sincere gratitude to Professor Y. Saeki, Nagoya Institute of Technology, for his measurements of the crystal orientation of the diamond using Laue diffraction. Mr. M. Goto, Toyota Technological Institute, presented important suggestions.

REFERENCES

1. M. Goto, F. Honda, and M. Uemura, Proc. of 25th Leeds/Lyon Symposium on Tribology, (1998) 313.
2. M. Goto, F. Honda, and M. Uemura, Wear, 252 (2002) 777.
3. M. Goto, K. Watanabe, and F. Honda, Surf. Sci., (2002) in press.
4. M. Martin, C. Donnet, Th. Le Mogne, and Th. Epicier, Phys Rev. B, 48 (1993) 10583.
5. D. H. Buckley, Japan J. Appl. Phys. Suppl. 2, Pt.1 (1974) 297.
6. H. Yoshizawa an J. Israelachivili, J. Phys. Chem., 97(1993) 11300.
7. M. N. Gardos and K.V. Ravi, Diam. Films Technol. 4(1994) 139.
8. T. Takahagi, I. Nagai, A. Ishitani, H. Kuroda, and Y. Nagasawa, J. Appl. Phys., 64 (1988) 3516.
9. G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Ragharachari, Appl. Phys. Lett., 56 (1990) 656.
10. G. Schulze and M. Henzler, Surf. Sci., 124 (1983) 336.
11. T. Sakurai andH. D. Hagstrum, Phys. Rev. B, 12 (1975)5349.
12. J. J. Boland, Surf. Sci., 244 (1991) 1.
13. Y.Gotoh and S. Ino, Japan J. Appl. Phys., 17 (1978) 2097.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

A novel experimental analysis of the rheology of ZDDP tribofilms

Clotilde Minfray [a], Jean-Michel Martin [a], Ton Lubrecht [b], Michel Belin [a] and Thierry Le Mogne [a]

[a] Ecole Centrale de Lyon, UMR 5513, 69134 Ecully cedex, France
[b] INSA Lyon, UMR 5514, 69621 Villeurbanne cedex, France

In this study, the rheology and durability of tribofilm has been investigated. A ZDDP tribofilm (6x8 mm) has been partially removed by Ar^+ ion etching (5keV), using an aluminum mask with a comb shape. A striped sample was obtained with a lateral alternation of etched zones and ZDDP tribofilm areas, respectively (about 1 mm width). On this striped sample, a friction test was performed using a reciprocating pin-on-flat friction machine. Tests were carried out in different lubricant environments at ambient temperature with the maximum contact pressure of 0.5 GPa. The friction signal and the Electrical Contact Resistance (ECR) were plotted as a function of position and number of cycles. The results can be explained by a mechanical process in the tribofilm zone (its removal) and chemical reaction in the etched zone.

1. INTRODUCTION

For many years, ZDDP (zinc dithiophosphate) additives have been used for their antiwear and anti oxidant properties in lubricating systems. In case of boundary lubrication, ZDDP additives are capable of generating a protective solid film on the friction surfaces. The chemical composition and the processes which generate the film are partially known. The film is constituted of a polyphosphate glass [1], [2] which is formed by the adsorption of the thermo-oxydative decomposition product of the ZDDP [3]. But only few articles treat the mechanical and rheological properties of these films. Furthermore, ZDDP additives cause pollution problems (phosphorus is poisonous for catalysts) [4], and they are commonly used in engine lubrication and Continuously Variable Transmition (CVT) gear boxes. For these reasons it is important to improve our understanding of the rheology of ZDDP nanometric films.

Rheology studies the flow and deformation of materials under a strain. This work investigates the way a ZDDP tribofilm flows under strain (in a tribological context) until its complete removal.

2. BIBLIOGRAPHY

Using Atomic Force Microscopy (AFM) techniques, Graham et al. [5] showed that the surface of the ZDDP alkyl tribofilm is morphologically heterogeneous. The film was composed of large and small pads and their nanomechanical properties were determined by Interfacial Force Microscopy (IFM) measurements. The centre of the large pads exhibited a high indentation modulus (209±38 GPa) and a reduced plastic deformation. Pad edges were found to be softer (87±23 GPa) as were small pads (74±20 GPa). Pad edges and small pads were also

more susceptible to plastic flow than the centre of large pads. Canning et al. [6] completed this study with X-Ray Absorption Near Edge Structure (XANES) analyses which showed that the length of the polyphosphate chains in the centre of the pad was larger than at the edge.

Aktary et al. [7] studied the evolution of the morphology and nanomechanical properties of this kind of ZDDP antiwear film as a function of the tribological contact time. A three stage growth mechanism was highlighted: "nucleation" of islands, "coalescence" of the islands on a large part of the surface, and large islands divided into smaller ones. But this morphological heterogeneity is not mechanically relevant: the rubbing time (5 min to 120 min) has only a minor influence on the nanomechanical properties of the films. For example, a 10 min tribofilm has an indentation modulus of 92.8±18.6 GPa and a hardness of 5.0±0.9 GPa. These values are also characteristic of soft polymeric materials with modest indentation moduli. Contrary to hard coatings, which can provide wear protection by highly elastic deformation and minimise plastic flow, ZDDP tribofilms are continuously formed and consumed.

On the other hand, Bell et al. [8] showed the presence of a soft overlayer covering the solid tribofilm. Bec et al. [9], [10], [11] comfirmed the presence of the overlayer of alkyl-phosphate precipitates on the polyphosphate film removed by solvent washing. It is a soft and mobile layer. They also measured the hardness profiles: when the penetration depth is increased, the hardness increases. So, the ZDDP tribofilm accommodates the indentation pressure by increasing its hardness (like a smart material).

Some nanomechanical data is available on ZDDP tribofilms but very little is known on the correlation between rheological properties and analytical data. Due to the heterogeneity of such films, a macroscopic investigation is performed to relate friction coefficient and rheological behavior. This is the aim of the test procedure described in this paper: the visualization of the film rheology and durability.

3. SAMPLE PREPARATION

First of all, a tribofilm was generated using a Cameron-Plint TE77 tribometer in cylinder-on-flat configuration (Figure 1.a), reciprocating mode, under boundary lubrication conditions. The samples (flat and cylinder: AISI 52100 steel) were polished with a 1 μm diamond solution. The lubricant used is a mixture of a base oil (PAO6) and a ZDDP antiwear additive (1.0 w%) supplied by Asahi Denka Kogyo K.K. The experiments were run at 80°C, during 60 min with a stroke length of 8 mm, at 6 Hz. The normal load was 350 N resulting in a maximum pressure of 600 MPa. The tribofilm is generated on a rectangular area of 6x8 mm.

The second stage of the sample preparation is the elaboration of the striped sample. A mask is prepared from an aluminum sheet, with stripe shaped windows of 1 mm width, separated by a 1 mm spacing. This mask is placed over the sample, using double face self-adhesive tape on the edge (Figure 1.b). Special care is taken to flatten the mask on the sample.

Thirdly, the sample is etched with an Ar^+ ion beam (5keV) in order to remove the tribofilm where it is not covered by the mask (Figure 1.c). AES (Auger Electron Spectroscopy) depth profiles are performed during this operation, in order to monitor the etching process and record analytical data. Etching is stopped when the tribofilm is completely removed (Figure 2).

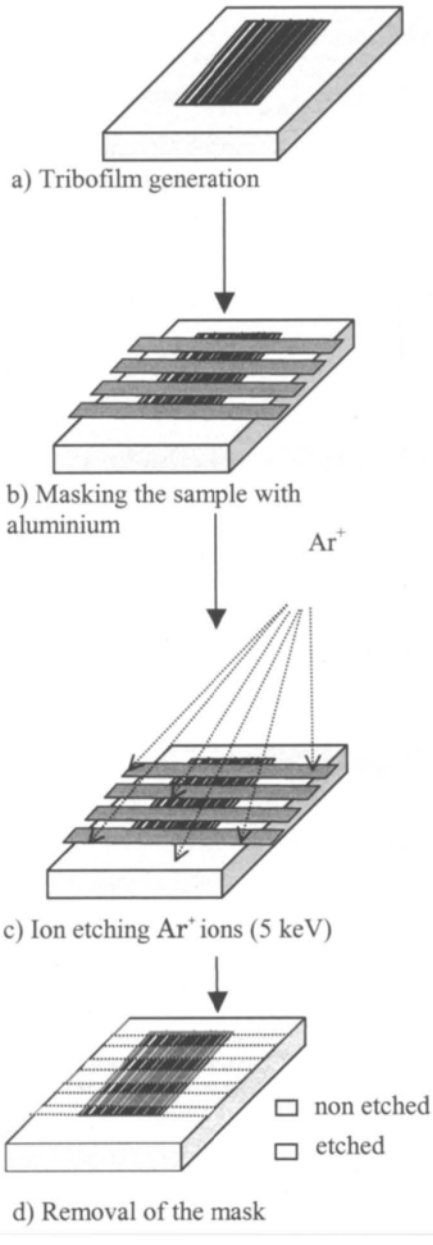


Figure 1. Schematic representation of the different steps in the sample preparation: a) generation of the tribofilm (Cameron-Plint tester, friction zone dimensions 6x8 mm), b) positioning the aluminum mask (stripes of 1mm width), c) Ar^+ ion etching, d) removal of the mask.

Finally, the mask is removed and a striped sample is obtained, with a lateral alternation of the original ZDDP tribofilm and etched zones, as shown in Figure 1.d.

From the AES depth profiles of Figure 2, it is possible to estimate the tribofilm thickness. Assuming that the etching speed in the iron oxide is the same as in the phosphate, and given that the oxide thickness is estimated at 4 nm by X-ray Photoelectron Spectroscopy XPS (ratio of iron oxide over iron metal peaks), the tribofilm thickness is estimated at about 14 nm. It is noted that AES depth profiling does not give any quantitative information.

4. THE DUAL ZONE FRICTION TEST (DZFT)

This paper studies the local evolution of the surface layer. The technique used is based on spatially resolved measurements in a reciprocating contact. This approach called *triboscopy* was described elsewhere [12]. It has been previously used to study phenomena occurring in a sliding interface, especially the build-up of a transfer layer and local wear of a solid lubricant in thin films [13].

In order to study the material deformation and displacement due to the sliding contact, this method has been improved here by using samples with two different surface treatments, the testing amplitude crossing these zones of different nature (Figure 3). It is called in this paper *Dual Zone Friction Test* (*DZFT*). Therefore, *DZFT* allows one to study the evolution of each zone during the test. Special attention is given to the evolution of the "border" between the two different zones.

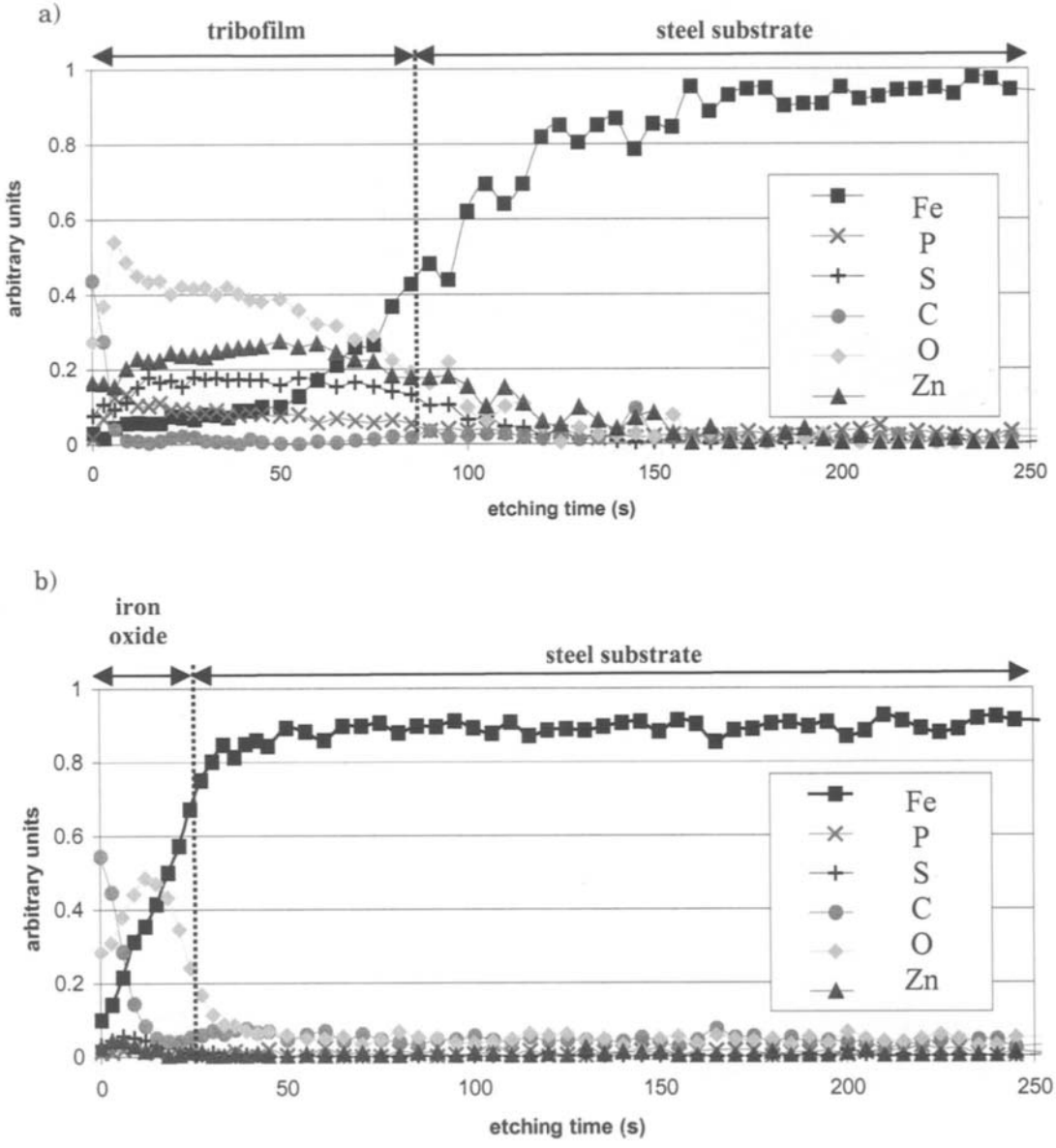


Figure 2. AES depth profiles: a) inside wear scar (mark a in Figure 3), b) outside wear scar (mark b in Figure 3).

For this purpose, a ball-on-flat friction machine (different device as the Cameron Plint one) was used. The ball is made from AISI 52100 steel which has a diameter of 6 mm. The friction coefficient and the electrical contact resistance (ECR) were recorded using dedicated electronics (measurement range : 1 to 10^7 ohms). Data is plotted in triboscopic diagrams, as a function of lateral position and number of cycles. The contact conditions were as follows: maximum pressure 0.5 GPa, stroke length 2.0 or 2.4 mm, frequency 0.5Hz, and a duration of 100 or 500 cycles. Under these conditions, the lateral resolution of the diagrams presented is 9.0 μm (distance corresponding to two successive pixels in the diagrams). Tests were carried out under dry conditions (open air, room temperature) and

in lubricated conditions, using different lubricants (dodecane and PAO).

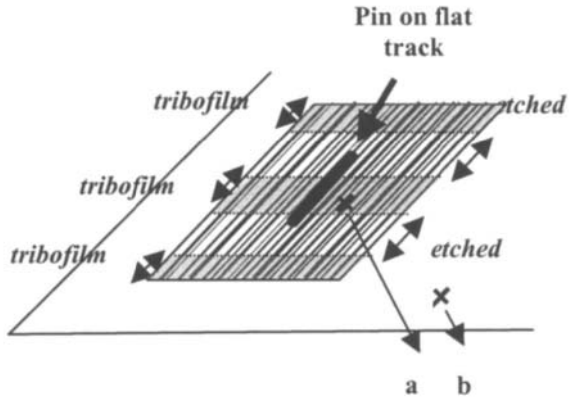


Figure 3. Schematic view of the track (2.0 or 2.4 mm lenght) in a *DZFT* experiment. The two marks a and b indicate where the AES depth profiles of Figure 2 were performed (mark a: inside wear scar, mark b: outside wear scar).

5. RESULTS AND DISCUSSION

5.1. Unlubricated DZFT

First, the response of the tribofilm as tested with unlubricated sliding contact was analyzed. A *DZFT* is run on the sample under the following conditions: stroke length 2.4 mm, frequency 0.5Hz, maximum Hertzian pressure 0.5 GPa, 100 cycles in ambient air. The triboscopic images are shown in Figure 4. The tribofilm is located near the centre of the image.

The local friction image (Figure 4.a), shows the initial situation turns to high friction with some stick-slip phenomena, inducing contact vibration, due to the low longitudinal stiffness of the test machine. The "low friction effect of the tribofilm" is found to last about 25 cycles.

On the Electrical Contact Resistance image (Figure 4.b), one clearly sees the presence of the highly insulating tribofilm (10^7 ohms). This feature completely disappears within 40 cycles, the ECR value decreases by several orders of magnitude. From such a diagram, one can estimate that the "tribofilm insulating effect" starts to disappear at about 25 cycles. One can deduce that the tribofilm has been partially degradated by the multipass friction test.

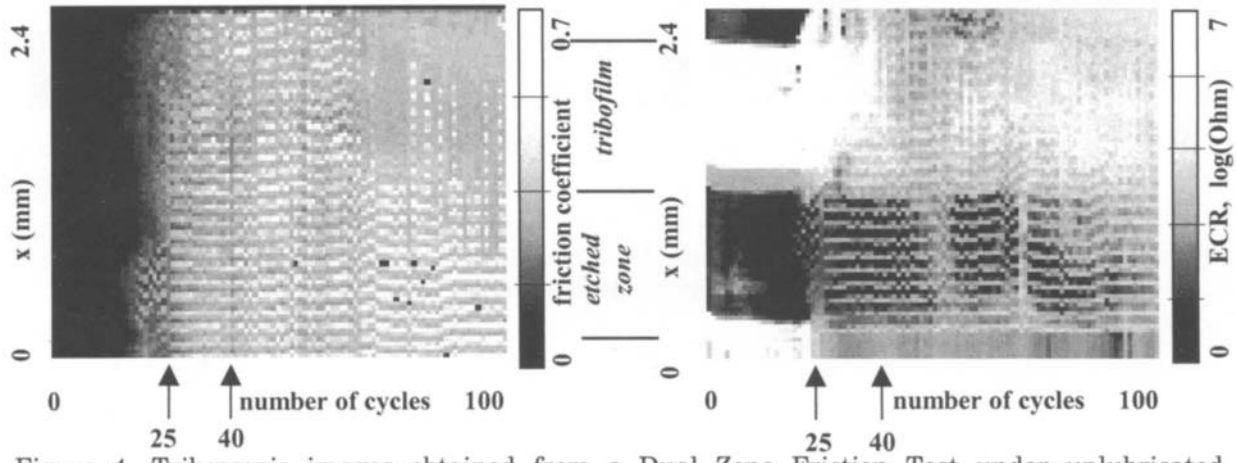


Figure 4. Triboscopic images obtained from a Dual Zone Friction Test under unlubricated conditions : a) local friction coefficient b) local ECR (test conditions: amplitude 2.4 mm, 100 cycles, max. contact pressure 0.5 GPa, ambient air).

One can also notice that the lateral edges of the high ECR zone are stable at the start, but both edges move laterally (after 25 cycles) before complete degradation occurs.

For the last 25 cycles, a high friction coefficient and low ECR occurred in both zones (compared to initial values in the tribofilm) showing that the tribofilm has been degradated.

From these first results, one can conclude that the method is suitable for the observation of the local removal of the initial tribofilm. The contact conditions used here are quite severe. In order to lower severity, a lubricated contact is studied with different fluids, maintaning identical contact conditions.

5.2. DZFT in dodecane

In this section, the response of the tribofilm is studied when lubricated with a low viscosity model lubricant: pure dodecane. The *DZFT* is run using the following conditions : stroke length 2.4 mm, maximum Hertzian pressure 0.5 GPa, 500 cycles. The triboscopic images are shown in Figure 5. The tribofilm zone is located in the centre of the figure.

On the friction image (Figure 5.a), one can notice a low friction zone, in the centre of the track, corresponding to the presence of the tribofilm. The "low friction effect of the tribofilm" lasts for 220 cycles. After that stick-slip occurs.

On the ECR image (Figure 5.b), one clearly sees the presence of the highly insulating tribofilm located in the centre part of the test stroke. This feature corresponds to the low friction zone described above. This feature is found to vanish completely within 400 cycles, the ECR value decreases sharply. The high ECR zone is gradually reduced in width, the two lateral edges moving smoothly towards the centre.

Hence, the severity of the unlubricated contact has been reduced by using dodecane.

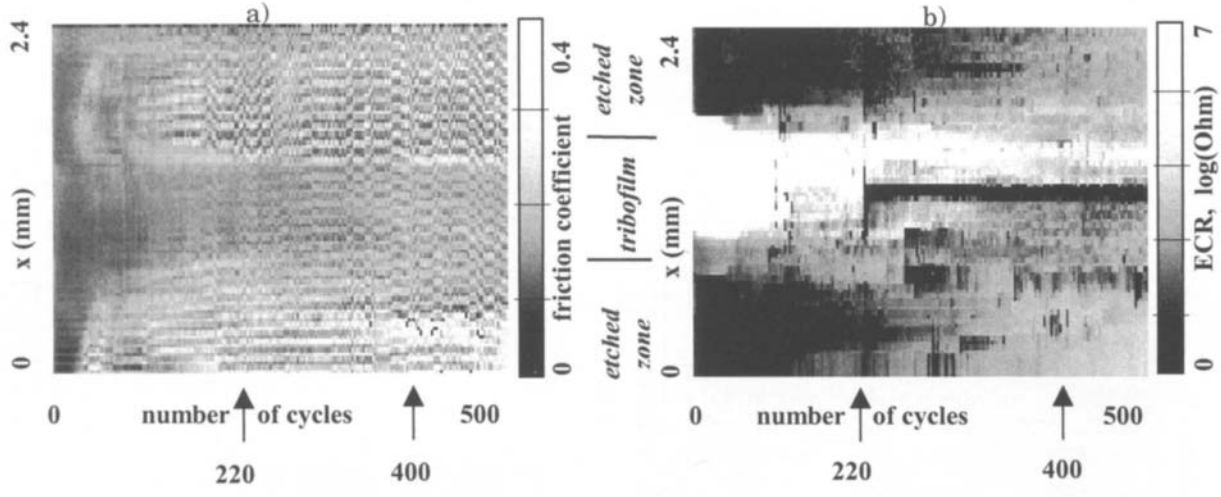


Figure 5. Triboscopic images obtained from a Dual Zone Friction Test, lubricated with dodecane: a) local friction coefficient b) local ECR (test conditions: amplitude 2.4 mm, 500 cycles, max. contact pressure 0.5 GPa).

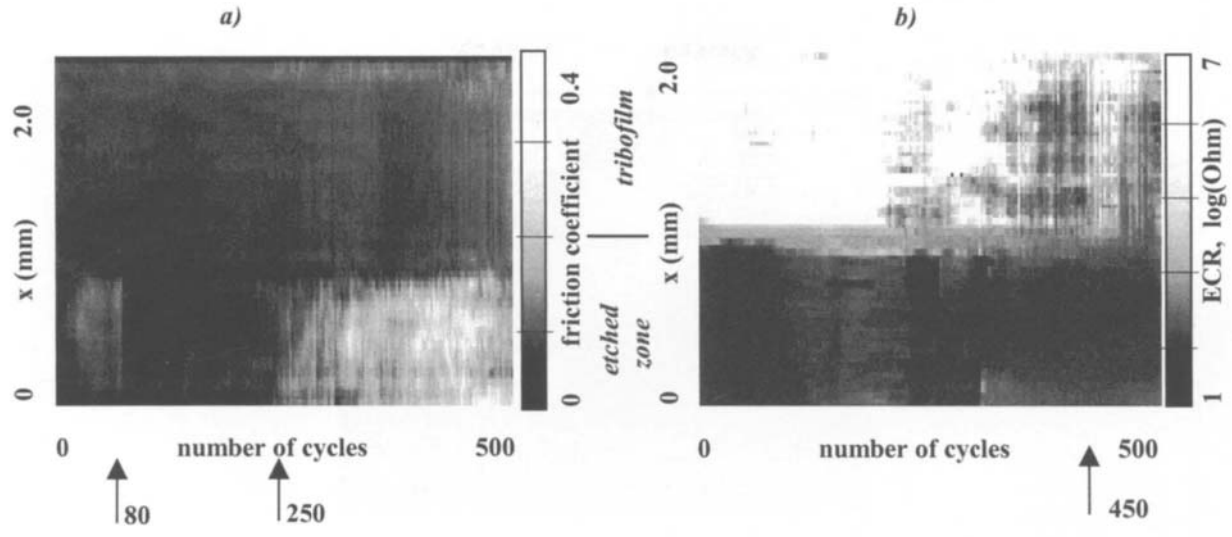


Figure 6. Triboscopic images obtained from a Dual Zone Friction Test in PAO lubricant, a) local friction coefficient b) local ECR (test conditions: stroke amplitude 2.0 mm, maximum pressure 0.5 GPa, 500 cycles).

5.3. DZFT in PAO base oil

5.3.1. DZFT results

In this section, the *DZFT* test is performed in the base oil (PAO6), the same one as used for the tribofilm generation in the Cameron-Plint machine. This situation corresponds to a particular case, when all active additives have been depleted. The *DZFT* is run on the sample under the following conditions: stroke length 2.0 mm (2.4 mm in §5.1 and §5.2), maximum Hertzian pressure 0.5 GPa, 500 cycles in ambient air. Figure 6. shows friction (Figure 6.a) and ECR (Figure 6.b) results. The tribofilm is located at the top of the pictures and the etched zone at the bottom.

Using the local friction information (Figure 6.a), the two zones can be easily distinguished: the tribofilm zone has a low friction coefficient compared to the etched part up to 80 cycles. From that point, the friction coefficient decreases significantly in the etched part. From 250 cycles onwards, the friction coefficient increases gradually to the end. It is interesting to note that the measured friction coefficient on the tribofilm at the beginning of the experiment is about 0.13 which is similar to the Cameron-Plint experiment of film formation (0.13). The reduced friction effect in the etched zone is about last for 250 cycles.

From the figure of the Electrical Contact Resistance (Figure 6.b), the difference between the two zones is also clear. Up to 200 cycles, the tribofilm zone has a high ECR value contrary to the etched one which is conductive. After that point, some conductive zones appear in the tribofilm. One edge moves laterally more than the other one (contrary to the unlubricated case where both edges move). The location was different from the unlubricated and dodecane lubricated cases. This will influence the removal process of the tribofilm. One can estimate the "tribofilm insulating effect" at about 200 cycles, much more than under unlubricated conditions (§5.1). The insulating feature of tribofilm in terms of ECR vanishes after around 450 cycles.

The friction coefficient reduces in the etched zone after 80 cycles. To understand this phenomenon, on additional analysis on the sample is performed.

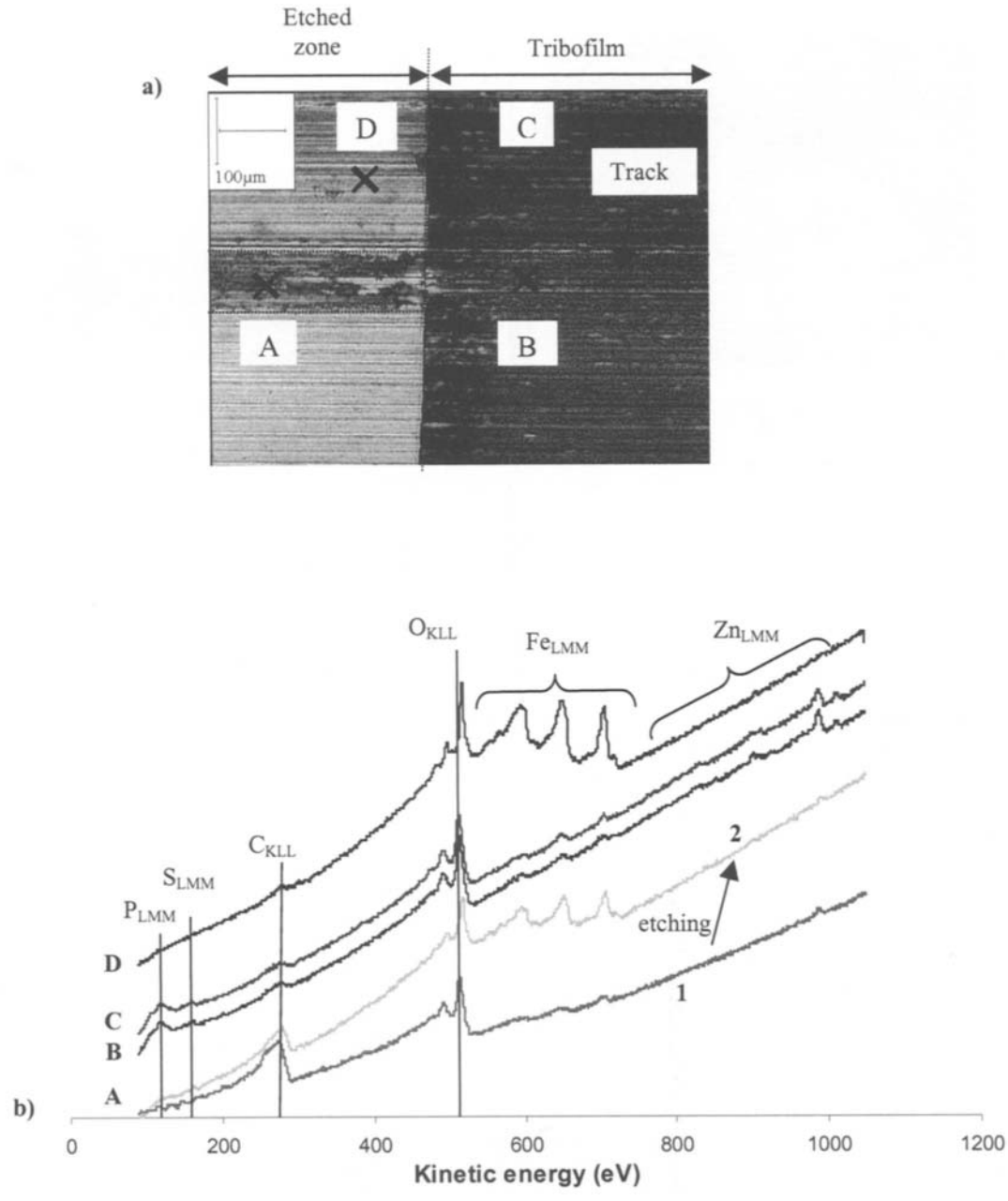


Figure 7. a) Optical observation of the track in the DZFT experiment performed in PAO
- mark A: track in the etched zone
- mark B: tribofilm in the DZFT track
- mark C: tribofilm
- mark D: substrate

b) AES spectra (lateral resolution less than 1 µm) were performed on the different marks of Figure 7.a.- For mark A, track in the etched zone, two spectra were performed:

1) after experiment
2) after few seconds of etching to remove carbon contamination

C_{KLL} is the notation of a Auger peak. KLL and LMM indicate the different electronic shells that are involved in the Auger transition.

5.3.2. AES spectra

AES spectra have been obtained in order to determine if the tribofilm is transferred to the etched zone causing a friction reduction. These analyses were performed on different points of the sample which are indicated in the Figure 7, after cleaning the sample with heptane.

- Spectrum A.1, (wear track in etched zone) shows a large carbon contamination, the presence of oxygen and zinc.
- After etching during a few seconds on the track in the etching zone (zone A), the spectrum A.2 was obtained. The carbon is still largely present but a few traces of phosphorus, sulfur and zinc were also detected.
- Spectrum B (in the track in the tribofilm zone), all peaks and especially the carbon one are equal to the one in the tribofilm zone (spectrum C).
- Spectrum C is characteristic of the tribofilm : Phosphorus, Sulfur and Zinc are detected together with oxygen and a little iron.
- Spectrum D shows that there is no carbon contamination on the substrate. Only oxygen and iron are detected. The oxygen peak position is different than in spectrum C (tribofilm). The oxygen could be bound to iron (like in spectrum A.2).

So the PAO seems to generate a carbon-based film (with oxygen and iron) in the etched zone and not in the tribofilm. That film could be partially responsible for the improvement in friction behavior.

Traces of characteristic tribofilm elements (P, S and Zn) were detected in zone D. These analyses were done after the entire experiment was finished. It would be interesting to perform the same analyses earlier, between 50 and 250 cycles, in order to study th evolution of the tribofilm with time.

Representative auger depth profile analyses (Figure 8) were also carried out in the tribofilm outside (Figure 8.a) and inside (Figure 8.b) the track. The level of zinc starts to decrease after 90 seconds of etching outside of the track, and after 70 seconds of etching, in the track. A thickness which correponds to 20 seconds of etching (under these etching conditions), was removed during the experiment. So, there remains a substantial amount of tribofilm material after the DZFT experiment.

At the end of the experiment, the ECR measurement can no longer detect the film. Maybe, this difference is due to the high roughness of the sample (some scratches of the Cameron-Plint track). The tribofilm could be present in the valleys but not on the tops. That could be the reason why, this film is detected in Auger but not in ECR measurements.

According to all this data, the removal process of the tribofilm could be explained by a combination of mechanical and chemical processes.

The tribofilm removal itself can be seen as a purely mechanical process. This sacrificial film fragments from the edges to the centre of the strip. But the precise process is not yet known and questions remain on the precise film transfer mechanisms, on the role of the counterface, on film deformation etc.

In the etched zone, for the *DZFT* under PAO lubricating conditions, it seems that a chemical reaction occurs (formation a carbon-rich film) with presence of tribofilm elements.

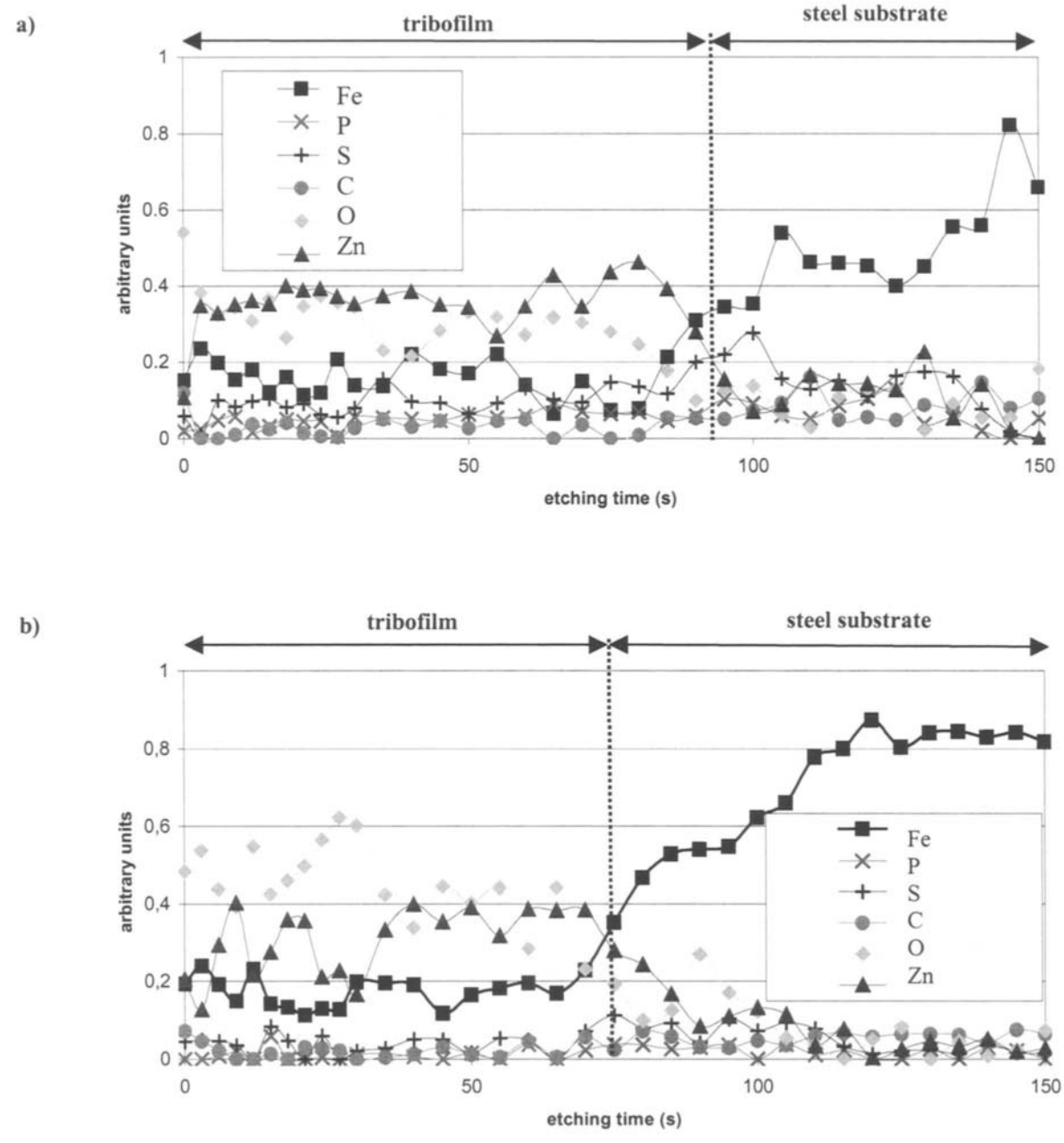


Figure 8. AES depth profiles carried out: a) outside the *DZFT* track in the tribofilm (mark C in Figure 7), b) inside the track in the tribofilm (mark B in Figure 7).

6. CONCLUSION

This paper shows the feasibility of a novel experimental analysis of the rheology of a ZDDP tribofilm. The tribofilm thickness is evaluated at about 14 nm. The test consists in partially removing the tribofilm on a steel surface and to run a friction test on the modified surface. Using triboscopic analysis of both friction and ECR measurements, information on the tribofilm durability are visualized under different lubrication conditions. Unlubricated contact conditions are quite severe, and the use of dodecane lowers this severity. The use of PAO resulted in a particular case where all additives of the lubricant have been depleted. Both mechanical and chemical removal mechanisms seem to be involved.

In the future, this kind of experiment should be improved and completed. For example, AES spectra can be performed after different numbers of cycles to study the evolution of the wear scar composition and the kinetics under different environments. The influence of the temperature should also be taken into account.

ACKNOWLEDGEMENTS

The authors would like to thank Asahi Denka Kogyo K.K. and French Ministery « Ministère de l'éducation nationale, de la recherche et de la technologie » for their support and Professor N. Omahe for his help during the preparation of this paper.

REFERENCES

1. J.M. MARTIN, Tribology letters, 6 (1999) 1-8.
2. J.M. MARTIN, C. GROSSIORD, T. Le MOGNE, S. BEC and A. TONCK, Tribology international, 34 (2001) 523-530.
3. M. FULLER, M. KASRAI, G. BANCROFT, K. FYFE and K. TAN, Tribology international, 31 (1998) 627-644.
4. K. INOUE, Japanese journal of tribology, 43 (1998) 1101-1111.
5. J.F. GRAHAM, C. McCAGUE, P.R. NORTON, Tribology letters, 6 (1999) 149-157.
6. G.W. CANNING, M. FULLER, G.M. BANCROFT, M. KASRAI, J.N. CUTLER, G. DE STASIO and B. GILBERT, Tribology letters, 6 (1999) 159-169.
7. M. AKTARY, M. McDERMOTT and G. McALPINE, Tribology letters, 12 (2002) 155-162.
8. J. BELL, R. COY and H. SPIKES, Proceedings of the Japan International Tribology Conference, Nagoya, (1990) 505-510.
9. S. BEC, A. TONCK, J.M. GEORGES, R.C. COY, J.C. BELL and G.W. ROPER, R. Soc. Lond. A, 455 (1999) 4181-4203.
10. A. TONCK, S. BEC, J.M. GEORGES, J.C. BELL, and G.W. ROPER, Proceedings of the 25th Leeds-Lyon Symposium (1999) 39-47.
11. D. MAZUYER, A. TONCK, S. BEC, J.-L. LOUBET, J.-M. GEORGES, Proceedings of the 27th Leeds-Lyon Symposium, (2001) 273-282.
12. M. BELIN, Wear, 168 (1993) 7-12.
13. K.J. WAHL, M. BELIN and I. SINGER, Wear, 214 (1998) 212-220.

This Page Intentionally Left Blank

SESSION XXII
TRIBOLOGICAL STUDIES OF DRY AND LUBRICATED SURFACES (4)

Chairman : Dr M Priest

Paper XXII (i)	**Friction from Fluid-Film to Boundary Lubricated Conditions**
Paper XXII (ii)	**A Macro-Micro- Fatigue Approach for Predicting Fatigue Life in Mixed - EHL Contact**
Paper XXII (iii)	**Progression of the Stick/Slip Zones in a Dry Wheel-Rail Contact : Updating Theories on the Basis of Tribological Reality**
Paper XXII (iv)	**Tribology of Electron Beam Excited Plasma Nitrided Tool Steels with Narrow Gap Features**

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Friction from fluid-film to boundary lubricated conditions

R. Bassani and E. Ciulli

Dipartimento di Ingegneria Meccanica, Nucleare e della Produzione, Università di Pisa
Via Diotisalvi, 2 - 56126 Pisa - Italy

Friction is an important feature for engineering systems. The friction coefficient, a fundamental parameter for machine design, is unfortunately not easy to determine because of the large number of variables influencing its value. Different lubrication regimes, ranging from fluid-film to boundary and corresponding to different ways of evolution of the friction coefficient, are usually considered. This work reviews studies on friction, mainly experimental, carried out with several kinds of lubricated contacts working under various conditions. The influence of several variables on the evolution of the friction coefficient is shown by using typical Stribeck-like and lambda curves. Some comparisons between experimental and theoretical results are also reported. Relatively simple formulas for the evaluation of the friction coefficient, based on appropriate models of the lubricant behaviour, have been developed that provide results in quite good agreement with the experimental ones under fluid-film conditions when thermal effects can be neglected.

NOMENCLATURE

f	friction coefficient
F	load
h	film thickness
h_c	central film thickness
p_H	Hertzian pressure
p_m	mean contact pressure
R_q	root mean square roughness
S	slide/roll ratio, $\Delta u/u$
SP	service parameter, function of η_0, u, F
u	rolling speed, $(u_1+u_2)/2$
α	lubricant pressure-viscosity coefficient
Δu	sliding speed, u_2-u_1
Λ	lambda ratio, h/σ
η	lubricant viscosity
η_0	lubricant viscosity at ambient pressure
σ	equivalent roughness, $(R_{q1}^2+ R_{q2}^2)^{0.5}$
τ_0	Eyring stress

Subscripts

1	referred to the body 1
2	referred to the body 2

1. INTRODUCTION

Friction related problems are very important for engineering systems, in particular for a good design with regard to components life and energy savings. Friction losses can be measured directly on real machines, but a preliminary tribological research, experimental and/or theoretical, can be very useful for time and costs reduction. One of the most required data for design is the friction coefficient, also employed in simulation programmes useful to reduce the number of experimental tests [1]. Unfortunately it is not always easy to find a realistic friction coefficient because there is a large number of variables (such as lubricant, velocity, load, geometry, roughness and materials) influencing its value.

The evolution of the friction coefficient is especially influenced by the parts of load supported by the lubricant and by the surface asperities of the solids, essentially depending on load, speed and lubricant viscosity values.

Three different lubrication regimes, ranging from fluid-film to boundary, are usually considered. Under the "fluid-film (or full fluid) lubrication" conditions the solids are completely separated by the lubricant film. The lubrication is called "hydrodynamic" or "elastohydrodynamic" depending primarily on the elastic properties and the degree of conformity of the solids. The term "boundary lubrication" is then used to describe the regime in which the film is reduced to absorbed molecular layers of lubricant on the surfaces of the solids. The contact area is comparable to the one that develops in dry contacts and the laws of dry friction apply with an actual value of the friction coefficient that depends upon the nature of the boundary lubricant; the load is supported by the surface asperities through the molecular layer of lubricant. The transition from fluid film to boundary lubrication does not take place instantaneously: the proportion of load carried by the surface asperities increases and the one carried by the lubricant film between the bodies decreases by increasing load or by decreasing speed or viscosity. This intermediate condition is called "mixed lubrication".

Useful ways to represent the evolutions of the friction coefficient f, evidencing the transition between the different lubrication regimes, are the so-called "Stribeck curves". In these diagrams f is usually plotted against a dimensionless number that is dependent on velocity, load and lubricant viscosity. This dimensionless "service parameter", SP, can be found in different forms and with different names (for instance Sommerfeld or Stribeck number) in the literature, as reported for instance in [2-3]. The Stribeck curves are usually obtained by varying only one parameter (often the speed) while the other quantities are maintained constant. Starting with hydrodynamic conditions, the value of the friction coefficient decreases by reducing velocity or viscosity or by increasing load till reaching a limiting value. An opposite effect occurs when the parameters are varied in the mixed lubrication regime, where the load is shared between the contacting asperities and the lubricant film. A single curve is not sufficient to describe the variation of the friction coefficient under these conditions. Different evolutions can be obtained for different load, speed and viscosity, as well as by varying other parameters such as the surface roughness and the hardness of the lubricated bodies, Fig.1.

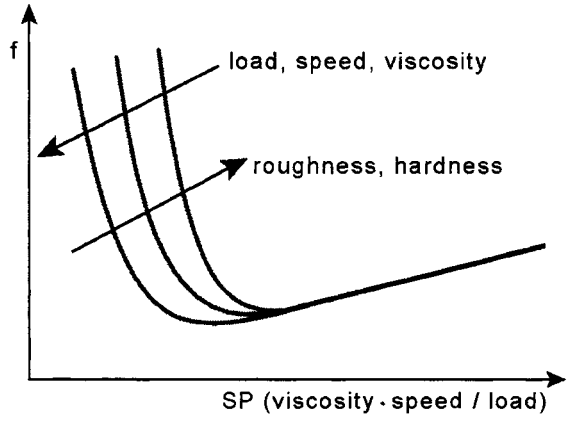


Figure 1. Typical Stribeck curves; influence of some parameters under mixed lubrication.

The evolutions qualitatively shown in Fig.1 are based on results obtained for journal bearings, as reported for instance in [4]. Under mixed lubrication regime, for a given value of SP, an increase in load or viscosity as well as a decrease in roughness or hardness tends to decrease f (constant-speed curves); an increase in speed tends to decrease f (constant-load curves). The differences tend to vanish at high values of SP (full fluid conditions), at least for journal bearings. Under these conditions, f varies almost linearly if an opportune SP function is chosen.

For inclusion of the effects of different surface roughness in only one curve, diagrams can also be used where f is plotted as a function of the so-called "lambda (Λ) ratio", a ratio of a representative value of the lubricant film thickness h and the equivalent

root mean square roughness of the two bodies in contact σ:

$$\Lambda = \frac{h}{\sigma} \tag{1}$$

where $\sigma=(R_{q1}^2+ R_{q2}^2)^{0.5}$, with R_{q1} and R_{q2} root mean square values of the roughness of the two bodies. It is easy to verify that the dimensionless film thickness Λ depends on the viscosity η_0, the speed u and the load F as the service parameter SP [3]. When putting typical expressions for calculating h under hydrodynamic or elastohydrodynamic lubrication conditions in the eq. (1), we obtain:

$$\Lambda = K\frac{\eta_0^l u^m}{F^n}\frac{1}{\sigma} \tag{2}$$

where K can depend on factors such as geometry, elastic properties of the bodies, piezoviscosity coefficient of the lubricant; l, m and n are numbers, which values depend on the type of lubrication (hydrodynamic, elastohydrodynamic, etc.) and of contact (conformal or nonconformal, line or point). Usually l=m, with values ranging from 0.5 to 2, and $0<n<2$.

Since the main difference between Λ and SP is the inclusion of the roughness in the first one, similar curves to those of the Stribeck diagrams can be obtained with lambda on the horizontal axis, Fig.2. The values of Λ shown in this figure are only suggested values for the change of lubrication regimes, but they can be rather different, as discussed later in the paper.

This work reviews experimental studies on friction carried out with several kinds of lubricated contacts working under various conditions. Typical nonconformal contacts were in particular investigated in moderately loaded conditions [5-7]. Some significant results are reported to show the influence of several parameters on friction. A comparison between experimental and theoretical results is also performed on the basis of expressions developed for the evaluation of f.

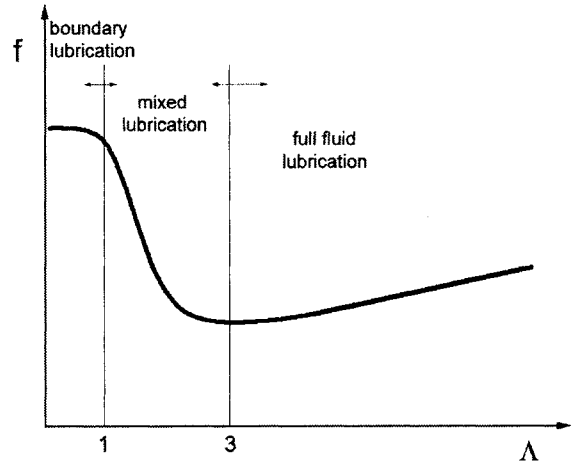


Figure 2. Typical lambda (Λ) curve with lubrication regimes.

2. EXPERIMENTAL WORKS

All experimental results reported in this work were obtained with the experimental rig for simultaneous measurements of the friction force and of the film thickness and shape shown Fig.3.

The lubricated contact is formed between the plane surface of the disc and the specimen. The load is applied with weights via a lever mechanism having a radial gas bearing as fulcrum. The longitudinal motion of the support is constrained only by a load cell that measures the friction force in the contact. The friction coefficient is calculated as the ratio of the measured friction force and the applied load. The contact between the disc and the specimen can be regulated by a vertical positioning system. The disc and the specimen can be driven by separate electric motors connected to the output of a digital acquisition board. Friction values are recorded by the data acquisition system every second for a total interval of one minute. The averaged values are normally used in the analysis.

An optical interferometry system makes it possible to investigate the film thickness and shape when glass discs are employed. Additional information about the apparatus is reported in [6].

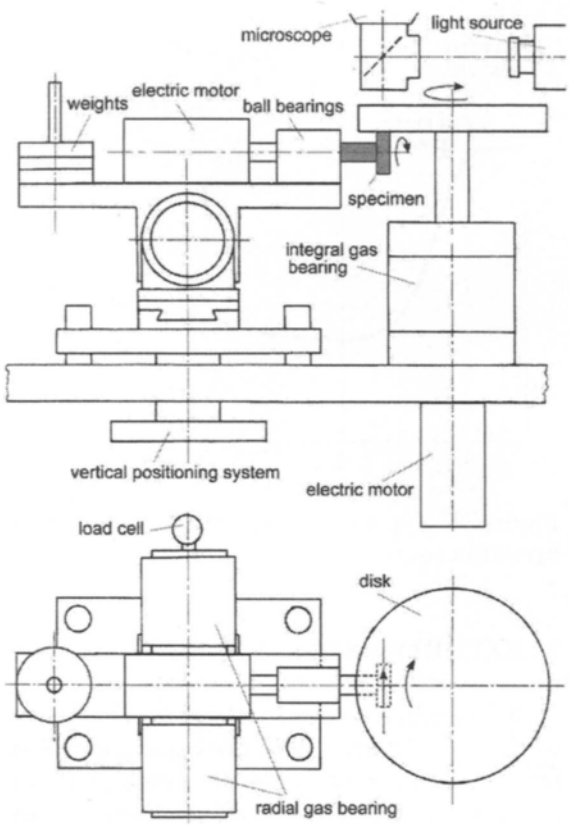


Figure 3. Schematic drawing of the experimental rig.

The experimental rig is extremely versatile: both conformal and nonconformal contacts can be tested with this machine. Lubricated contacts between discs and cylindrical and spherical specimens [5-7] and thrust bearing tilting pads [8-9] were investigated.

The influence of roughness and materials on the friction trends was firstly investigated, particularly under mixed and boundary lubrication conditions [5]. Rectangular contacts between the flat surfaces of discs and cylindrical specimens, all with the same shape, were tested. Some wear effects were found after tests under mixed lubrication.

Steel specimens with similar roughness were then used in [6] against a glass disc. The influence of the shape was particularly investigated by using both spherical and cylindrical specimens. To limit the influence of the different parameters, the tests were carried out at constant load and lubricant temperature with specimens made of the same steel. These specimens were also tested against some other steel discs with different surface roughness and hardness in [7]. The dimensions of the specimen were chosen to have similar Hertzian contact characteristics (maximum or mean pressure and contact area) or lubricant film thickness (minimum or central) in order to find the most important variables with regard to friction. All the cylindrical specimens had the same axial width of 3mm, selected to avoid significant variations of speed along the radial direction of the disc and to limit the border effects.

3. FLUID FILM RESULTS

3.1. Nonconformal contacts

Some of the cylindrical and spherical specimens used in [6-7] are shown in Fig.4.



Figure 4. Steel specimens.

All specimens were tested at the same values of the rolling speed $u=(u_1+u_2)/2$, with u_1 and u_2 respectively surface speed of the disc and of the specimen (u=0.0125, 0.025, 0.05, 0.075, 0.1, 0.2, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1m/s). For each speed four values of the slide/roll ratio, $S=(u_2-u_1)/u$, were investigated (S=0, 0.25, 0.5, 1) with the measured traction forces in tests with S=0 (pure rolling conditions) used as a reference value. The load and the lubricant temperature were respectively 20N and 30°C. Viscosity and pressure-

viscosity coefficient of the lubricant, bis(2-ethylhexyl)phthalate, were η_0=0.042Pa s and α=1.8 10^{-8} Pa^{-1}.

As expected, the highest film thicknesses were associated to the biggest specimens, a spherical one with a diameter Φ=41.275mm and a roughness R_q=0.03μm (identified by the label S4) and a cylindrical one (C4, with Φ=42mm and R_q=0.14μm). C4 and S4 worked under full fluid conditions starting from lower speeds compared to other smaller specimens. For a better investigation under fluid film conditions C4 and S4 have been recently tested with the same lubricant at a lower temperature (T=10°C). Typical results obtained at the two temperatures with the spherical specimen S4 against a glass disc (D7) with a SiO_2 coating (R_q=0.03μm) and two similar fine ground steel discs (A9 and A10, respectively with R_q=0.07μm and R_q=0.05μm) are shown in Fig.5 and in Fig.6. Since the lubricant viscosity and the load are kept constant, the trends of f, plotted as a function of u, are similar to the ones of the Stribeck curves (for this reason these diagrams are also called "Stribeck-like" diagrams). The two vertical lines reported on the diagrams correspond to the significant values of Λ=1 (dotted line) and Λ=3 (solid line). The values of Λ have been calculated using the minimum film thickness formulas reported in [10] and the out-of-contact equivalent roughness. Note that the diagram shown in Fig.6b is reported in this section only for comparison, since, due to the high values of the equivalent roughness, full fluid conditions are not established (Λ=3 is reached in this case at u≈1.2m/s).

It is worth noting that thermal effects were more important for the low than for the high temperature tests: at the highest speed an increase of 4°C was measured between the outlet and the inlet oil temperature of 10°C, while no significant differences were measured during tests at 30°C (two thermocouples were used for these measurements).

When comparing the diagrams of Fig.5 with those of Fig.6, the disc material seems to play a relevant role. The roughness differences of the discs cannot be sufficient to justify the differences, and probably also other variables, such as hardness and thermal parameters (e.g. material conductivity) have a certain influence on the lubricated contacts.

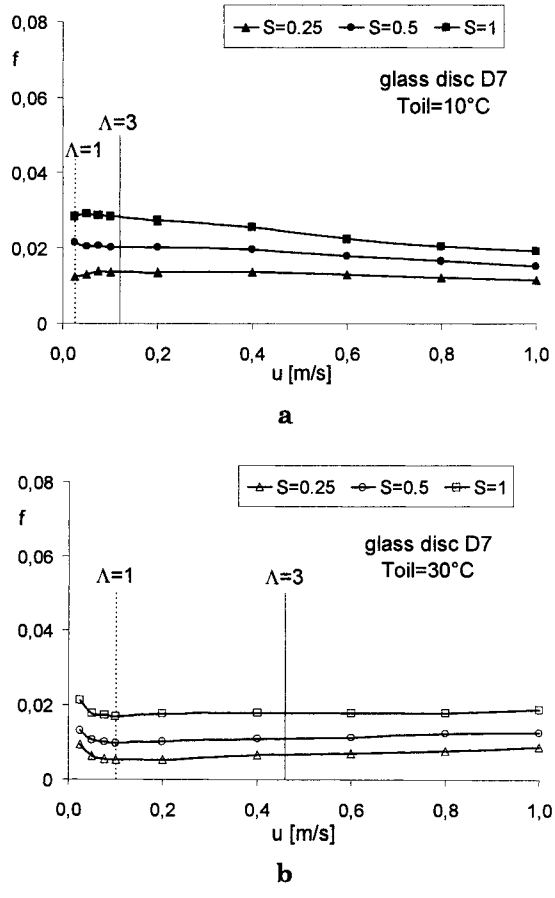
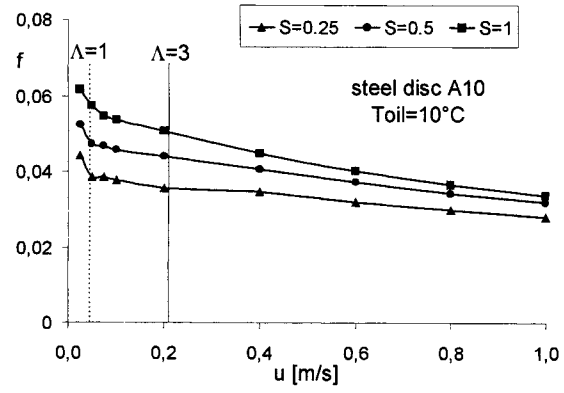


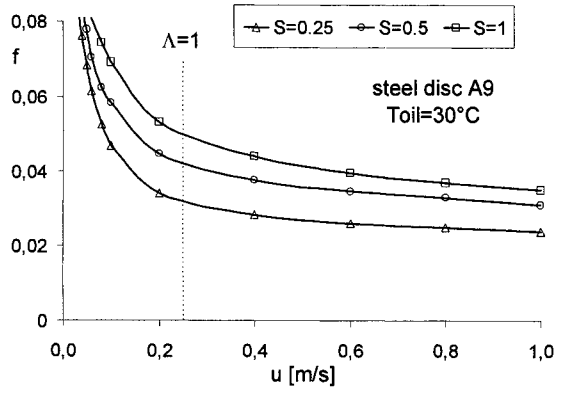
Figure 5. Friction coefficient for the spherical specimen S4 against a glass disc at two lubricant temperatures: 10°C (a) and 30°C (b).

In Fig.7 a similar diagram for the cylindrical specimen C4 is shown.

The typical inversion related to the different viscosity, reported for instance in [11], appears from the comparison of the tests made at the two temperatures. The lubricant with the lowest viscosity (the oil at 30°C) produces lower friction for the same speed under full fluid conditions, while an opposite effect arises under mixed lubrication, Fig.8.



a



b

Figure 6. Friction coefficient for the spherical specimen S4 tested against two similar steel discs at two different lubricant temperatures: 10°C (a) and 30°C (b).

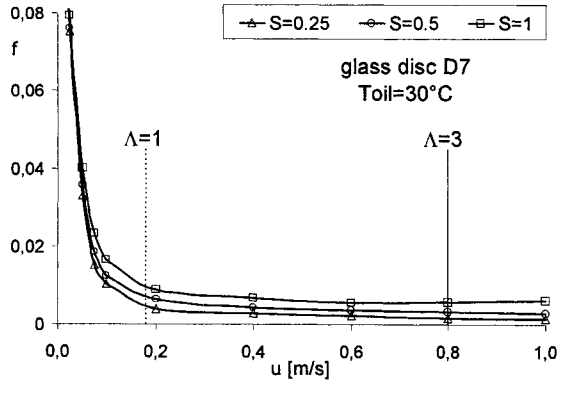


Figure 7. Friction coefficient for the cylindrical specimen C4 against a glass disc at 30°C.

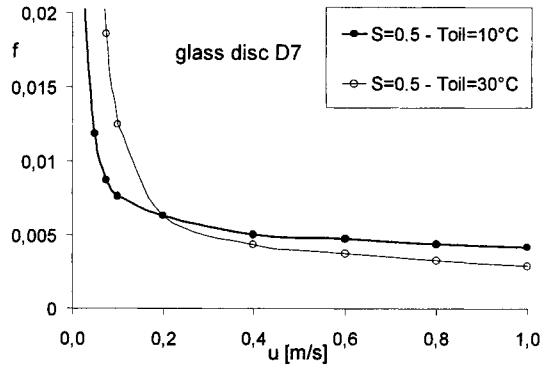


Figure 8. Comparison between the friction coefficient for the cylindrical specimen C4 against a glass disc at 10° and 30°C (S=0.5).

3.2. Conformal contacts

An experimental investigation was carried out on lubricated thrust tilting pads against different discs [8-9]. The lubricated contact occurred between the plane surface of a disc, simulating the collar supported by the thrust hydrodynamic bearing, and the pad, placed on a suitable case (the rotation of the case was obviously prevented), Fig.9. The steel pad had an outer radius of 42mm, an inner radius of 24mm and a sector angle of 30°, with a surface area of about 3cm^2. It had a spherical pivot placed at the mean radius at a distance of 0.8mm from the pad centre (pivot offset).

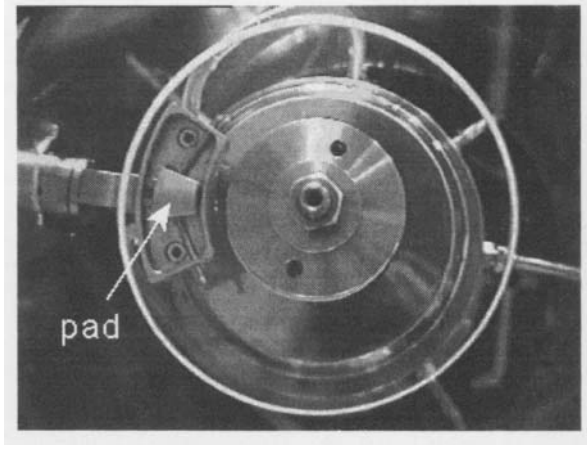


Figure 9. Pad positioned under a transparent glass disc.

Simultaneous measurements of film thickness and friction were made, while the tilt angle of the pad was also measured by optical interferometry with monochromatic light. Three different loads, F=10, 20 and 30 N, and six different speeds, Δu=0.05, 0.10, 0.15, 0.2, 0.3 and 0.4 m/s, were tested (the range of speed was chosen in order to avoid cavitation and starvation phenomena, evidenced in [8]). With Δu is indicated the circumferential speed at the mean radius (pivot position). The same lubricant as for nonconformal tests was used, with η_0=0.075 Ns/m^2 at the test temperature (20°C).

Typical results obtained using a glass disc with a chromium semi reflecting layer protected by a SiO_2 coating and a pad with $R_q \approx 0.02\,\mu m$ are reported in Fig.10. Tests were carried out in typical hydrodynamic conditions (the minimum film thickness was 10 μm under the worst tested conditions so that Λ values were always greater than some hundreds). The friction coefficient increases with Δu and decreases with F, as expected.

3.3. Comparison with theory

Under full fluid conditions the influence of the roughness is reduced. If also thermal effects can be neglected, f can be calculated as a function of the parameters present in the Stribeck number (viscosity, speed and load) by using the isothermal theory.

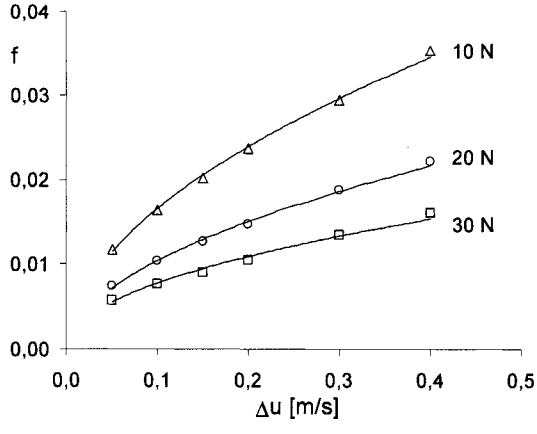


Figure 10. Friction coefficient for a tilting pad tested against a glass disc.

The following observations for nonconformal contacts regard circular contacts, but similar formulae could be developed for elliptical and line contacts.

By assuming a Newtonian behaviour of the lubricant and using a mean viscosity calculated introducing the mean contact pressure in the Barus formula ($\eta=\eta_0 e^{\alpha p_m}$, where α is the pressure-viscosity coefficient and $p_m=2/3 p_H$, with p_H Hertzian pressure), a mean velocity gradient $\Delta u/h_c$ (with $\Delta u=S\,u$ sliding speed and h_c central film thickness) and the Hertzian contact area as a reference area, it is easy to find:

$$f = \frac{\eta_0 e^{\alpha p_m}}{p_m}\frac{u}{h_c} S \qquad (3)$$

By introducing in (3) one of the commonly used formulas for h_c, e.g. the one reported in [10], the following expression can be found for a circular contact:

$$f = 1.614\,\frac{\eta_0^{0.32} u^{0.32}}{F^{0.27}}\,\frac{R^{0.22}}{\alpha^{0.53} E^{0.58}}\, e^{\alpha p_m} S \qquad (4)$$

where $E=2[(1-\nu_1^2)/E_1+(1-\nu_2^2)/E_2)]^{-1}$ is the equivalent elastic modulus (ν_1 and ν_2 Poisson ratios) and R is the radius of the spherical specimen.

It is worth noting that the Stribeck variables η_0, u and F are present in eq. (4) and that the influence of F is not immediately clear because F affects also p_m.

In Fig.11 the experimental results of Fig.5b are compared with those calculated using the formula (4) (dotted lines).

For the tested conditions, the lubricant behaviour is probably at the border between the Newtonian and the Eyring one (according with [12]; for instance, values of $\alpha p_m \approx 4$ and $\alpha\eta_0 u/R = 5\cdot10^{-10}$–$4\cdot10^{-8}$ correspond to the results shown in Fig.11). Results are therefore also comparable with those obtained by means of the formula developed in [13] for Eyring fluids under isothermal conditions. By rearranging this formula to include the variables used in this work, the

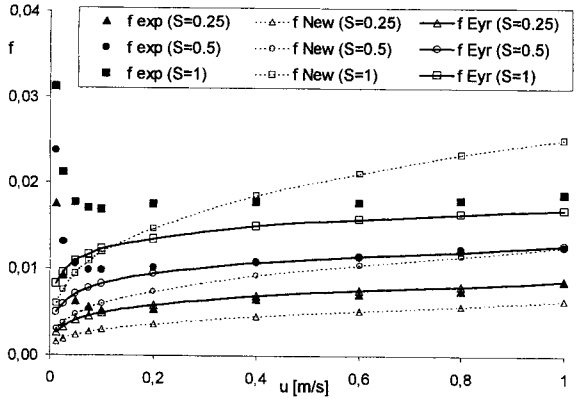


Figure 11. Experimental friction coefficient of Fig.5b compared with numerical results based on Newtonian (eq. 4) and Eyring (eq. 5) behaviour of the lubricant.

following expression can be obtained:

$$f = \frac{\tau_0}{p_H} \operatorname{arcsinh}\left(\frac{2}{5}\frac{\eta_0 e^{\alpha p_H}}{\tau_0}\frac{u}{h_c} S\right) \qquad (5)$$

where τ_0 is the Eyring stress. The evaluation of τ_0, which varies with temperature and pressure, as reported in [14], is one of the biggest problems for the use of this formula. Due to the relatively low value of the Hertzian pressure for the steel specimen S4 under the load of 20N (0.32GPa), an estimated value of τ_0=2MPa has been used which gives the results shown in Fig.11.

By introducing the dimensionless parameters used in [13], $\bar{\tau}_c = 2\eta_0 e^{\alpha p_H} uS/(h_c \tau_0)$ and $\bar{\mu} = f p_H/\tau_0$, the curves shown in Fig.12 have also been obtained. In this diagram the three theoretical curves of Fig.11 for the different S values reduce to a single curve. Under full fluid regime numerical and experimental results seem to be in good agreement. When thermal effects become evident, as for the results shown in Fig.5a, with decreasing values of f by increasing u, very different trends arise, as on the other hand pointed out by the authors of [13].

The influence of the variables present in SP is usually greater for conformal contacts

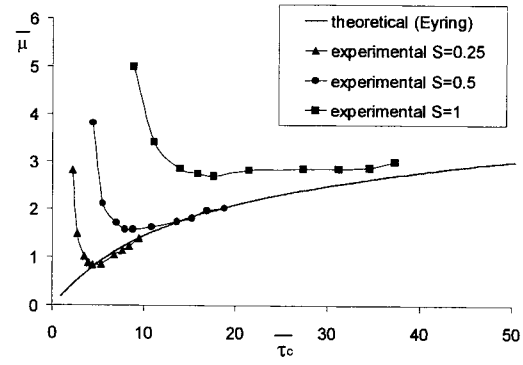


Figure 12. Comparison between experimental and numerical friction results by using the dimensionless variables of [13].

than for nonconformal contacts: f can be proportional to $(\eta_0 u/w)^x$, with x=0.5–1 (w is a load per unit length). For hydrostatic thrust bearings, it was found that x=1 [15]. Experimental results obtained in [9] with thrust bearings tilting pads are in good agreement with those obtained by a FEM code developed for the solution of the two-dimensional Reynolds equation [8]. Some results, obtained under different loads by using two glass discs (with and without SiO_2 coating) and a steel disc [9], are plotted against the service parameter $[\eta_0 \Delta u/(F/L)]^{0.5}$ (where L is the pad length in the radial direction) together with the numerical ones in Fig.13.

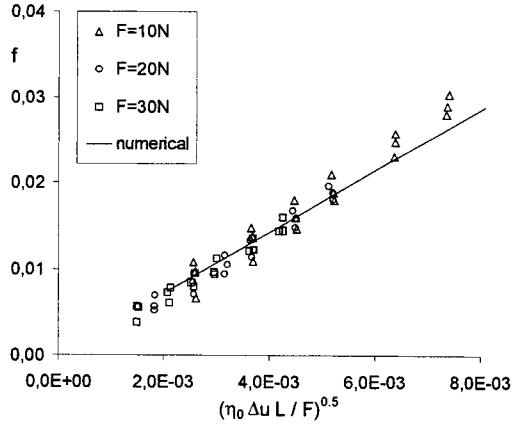


Figure 13. Experimental friction coefficient for a tilting pad against glass and steel discs compared with numerical results (line).

Note that $\Delta u=2u$ is used under these pure sliding conditions (S=2).

Experimental and numerical results show a quite linear variation with the service parameter used, and the following relation can be found for the friction coefficient:

$$f = C \frac{\eta_0^{0.5} \Delta u^{0.5}}{F^{0.5}} L^{0.5} \qquad (6)$$

where C is a constant (3.57 for the shown results).

4. MIXED AND BOUNDARY RESULTS

4.1. Experimental nonconformal data

Mixed lubrication conditions, with their typical decreasing in f by increasing u, were established for nonconformal contacts by using specimens and discs with higher roughness and/or by testing smaller specimens. Actually, lower values of Λ can be obtained both with an increase in σ and with a decrease in h (see the formula (1)). Cylindrical specimens with a diameter Φ=42mm and roughness up to R_q=0.6μm were tested in [5] against discs with surface roughness up to R_q=1.8μm. In [6] and [7] cylindrical specimens with diameters of 8, 4 and 2mm and a spherical one with Φ=24.606mm were tested in addition to the specimens C4 and S4 mentioned in the previous section. An example of results is shown in Fig.14 (spherical specimen S4, against the steel disc A5 with R_q=0.15μm). Once again, the dotted vertical line indicates Λ=1 (Λ=3 was reached at u=1.6m/s in this case).

The results obtained in [5-7] proved that friction is influenced by dimensions and shape of the specimens, as well as by roughness and materials. In particular, it was found that, with regard to friction, the minimum film thickness is more important than the central one and than the Hertzian pressure for a comparison between bodies with different shape.

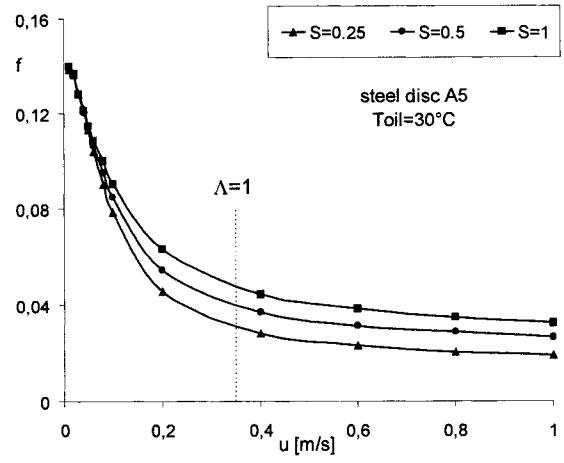


Figure 14. Friction coefficient for the spherical specimen S4 against a steel disc with a relatively high roughness; T=30°C.

In Fig.15 the evolutions of the friction coefficient for two cylindrical and two spherical specimens tested against the steel disc A9 are shown as a function of Λ. The cylindrical specimen R2, with Φ=4mm, and the spherical one S3 (Φ=24.606mm), which show very close values of f, worked almost with the same minimum film thickness under the tested conditions (the same was for R4, with Φ=8mm, and S4, with Φ=41.275mm). Evolutions similar to the ones of Fig.15 were also found by using the other steel discs.

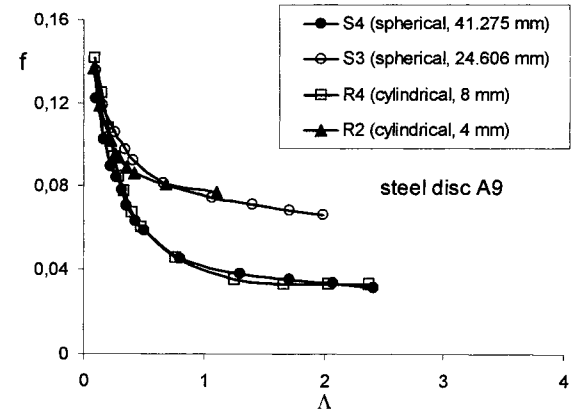


Figure 15. Comparison of friction evolutions for different specimens against a steel disc (S=0.5).

This concurrence of results was also present in tests with the glass disc (although the values of f for the specimens with similar minimum film thickness were not so close as for the tests with the steel discs).

Because of the importance of the minimum film thickness found in these tests, this parameter has been used for the evaluation of the limiting Λ values between lubrication regimes in this work.

Some indications of arising of boundary lubrication conditions, with almost constant f, were found with the most rough discs. In Fig.16 results obtained with the spherical specimen S4 and the cylindrical one C4 against the steel disc A5 are shown. The limiting values of f under boundary conditions were about 0.15 for all specimens.

It is worth stressing that the evolutions of the friction coefficient under mixed and boundary conditions can be quite different when different lubricants are used [16-17]; the presence of antifriction and antiwear additives in particular can alter the shape of these curves [18-21]. In particular the value of the friction coefficient under boundary lubrication conditions can vary depending on many factors, such as materials, roughness and boundary layer properties [22-26].

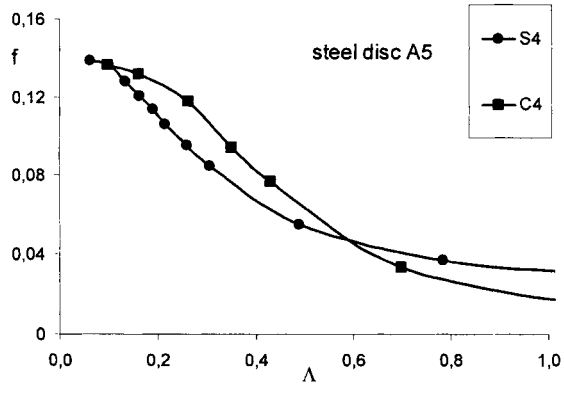


Figure 16. Comparison of friction evolutions for the cylindrical specimen C4 and the spherical specimen S4 against a steel disc with a relatively high roughness; low speed conditions (S=0.5).

4.2. Wear and other problems

Depending on many parameters, such as surface roughness and hardness of materials, the surfaces of the lubricated bodies can be worn in different ways under mixed and boundary lubrication conditions. Actually, under the most critical working conditions experienced in [5], some wear occurred, as evidenced by a subsequent analysis of the discs and specimens surfaces [27]. The instantaneous friction values recorded each second for the two cylindrical specimens C4 and C2 (respectively made of steel and aluminium) are shown in Fig.17. As already mentioned, for each speed and slide/roll ratio sixty friction values were usually recorded for one minute (each point reported in all previous diagrams is the mean of these values). The decrease of f by increasing the S value, shown at low speeds in particular for the aluminium specimen C2, clearly indicates the presence of running-in phenomena. On the other hand, when wear becomes negligible (as for the steel specimen at high speed), the usual increase of f with S was found.

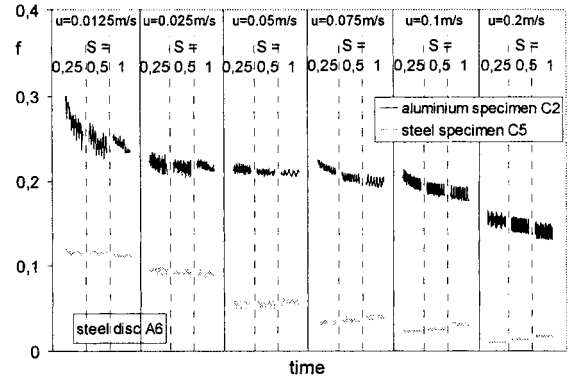


Figure 17. Evolutions of friction coefficient for the cylindrical specimens C2 and C5 under mixed lubrication conditions in presence of wear.

Wear introduces an additional variable in the friction studies; it causes a variation of roughness during tests, making the compari-

son between results more difficult. The calculated Λ is not really significant when wear occurs. A more adequate Λ could be evaluated considering the variation of roughness during the process, as for instance reported in [28].

It is also worth noting that roughness can both increase (e.g. in presence of scuffing) or decrease (running-in process) its value. Occurrence of wear is also related to vibrations and friction fluctuations, as evident in Fig.17. Results obtained by means of the other specimens showed some fluctuations too, but in a reduced manner than for C2. These variations were also related to the rotational frequency of the disc. Some defects in machining and in positioning produced vibrations and possible load fluctuations. The one minute averaged values were considered a sufficiently good mean value of friction under most tested conditions. However, the influence of a variable load with particular frequencies can produce some variations on the results, as shown in [29], and a more detailed investigation should be necessary.

To limit the occurrence of wear, tests described in [6] and [7] were made using only steel specimens and starting from the highest speed.

4.3. Theoretical observations

It is not easy to find a formula to evaluate the friction coefficient under mixed lubrication conditions. As reported in many studies (see for instance [24]), complex phenomena related to the asperity contacts and to the chemical-physical characteristics of lubricant and materials are not easy to combine in a single expression. For instance, to include boundary lubrication effects, roughness should be present in each potential formula, but it cannot easily be introduced by using only one parameter such as R_q; the above mentioned wear phenomena introduce additional variables.

Two terms are usually present in the expressions for the evaluation of the friction coefficient under mixed lubrication conditions, as reported for instance in [3]:

$$f = f_s \frac{F_s}{F} + f_h \frac{F_h}{F} \tag{7}$$

where f_s is the friction coefficient under boundary lubrication, f_h is the one for the full lubricated conditions, F_s and F_h the part of the total load F carried respectively by the asperity contacts and by the full lubricated zones. As evident, the two terms are respectively related to boundary and full fluid lubrication regimes. Examples of formulae, developed for different specific applications, are reported for instance in [4] and [30].

5. DISCUSSION

Although not present in the usual Stribeck diagrams, a reduction of f, probably related to thermal effects, can also occur by increasing the speed after reaching the full fluid regime for nonconformal contacts. However, shape, dimensions and temperature can alter the trends of the friction curves, as shown in Fig.18, and in particular the limiting Λ values for which the transitions between the different regimes of lubrication occur.

The zone with almost constant friction, usually considered as typical for the EHD

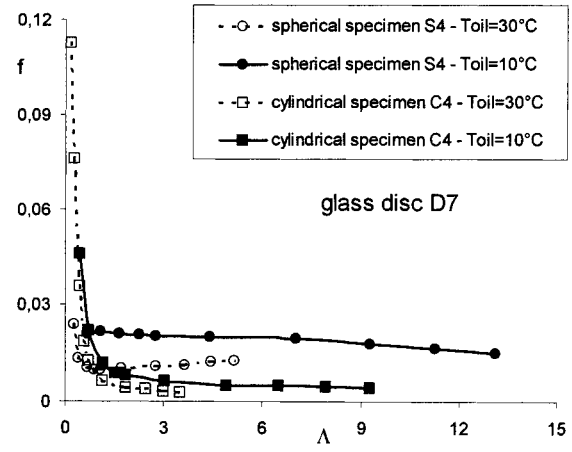


Figure 18. Comparison of friction evolutions for the specimens C4 and S4 at different lubricant temperatures (S=0.5).

regime [17], can be more or less evident depending on geometry and thermal effects.

The limiting Λ values for the different regimes can be only roughly estimated, also because it is not easy to define a significant value for the roughness to put in the Λ formula. The composite roughness has been calculated in this work by introducing the initial out-of-contact roughness of the bodies, which can vary because the asperity heights can be reduced by the high pressures in the contact (particularly if nonconformal) and in presence of wear. However, the results obtained seem agree quite well with the usual limiting values found in the literature (see for instance [10], [22], [28]):

$\Lambda < 0.1 - 1$: boundary lubrication regime;

$0.1 - 1 \leq \Lambda \leq 2 - 5$: mixed lubrication regime;

$\Lambda > 2 - 5$: full fluid lubrication regime.

It is worth noting that, even if a value of $\Lambda=3$ is often considered as the limiting value between mixed and full fluid regimes for elastohydrodynamic lubricated contacts, different values can be found in the literature (for instance, $\Lambda=2$ [31], $\Lambda=1$–1.5 [17] and $\Lambda=5$ [32] are indicated for ball-plane contacts).

Some possible different trends of f are shown in Fig.19. The borders between the lubrication regimes as well as the one between the thermal and the isothermal zones can shift so that different trends, such as the ones reported with dashed lines, can occur.

Under boundary lubrication, an evolution of f as that indicated with the dashed line can be produced for instance by destructive wear. Such phenomena were generally not present in the experimental results carried out in [6-7], because, as already mentioned, the working conditions, the dimensions and the materials of the specimens were chosen to avoid (or at least to limit) wear problems.

Under full fluid conditions, the four cases of the dashed lines identified by A, B, C and D in the graph, can be obtained particularly by varying the thermal boundary. The zone with constant f (line B), found in some EHL experiments, could be absent if thermal effects are predominant (line A). The decrease of f could not take place for conformal junctions, where the lubricated contacts occur over larger areas than for the nonconformal ones. However, heating problems can also arise in this case at very high speeds, when also turbulence effects (that usually increase friction) can be present; depending on which effect prevails, f could follow line D or C.

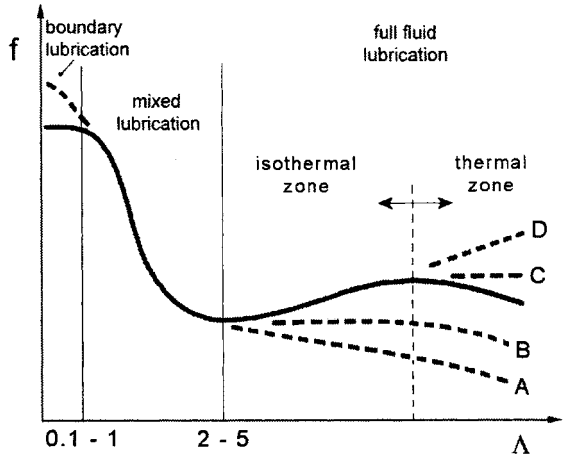


Figure 19. Generalized lambda (Λ) diagram with lubrication regimes.

Particularly at very high speed, the trend of f can also be influenced by the possibility of the lubricated contact to increase its film thickness without limitations. The behaviour of the nonconformal contacts investigated in the experimental works reported in this paper can be different from those of the contacts between the rolling bodies and the race of a real rolling bearing, where the increase of the film thickness is limited by the geometry of the bearing. The same situation occurs in a journal bearing, where the film thickness cannot increase over the radial clearance.

These limiting values of the film thickness, related to the geometry of the machine component, were usually not approached in the past, but today very reduced clearances, e.g. in precision bearings, and extreme working conditions can be reached. Advanced precision rolling bearings and journal bear-

ings have been for instance recently tested up to a speed of 25000rpm (corresponding to a journal sliding speed of 70m/s) by means of a new purposely designed and manufactured rig [33-34]. The journal bearings have been in particular designed to work also under extremely severe conditions of load (up to 45000N) by using a purposely developed computer program that takes into account both elastohydrodynamic and thermal aspects [35]. With this program a particular shape of the shaft has been designed, which has proved to work well during some experimental tests, by showing how the scientific understanding can be conveniently translate into design methodology. Extreme working conditions have also been tested in which mixed and also boundary lubrication regimes were reached, where the characteristics of the coatings used become very important [36]. Micro and also nano aspects [37] are today of fundamental importance, for instance in micromechanics components.

6. CONCLUSIONS

This work reviews studies on friction, mainly experimental, carried out with several kinds of lubricated contacts working under various conditions. Theoretical and experimental results are compared under full lubricated conditions.

The evolutions of the friction coefficient, in particular for nonconformal contacts, can be very different from those of the typical Stribeck or Λ diagrams. Many variables such as shape and dimension of a lubricated contact, roughness, materials, lubricant characteristics and thermal effects influence the friction trends.

It is not easy to find a formula for the evaluation of f. However, under full fluid conditions, relatively simple formulae, based on appropriate models of the lubricant behaviour, have been developed that provide results in sufficiently good agreement with the experimental ones when thermal effects are negligible.

Under mixed and boundary lubrication conditions the evaluation of the friction coefficient is more difficult, also for the possible presence of wear.

A value of Λ based on the minimum film thickness seems more significant than that calculated with the central film thickness with regard to the results under mixed lubrication conditions and to the limiting values between lubrication regimes. However, the central film thickness should be used for the evaluation of the friction coefficient under full fluid conditions.

REFERENCES

1. E. Ciulli, G. Rizzoni and J. Dawson, Modeling of SI and CI Engines (SP-1168), SAE Paper 960357 (1996) 181.
2. D. Dowson, Fundamentals of Friction: Macroscopic and Microscopic Processes, Kluwer Academic Publishers, Dordrecht, Netherlands, (1992) 325.
3. E. Ciulli, Fundamentals of Tribology and Bridging the Gap Between the Macro-and Micro/Nanoscales, NATO Sciences Series, Kluwer Academic Publishers, Dordrecht, The Netherlands, (2001) 725.
4. M. Nonogaki and T. Nakahara, Japanese Journal of Tribology, 38, 4, (1993) 499.
5. R. Bassani and E. Ciulli, Thinning films and tribological interfaces, Elsevier, Amsterdam, (2000) 759.
6. R. Bassani, E. Ciulli, Tribology Research: From Model Experiment to Industrial Problem, Elsevier, Amsterdam, (2001) 403.
7. R. Bassani, E. Ciulli and M. Lorenzetti, 2nd World Tribology Congress, paper 920, Vienna, 3-7 September 2001.
8. R. Bassani, E. Ciulli and P. Forte, Proceedings of the Institution of Mechanical Engineers, 212, J, (1998) 271.
9. R. Bassani, E. Ciulli and P. Forte, 2nd World Tribology Congress, paper 921, Vienna, 3-7 September 2001.

10. B.J. Hamrock, Fundamentals of fluid film lubrication, McGraw-Hill, New York, 1994.
11. C. Bovington, S. Korcek and J. Sorab, Lubrication at the frontier: The role of the interface and surface layers in the thin film and boundary regime, Elsevier, Amsterdam, (1999) 205.
12. C.R. Evans and K.L. Johnson, Proc Instn Mech Engrs, 200, C, (1986) 313.
13. B. Jacod, C.H. Venner, P.M. Lugt, ASME Journal of Tribology, 123 (2001) 248.
14. C.R. Evans and K.L. Johnson, Proc Instn Mech Engrs, 200, C, (1986) 303.
15. R. Bassani and B. Piccigallo, Hydrostatic Lubrication, Elsevier, Amsterdam, 1992.
16. H. Yu and J.B. Medley, Elastohydrodynamics-'96: fundamentals and applications in lubrication and traction, Elsevier, Amsterdam, (1997) 475.
17. S. Bair and W.O. Winer, ASME Journal of Lubrication Technology, 104 (1982) 382.
18. J. Sugimura, J. Kim, S. Gondo and Y. Yamamoto, Lubrication at the frontier: The role of the interface and surface layers in the thin film and boundary regime, Elsevier, Amsterdam, (1999) 829.
19. G. Tripaldi, A. Vettor and H. Spikes, Subjects in engine oil rheology and tribology (SP-1209), SAE Paper 962036, (1996) 73.
20. G. Guangteng and H.A. Spikes, Tribology Transactions, 40, 3 (1997) 461.
21. G. Tripaldi, S. Fattori, R. Nodari and A. Vettor, Lubrication at the frontier: The role of the interface and surface layers in the thin film and boundary regime, Elsevier, Amsterdam, (1999) 751.
22. J.A. Williams, Engineering Tribology, Oxford University Press, Oxford, 1994.
23. S. Kato, E. Marui and M. Hashimoto, Tribol. Trans. 41, 3, (1998) 341.
24. K.A. Blencoe, G.W. Roper and J.A. Williams, Proc. Instn. Mech. Engrs., 212, J, (1998) 391.
25. H.A. Spikes, Proc. Instn. Mech. Engrs., 214, C, (2000) 23.
26. Y.-Z. Hu and S. Granick, Tribology Letters 5, (1998) 81.
27. R. Bassani, E. Ciulli, F. Di Puccio and M. Lorenzetti, Mecânica Experimental, 6, (2001) 117.
28. H. Czichos, Tribology, Elsevier, Amsterdam, 1978.
29. A.A. Polycarpou and A. Soom, ASME Journal of Tribology, 117, (1995) 261.
30. D.P. Hess and A. Soom, ASME Journal of Tribology, 112, (1990) 147.
31. G. Guangteng and H.A. Spikes, Elastohydrodynamics-'96: fundamentals and applications in lubrication and traction, Elsevier, Amsterdam, (1997) 159.
32. A.R. Lafountain, G.J. Johnston and H.A. Spikes, Tribology Transactions, 44, 4 (2001) 648.
33. R. Bassani, E. Ciulli, S. Manconi, B. Piccigallo and G. Oddone, AITC2002, 3rd AIMETA International Tribology Conference, Vietri sul Mare, Salerno, 18-20 September 2002.
34. R. Bassani, E. Ciulli, F. Di Puccio, B. Piccigallo, R. Squarcini and I. Moretto, AITC2002, 3rd AIMETA International Tribology Conference, Vietri sul Mare, Salerno, 18-20 September 2002.
35. R. Bassani, E. Ciulli, B. Piccigallo and G. Oddone, Proceedings of the first 2000 AIMETA International Tribology Conference, Gruppo Tipografico Editoriale srl, L'Aquila, (2000) 347.
36. R. Bassani, E. Ciulli, M. Labardi and M. Allegrini, 2nd World Tribology Congress, paper 902, Vienna, 3-7 September 2001.
37. B. Bhushan, J.N. Israelachvili and U. Landman, Nature, 374, (1995) 607.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

A Macro-Micro-Fatigue approach for predicting fatigue life in mixed-EHL contact.

Dmitry Epstein[a], Q. Jane Wang[a], Leon M. Keer[a], Herbert S. Cheng[a], Dong Zhu[b], Stephen Harris[c], and Arup Gangopadhyay[c]

[a] Department of Mechanical Engineering, Northwestern University,
Evanston, Illinois 60208-3109, USA

[b] Eaton Corporation
26201 Northwestern Highway, Southfield, Michigan 48037, USA

[c] Ford Motor Company,
Dearborn, Michigan 48124, USA

Surface contact fatigue is one of the major failures that threaten the life of engineering components. A Macro-Micro-Fatigue approach for analyzing the effect of roughness on fatigue life in mixed elastohydrodynamic lubrication contact is developed and presented. This method involves two macro-micro approaches, a macro-micro mixed-lubrication approach for lubrication analysis and a macro-micro contact stress approach. The macro-micro approaches allow not only efficient modeling of pressure and film thickness distributions over the entire area of contact, but a detailed account of asperity interaction effects as well. A three-stage process is designed to comprise a rough-surface contact solver, a simplified mixed elastohydrodynamic lubrication solver, and a hydrostatic rough-contact stress solver. Subsurface stresses are calculated by means of the fast Fourier transform, and fatigue life is estimated using a version of the Zaretsky fatigue model. Typical point mixed-EHL contacts of rough surfaces are analyzed and results are presented on a comparative basis.

1. INTRODUCTION

Compact design of mechanical components, such as gears, CVT parts, and bearings, requires engineering surfaces to work under high contact stresses. Improving the fatigue life of such components is a challenge, which demands the ability to analyze the performance of these components in mixed lubrication and predict their life in service. In mixed lubricated contacts, asperity interaction induces high and irregular stresses in the near-surface material layer [1]-[3], causing pitting failure of surfaces. Fatigue life analysis of components working in a mixed lubrication regime must therefore include the effect of surface roughness. In the case of dry contacts, or boundary lubricated contacts in which hydrodynamic effects can be neglected, it is often possible to isolate a fairly small portion of the contact that experiences the most severe stresses and use that critical area for fatigue analysis. However, in mixed lubrication where hydrodynamic effects exist, the entire contact area has to be analyzed on some level in order to find the distribution of fluid pressures. This demands an approach capable of processing a large area of surface interaction.

Analyzing the fatigue life of such lubricated rough contact requires the solutions of mixed lubrication, asperity contact, and subsurface stresses to be incorporated into a contact fatigue model. A full-scale micro EHL model [4]-[5] may be used for the mixed lubrication analysis. However, contacts arising in

many applications can be relatively long and/or wide, and significant computation times are anticipated if a deterministic micro-EHL model is used for such problems. Moreover, performing adequate analysis of roughness-induced contact stresses and surface pitting in large contacts can present serious difficulty. The average flow model of Patir and Cheng [6] has been proven effective in the studies of some conformal-contact problems where the asperity contacts are localized in a few characteristic regions [7-9]. In these analyses, lubrication is processed through the macro-scale average flow model while the contact is calculated by means of an off-line deterministic asperity contact model. This method has the advantage of being able to analyze large surfaces in mixed lubrication and is evolving to solve counterformal-contact problems [10]. It is, by nature, an interactive deterministic-stochastic modeling method and named by Jiang et al. [10] as a macro-micro approach. By comparing the result from the macro-micro approach with those from Zhu and Hu's full-scale micro EHL model [4-5], a recent investigation [11] has proven this approach to be applicable in the region where the ratio of average film thickness to the RMS roughness of surfaces (h_a/R_q) is no less than 1.

Using the macro-micro mixed-lubrication approach, the average flow model analyzes the hydrodynamic pressure distribution over the entire contact area and interfaces with a micro-asperity contact model. First, the entire area of contact is treated as the contact of two pseudo-smooth surfaces, where their actual roughness is taken into account through flow factors and a pressure-gap curve, which will be described later. The macro-contact area is then divided into patches of a suitable size, and taking into account asperity interaction, analyzed using a fine mesh. This procedure defines the second macro-micro approach. Subsurface stresses for each patch are then calculated and their contributions to fatigue life summed up. The processes of these simplified macro-micro analyses result in an efficient approach, the Macro-Micro-Fatigue (MM-Fatigue) approach for contact fatigue analysis for components under mixed lubrication conditions.

Nomenclature

a = half-width of the Hertzian contact
c = stress exponent
d = depth exponent
e = Weibull slope
$erf(x)$ = $\int_0^x \frac{2}{\sqrt{\pi}} e^{-y^2} dy$
F = fatigue integral
h = film thickness, or gap between surfaces
h_0 = average gap in a patch
h_a = average gap in the entire contact area
h_{cp} = compliance (separation of undeformed rough surfaces)
$\overline{h}$, $\overline{h}_{cp}$ = average (smooth) gap and compliance
$\overline{H}$, $\overline{H}_{cp}$ = dimensionless gap and compliance, $\overline{H} = \overline{h}/R_q$, $\overline{H}_{cp} = \overline{h}_{cp}/R_q$
n_x, n_y = number of nodes in a fine (microcontact) mesh
N_x, N_y = number of nodes in a coarse (macrocontact) mesh
N = fatigue life (the number of cycles until failure)
p = total pressure
p_0 = average pressure over a patch
p_c, p_f = contact (asperity) pressure and fluid (lubricant) pressure
p_{co} = plastic cutoff pressure
p_h = maximum Hertz pressure
R_q = composite root mean square roughness
S = probability of survival
u_1, u_2 = velocities of the interacting surfaces
v = elastic surface displacement (deformation)
v_{cs} = center-surface deformation
V = stress-affected material volume
V_0 = characteristic volume
w = composite roughness
x = coordinate in rolling direction
y = coordinate perpendicular to rolling direction
z_0 = characteristic depth
η = lubricant viscosity
κ = constant stress factor
λ = lambda ratio, h_a/R_q
σ_0 = characteristic stress
σ_e = equivalent stress in an elementary volume
σ_{fl} = fatigue limit stress
σ_{eff} = effective stress
ϕ_x, ϕ_y, ϕ_s = Patir-Cheng flow factors
$\phi'_x, \phi'_y, \phi'_s$ = modified flow factors

This paper gives the formulation and solution methods for this Macro-Micro-Fatigue (MM-Fatigue) approach. Numerical simulations are performed for several cases of point contacts of real rough surfaces, and results are presented on a comparative basis.

2. SOLUTION OF THE CONTACT PROBLEM

2.1. The macro-micro approach

The macro-micro approach is an interactive deterministic-stochastic modeling process. Like the full-scale micro EHL model, it considers that the lubricant film is continuous in non-contacting regions and that asperity contact occurs when the local film is sufficiently thin. Although this approach computes the film thickness from the deformed average gap, it treats flows and asperity contact in two separate submodels: a flow model for lubricant flows and an off-line contact model for asperity contact pressure. Compared with the previous use of the macro-micro approach in counterformal contact [10], the current work employs the average flow model of Patir and Cheng [6] to describe the influence of asperities on lubricant flows through modified flow factors and an improved asperity contact model to calculate asperity pressure. Instead of digitizing the Reynolds equation over a fine mesh that reflects the topography of a rough surface, the average Reynolds equation below may be solved in a much coarser mesh, where the asperity influence is averaged in the vicinity of each computation point.

$$\frac{\partial}{\partial x}\left[\phi_x \frac{\bar{h}^3}{12\eta}\frac{\partial p_f}{\partial x}\right]+\frac{\partial}{\partial x}\left[\phi_y \frac{\bar{h}^3}{12\eta}\frac{\partial p_f}{\partial y}\right] = \frac{(u_1+u_2)}{2}\frac{\partial \bar{h}}{\partial x}+\frac{(u_1-u_2)}{2}R_q\frac{\partial \phi_s}{\partial x} \quad , \tag{1}$$

where ϕ'_x, ϕ'_y, and ϕ'_s are the flow factors, which will be explained next.

The lubricant film thickness, or gap, h can be approximated as the sum of the separation between the average profiles of the deformed surfaces and the combined roughness of the two surfaces (assuming that the deformation of asperities is relatively small). The latter affects the micro-flows and should be considered in the average flow (AF) model, while the former should be analyzed through a macro-scale computation:

$$h = h_0 + B_x x^2 + B_y y^2 + \underbrace{v_{cs}(x,y)}_{\text{to be considered in macro scale}} + \underbrace{w(x,y)}_{\text{to be considered in the AF model}} \quad , \tag{2}$$

where v_{cs} is the elastic deformation of the center-surfaces and w is the composite roughness.

The macro-micro approach includes three components: (1) the macro center-surface determination, (2) the flow-factor and pressure calculation, and (3) micro-scale asperity stress calculation. The film thickness used in Equation (1) is the average film thickness, defined by $h = \bar{h} + w$. The roughness effect is accounted for through the AF model, utilizing modified Patir-Cheng flow factors. Patir and Cheng [6] defined their flow factors in terms of undeformed geometry, or compliance $\bar{h}_{cp}$. The flow factors used in the present study differ in that they take into account the deformation of the interacting surfaces, by using the following approximate relation [12]:

$$\bar{H} = \bar{H}_{cp}\left[0.5 + 0.5 erf\left(\frac{\bar{H}_{cp}}{\sqrt{2}}\right) + \frac{e^{-0.5\bar{H}_{cp}^2}}{\sqrt{2\pi}\bar{H}_{cp}}\right], \tag{3}$$

where $\bar{H} = \bar{h}/R_q$, $\bar{H}_{cp} = \bar{h}_{cp}/R_q$, and R_q is the root mean square composite roughness. Consequently, the relationship between flow factors in (1) and Patir-Cheng flow factors, ϕ_x, ϕ_y, and ϕ_s, can be expressed as

$$\phi'_x = \phi_x\left(\frac{\bar{h}}{\bar{h}_{cp}}\right)^3, \quad \phi'_y = \phi_y\left(\frac{\bar{h}}{\bar{h}_{cp}}\right)^3, \text{ and } \phi'_s = \phi_s .$$

Equation (1) relates lubricant film thickness $\bar{h}$ to the lubricant pressure p_f. In a mixed EHL problem asperity contact pressure has to be modeled as well. The present study employs a simplified interactive approach similar to that in Jiang at al. [10]. This approach utilizes a semi-empirical contact pressure-gap relation tabulated with an off-line dry contact

solver. The relation is applied to the mixed lubrication based on the information of the lubricant film thickness and load balance.

The mixed-EHL point contact is analyzed. Figures 1 and 2 illustrate the average central film thickness, as influenced by rolling velocity (Fig. 1) and roughness (Fig. 2). The results, obtained by means of Zhu and Hu's full-scale micro-EHL model [4-5], are also presented for comparison. Here, the radius of curvatures of the mating bodies are: R_1 = 19.05 mm and R_2 = 8 m (infinitely large). The material properties are: Young's modulus, E = 200 GPa, Poisson's Ratio, ν = 0.3, and hardness. H = 7 GPa.

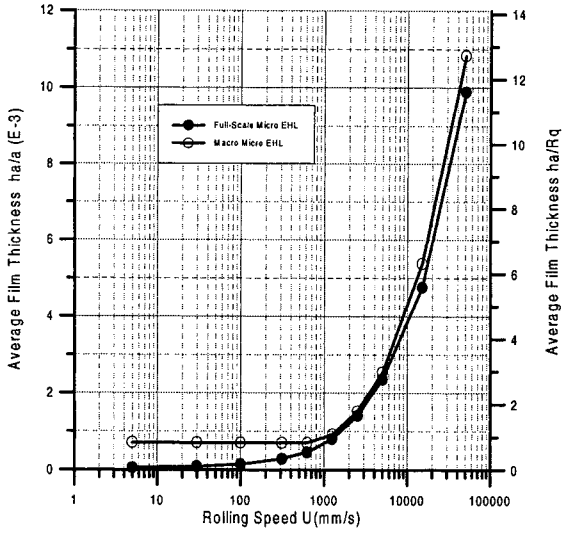


Figure 1. Variations of the average central film thickness as a function of rolling speed (W = 800 N, R_q = 0.4 μm).

Data for a typical mineral lubricant is used in the calculation, and the lubricant properties are: viscosity, $\eta_0 = 0.96 \times 10^{-7}$ Ns/mm^2, pressure-viscosity coefficient, α = 0.01820 mm^2/N, and index in the Roelands viscosity-pressure relationship is 0.67.

The comparison suggests that the use of the macro-micro approach may yield reasonable film thickness and pressure distributions on a local average sense for operating conditions corresponding to the ratio of average film thickness to the RMS roughness (h_a/R_q, or the λ ratio) no less than 1 [11]. In the region of $\lambda \geq 1$, the film pressure may be approximated by the superposition of the hydrodynamic pressure and the asperity contact pressure. In the region of $\lambda < 1$, the film pressure may be simply estimated by an asperity contact model (such as DryContFast). Such dry contact pressure may represent the upper boundary of the mixed-lubrication pressure.

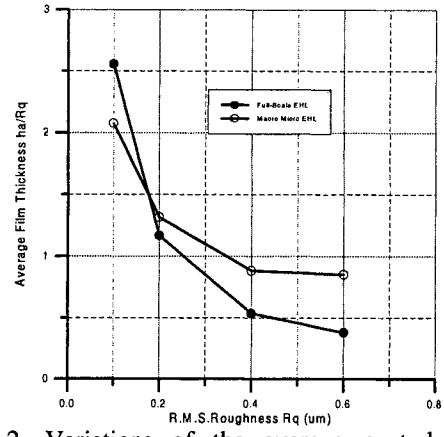


Figure 2. Variations of the average central film thickness as a function of roughness (W=800N, U=625mm/s).

Most of the disagreement with the more exact solution appears in the low speed and heavy asperity contact regions where asperity interaction is predominately the mechanism supporting the load. Compared to the universal treatment of the hydrodynamic pressure and asperity contact in the full-scale micro EHL model [4-5], the macro-micro approach superimposes the asperity contact pressure obtained from an off-line contact simulation to the hydrodynamic pressure from the average flow analysis, resulting in an under-estimation of asperity deformation but an over-estimation of the average asperity pressure. The next step in the macro-micro analysis partly compensates for this effect.

2.2. The macro-micro contact stress approach

With the consideration that the distance between the contacting surfaces at a given surface location is the gap h, and w is the composite roughness, the

discrete formulation of the contact problem is as follows:

$$h_{ij} = -v_{ij} - w_{ij} \quad (0 \le i < n_x,\ 0 \le j < n_y); \tag{4a}$$
$$h_{ij} = 0 \quad (p_{co} > p_{ij} > p_f); \tag{4b}$$
$$h_{ij} \ge 0 \quad (p_{ij} = p_f); \tag{4c}$$
$$h_{ij} \le 0 \quad (p_{ij} = p_{co}); \tag{4d}$$
$$p_{co} \ge p_{ij} \ge p_f \quad (0 \le i < n_x,\ 0 \le j < n_y); \tag{4e}$$
$$\sum_{i=0}^{n_x-1} \sum_{j=0}^{n_y-1} p_{ij} = p_0; \tag{4f}$$
$$(n_x n_y)^{-1} \sum_{i=0}^{n_x-1} \sum_{j=0}^{n_y-1} h_{ij} = h_0 . \tag{4g}$$

Here v_{ij} are the nodal displacements, h_0 and p_0 are the average gap (film thickness) and the average pressure specified for the problem. The cutoff pressure p_{co} is introduced to account approximately for the possibility of plastic yielding at highly stressed contact locations.

The present numerical model of a rough contact in the mixed lubrication regime is called hydrostatic because the fluid pressure is assumed to have the same value, p_f, throughout the considered micro-contact patch. The fluid is assumed to redistribute itself around the surface asperities and balance any local pressure differential without generating significant viscous forces. On the other hand, EHL action is still assumed to take place at the macroscopic scale, preventing the fluid from seeping out of the macrocontact area. Hence, the fluid pressure does not equal the ambient pressure throughout the contact area, but varies on the macrocontact scale. The present model corresponds to a situation where the rolling and sliding speeds are insufficient to produce significant hydrodynamic effect on the individual patch length scale. The present hydrostatic model is similar to that described by Harris [3], who uses the Greenwood-Williamson [15] approach to describe the dry part of the contact. A direct numerical approach to rough contacts is used here.

The system of equations and inequalities (4) is solved by iteration, using two levels in the iteration algorithm. The inner iteration loop is used to solve Eqs. (4a)–(4f) with a fixed value of p_f for the nodal pressures. This level of iteration is performed using the same iteration scheme as in the dry contact solver. The outer iteration loop is used to find the value of p_f for which Eq. (4g) is satisfied. The hydrostatic fluid pressure p_f is thus determined as part of the solution, rather than obtained from the MEHL solution along with the average total pressure p_0 and the average gap h_0. The reason for calculating p_f is to improve the estimation of detailed contact pressure with the given average pressure (4f) and average gap (4g) conditions. Apart from this additional unknown and the extra equation (4g), the system of equations and inequalities (9) is similar to the discrete formulation for dry contact [13-14]. For p_f=0, the problem (4a)–(4f) reduces to a discretized dry contact problem.

Tangential contact stresses are calculated using the Amontons-Coulomb friction law, as was done for the dry contact. Subsurface stresses are calculated using a combination of FFT and MLMS [13-14].

3. FATIGUE LIFE ANALYSIS

The fatigue life model used in the present study is based on a number of published fatigue models, all of which are based on Weibull's stochastic model [16]. For an overview of fatigue life theories see Zaretsky et al. [17]. The contact fatigue life N, the number of stress cycles until failure, corresponding to a given probability of survival S is determined from the following equation:

$$\ln\frac{1}{S} = \frac{N^e}{V_0} \iiint_V \left(\frac{\sigma_{\text{eff}}}{\sigma_0}\right)^{ec} \left(\frac{z_0}{z}\right)^d dV = \frac{N^e F}{V_0} . \tag{5}$$

Here e is the Weibull slope, V_0 is the characteristic volume, V is the stress-affected material volume, c is the stress exponent, σ_{eff} is the effective stress experienced by the elementary material volume dV, σ_0 is the characteristic stress, d is the depth exponent, z is the depth at which the elementary volume dV lies, and z_0 is the characteristic depth. The physical meaning of the constants will be described later. In what follows, Eq. (5) will be referred to as the fatigue life equation, and the integral in the right-hand side of this equation will be called the fatigue integral F. The effective stress σ_{eff} is defined as follows:

$$\sigma_{\text{eff}} = \begin{cases} \kappa\sigma_e - \sigma_{fl}, & \kappa\sigma_e > \sigma_{fl}; \\ 0, & \kappa\sigma_e \le \sigma_{fl}; \end{cases} \quad (6)$$

where κ is the constant stress factor, σ_e is the equivalent stress affecting the elementary volume dV, and σ_{fl} is the fatigue limit stress. The constant stress factor κ is used to take into account the effect of the constant (i.e., non-cyclic) stresses (such as residual stresses or centrifugal hoop stresses) existing in the volume V on the fatigue life. In the present mode, the octahedral Mises stress is used as the equivalent stress:

$$\sigma_e = \frac{1}{3}\sqrt{(\sigma_x-\sigma_y)^2+(\sigma_y-\sigma_z)^2+(\sigma_z-\sigma_x)^2+6(\tau_{xy}^2+\tau_{yz}^2+\tau_{zx}^2)}. \quad (7)$$

The present fatigue life model, expressed by Eqs. (5-7), is based on three existing fatigue life theories. The basic concept of fatigue integral is due to Ioannides and Harris [18]. The incorporation of fatigue limit stress, the mathematical form of the depth factor, and the use of the Mises stress as the equivalent stress are also part of the Ioannides–Harris theory [17-18].

Stresses are produced by a cyclic rolling/sliding contact, and subsurface stress distributions computed earlier correspond to the peak contact loading, when the moving contact is directly above the material volume considered. Away from the contact, the stresses rapidly decay and are negligible. Thus, the computed stresses can also be regarded as the range of the cyclic contact stress variation during the contact cycle. Metal fatigue is known to be governed primarily by the cyclic stress range. The constant component in the applied stress may also have a noticeable effect on fatigue, and in structural fatigue, this effect is usually accounted for by a correction factor. The primary (cyclic) stress is multiplied by this factor, and the result used to estimate the fatigue life. The constant stress factor κ appearing in Eq. (6) is taken from the Ai [20] fatigue life model. The nature of the constant stress, and hence the mathematical form of the factor κ is application-specific, where one specific form of the constant stress factor can be found in [20].

The definition of the stress exponent c as a physically meaningful material constant independent of the Weibull slope is adopted from [16]. In addition, the "material factor" A, which appears in all of the above fatigue life models, has a fractional dimension and an obscure physical meaning, and as a consequence, has been removed from the present model and replaced by three constants: V_0, σ_0, and z_0.

The depth factor used in the Ioannides–Harris theory is known to fit fatigue life data for a number of rolling bearing types; however, as discussed in [19], its physical meaning is also obscure, and for this reason, Zaretsky et al. [17] abandoned this concept. On the other hand, Ai [20] introduced a different form of the depth factor but did not present an argument to justify his choice. In the present study no depth factor was used.

Subsurface stresses are computed simultaneously for points lying at the same depth directly below the grid nodes. Therefore, the fatigue integral F is also computed layer by layer, as follows:

$$F = \iiint_V \left(\frac{\sigma_{\text{eff}}(x,y,z)}{\sigma_0}\right)^{ec}\left(\frac{z_0}{z}\right)^{d} dV = \int_0^{z_{\max}} \left(\frac{z_0}{z}\right)^{d}\left[\iint_{S_g}\left(\frac{\sigma_{\text{eff}}(x,y,z)}{\sigma_0}\right)^{ec} dxdy\right]dz \quad (8)$$

where S_g is the grid area, and $z_{\max}$ is the maximum depth of interest, below which stresses contribute little to the fatigue life.

After the fatigue integral has been computed, the fatigue life is calculated using Eq. (10). The computed absolute value of fatigue life is not a meaningful number unless all of the material constants appearing in the fatigue life equation can be determined by independent experiments. The main purpose of the present research however is to provide a means of studying relative trends in the contact fatigue life for the same substrate material by comparing the calculated lives for different geometric and mechanical parameters such as surface roughness, contact macro-geometry, load, etc.

4. SOLUTION PROCEDURE

The contact fatigue analysis with the MM-Fatigue approach consists of two major parts. First, a

contact pressure distribution for the entire macro-contact is computed on a coarse grid. Micro-geometry (surface roughness) is not directly taken into account at this stage. For heavily loaded and/or low-speed contacts in the region of $\lambda < 1$, the dry contact analysis can be used. Otherwise, the lubrication model described in Section 2 should be applied. In the latter case, a coarse-grid distribution of the lubricant film thickness will be obtained in addition to the macro-contact pressure distribution.

Subsurface stresses corresponding to the macro-contact solution are calculated in order to determine their contribution to the "macroscopic" fatigue life integral F_{macro}. This integral describes the probability of fatigue crack initiation due to the macro-contact stress field, but does not include the effect of surface roughness. In most situations, fatigue life is dominated by the contribution of subsurface material layers lying at depths comparable to the macro-contact dimensions. The macroscopic fatigue life integral is usually important to rolling bearings, where subsurface-originated pitting is often the dominant failure mode. It is less important to gears, but can still be useful as a benchmark for the fatigue life integral corresponding to the roughness-induced stresses.

Next, the "microscopic" part of the macro-micro analysis is performed. To analyze roughness-induced stresses and their effect on surface pitting, the macro-contact area is divided into rectangular patches corresponding to the nodes of the coarse grid. The patches are analyzed individually, using high-resolution surface topography data. These roughness data sets are assumed to be representative of the rough surfaces in contact, and their resolution must be sufficient to contain adequate information about the short-wavelength roughness components that are potentially important to pitting initiation.

Suppose the coarse grid corresponding to the macro-contact solution has $N_x \times N_y$ nodes. In the micro-contact solution, each of these nodes corresponds to a rough surface patch discretized by an $n_x \times n_y$ grid. The average pressure p_0 and the average gap h_0 over the entire patch are:

$$p_0 = P_{IJ}, \tag{9a}$$

$$h_0 = H_{IJ}, \tag{9b}$$

where P_{IJ} is the macro-contact pressure at the node (I, J) and H_{IJ} is the macro-contact film thickness at that node.

The patch contact problem is solved by iteration, as discussed in section 2.2. At the end of the iterative solution, the micro-contact area and micro-contact pressure distributions are determined for the current patch. The corresponding tangential contact stress distribution is obtained under the proportional traction approximation (Coulomb friction). The subsurface stresses corresponding to these surface stress distributions are computed using the fast Fourier transform (FFT) technique, and the fatigue life integral F_{IJ} for the current patch is obtained by integrating the subsurface stresses.

After all patch problems corresponding to the $N_x \times N_y$ coarse-grid nodes have been solved, the fatigue life integral for the entire rough contact is obtained by summing up the contributions of all patches:

$$F_{\text{rough}} = \sum_{I=0}^{N_x-1} \sum_{J=0}^{N_y-1} F_{IJ} \,. \tag{10}$$

The integral F_{rough} describes the probability of near surface-originated pitting due to the roughness-induced contact stresses.

Due to the relatively small dimensions of each patch, the micro-contact grid dimensions n_x and n_y will remain moderate even if grid spacings need to be made small to capture the effect of short-wavelength roughness. It can be shown that successive solution of $N_x \times N_y$ patch problems defined on $n_x \times n_y$ grids requires much less CPU time than solution of a single problem for the entire rough contact on an $(N_x \times n_x) \times (N_y \times n_y)$ grid. In addition, the patch contact analysis will require only a moderate amount of computer memory. In practice, since macro-contact distributions of pressure and gap are usually smooth, the macro-micro approach outlined above can be made even faster by grouping patch contacts with close values of normal loads and analyzing only one contact per group. Thus, the number of distinct terms in Eq. (10) will generally be much less than $N_x \times N_y$.

5. RESULTS AND DISCUSSION

Two real machine surfaces were used in the simulation (data provided by Ford); their topography was measured with an optical profilometer. To better capture the effects small-scale roughness may have on contact stresses and fatigue life, scanning was performed at the relatively high resolution of half a micron in both directions. Figure 3 presents a sample of the scanned rough surface. At such high resolution a sample patch of 512 by 512 nodes has dimensions of only a quarter of a millimeter, which is smaller than the contact area in many common applications. Contact and subsurface stress calculations using a mesh of this size requires on the order of an hour on a 1.4 GHz PC, when fast numerical techniques such as FFT are used.

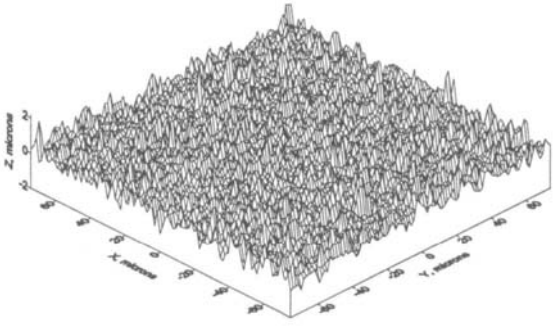


Figure 3. Rough surface used for fatigue analysis.

The objective of this particular example was to demonstrate the effect that different surface roughness has on contact fatigue. For this purpose, the original surface heights of the two contacting surfaces were scaled by factors ranging from 0.2 to 1.2, producing five pairs of surfaces with composite RMS roughness ranging from R_q = 0.15 μm to R_q = 0.94 μm. The former is typical for highly polished components such as bearings; the latter is typical for some gears and other honed or ground machine surfaces. The coefficient of friction was 0.05, which is typical for mixed lubricated contacts of steel tribological components.

The type of contact considered in this simulation was ball-on-flat, the ball having the radius of 11 mm. (Generally, our software is capable of modeling contacts of surfaces that can be locally approximated as ellipsoids or cylinders.) The velocities of both surfaces were 7 m/s. The applied normal load was 150 N. The material of both surfaces was steel with Young's modulus of 210 GPa. In the Hertzian approximation (smooth surfaces with no lubricant), the maximum pressure in the center of the contact would be p_h =1.5 GPa with a circular contact area of radius a = 0.22 mm.

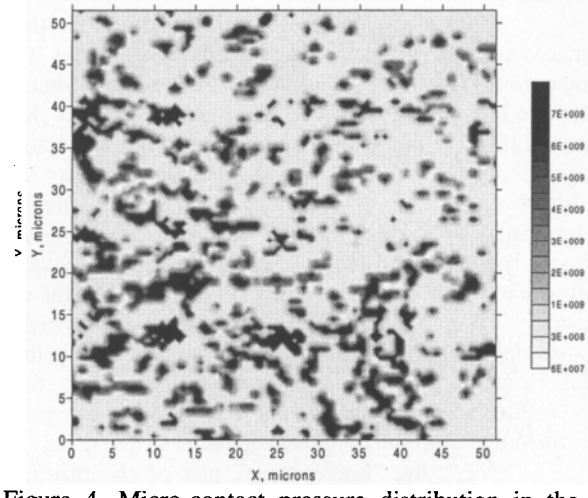


Figure 4. Micro-contact pressure distribution in the highest-pressure zone (Rq = 0.78 μm case). Note the constant hydrostatic pressure in the background.

Following the MEHL calculation, macro-contact distributions of pressure and gap were used to calculate stresses on a fine mesh, using the actual surface geometry. In the present study macro-contact pressure distributions were divided into 5 zones, and fatigue life contributions from all zones were summed (see section 3). A plastic cutoff pressure of 8 GPa was used when calculating surface pressures. Figures 4 and 5 illustrate surface pressure and Mises stress in a vertical plane, calculated for the zone with the highest pressures. From these results it is evident that asperity interaction produces very high stresses in the near-surface layer, which are likely to result in surface-initiated fatigue.

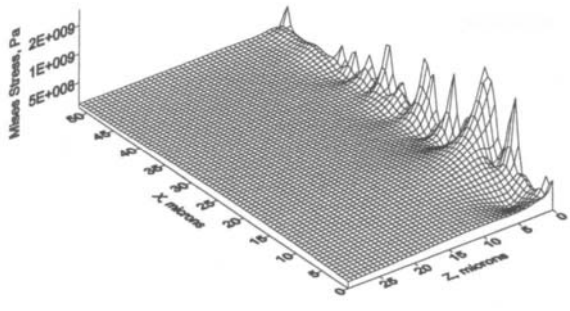


Figure 5. Von Mises stress profile in a vertical plane in the highest-pressure zone (Rq = 0.78 μm case).

Fatigue life was calculated for 50% probability of survival (Figure 6). Reasonable values for the Weibull slope (2.0) and the stress exponent (5.0) were used. However, since most of the other material constants were not known, comparative results are presented with the highest fatigue life normalized to unity. The projected fatigue life decreases as more asperities come into contact and the load ratio (the fraction of load supported by asperities) increases from nearly zero in the R_q = 0.15 μm case to near unity in the R_q = 0.94 μm case. In all of the cases roughness-induced stresses made by far the largest contribution to the fatigue life, when compared to macro-contact (Hertzian) stresses. Saturation occurs as roughness approaches its highest value in this study. The bulk of the contact pressures in areas of direct asperity contact correspond to the plastic cut-off pressure. As a result, the contribution to the fatigue life depends mainly on the density of asperity contacts, which in turn experiences a saturation as the load ratio approaches unity. One should view this result with caution since a large-scale plasticity effect may require an exact elastic-plastic analysis rather than the use of a cut-off stress to represent plasticity.

6. CONCLUSIONS

The paper presents the formulation and solution method of the Macro-Micro-Fatigue (MM-Fatigue) approach. This method involves two macro-micro approaches, a macro-micro mixed-lubrication approach for lubrication analysis and a macro-micro contact stress approach for contact stress evaluation. It includes the use of the Macro-Micro technique to calculate average pressure and gap in the contact area; the use of a similar Macro-Micro principle to conduct stress analysis of contacts using real roughness data; and the use of a Weibull-type stochastic cyclic fatigue model. What distinguishes the present approach from previous efforts in this area is the ability to adequately capture effects of small-scale roughness on contact stresses and fatigue for relatively large contact areas, while keeping computation times and memory usage within reasonable limits. The MM-Fatigue approach was illustrated by a number of calculations that used real machined surfaces with varying roughness under relatively severe contact conditions. The importance of taking roughness effects into account is underscored by the fact that the character of the contact as well as the predicted fatigue life had a strong correlation with the roughness of the contacting surfaces.

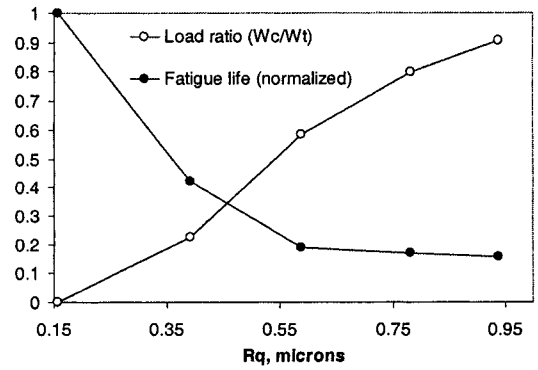


Figure 6. Load ratio (fraction of load supported by asperities) and fatigue life (normalized) vs. surface roughness.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the supports from US National Science Foundation and Office of Naval Research, Ford Motor Company, as well as the Co-Center for Surface Engineering and Tribology at Georgia Institute of Technology, and Northwestern University. The authors also wish to thank Dr. T. Yu for some of the mixed lubrication calculation.

REFERENCES

1. Littmann, W. E. and Widner, R. L., 1966, "Propagation of contact fatigue from surface and subsurface origins". ASME Journal of Basic Engineering, Vol. 88, pp. 624-636.
2. Miller, G. R., Keer, L. M., Cheng, H. S., 1985, "On the mechanics of fatigue crack growth due to contact loading", Proceedings of the Royal Society of London, Vol. A397, pp. 197-209.
3. Harris, T. A., 1991, Rolling Bearing Analysis, 3rd Edn., Wiley, New York, Chap. 13.
4. Zhu, D. and Hu, Y.Z., 2001a, "A Computer Program Package for the Prediction of EHL and Mixed Lubrication Characteristics, Friction, Subsurface Stresses and Flash Temperatures Based on Measured 3-D Surface Roughness," Tribology Transactions, Vol. 44, pp.383-390.
5. Zhu, D., and Hu, Y.Z., 2001b, "Effects of Rough Surface Topography and Orientation on the Characteristics of EHD and Mixed Lubrication in Both Circular and Elliptical Contacts," Tribology Transactions, Vol. 44, pp.391-398.
6. Patir, N. and Cheng, H. S., 1978, "An Average Flow Model for Determining Effects of Three-Dimensional Roughness on Partial Hydrodynamic Lubrication", Journal of Lubrication Technology, Vol. 100, pp. 12-17.
7. Zhu, D., Cheng, H. S., Arai, T., and Hamai, K., 1992, "A Numerical Analysis for Piston Skirt in Mixed Lubrication - Part I: Basic Modeling", ASME Journal of Tribology, Vol. 114: 553-562.
8. Wang, Q., Shi, F. and Lee, S., 1998, "A Mixed-TEHD Model for Journal Bearing Conformal Contacts, Part II: Contact and Performance Analyses," ASME Journal of Tribology, Vol. 120, pp. 206-213.
9. Wang, Y., Wang, Q. and Chih, L., 2001, "Mixed Lubrication of Coupled Journal-thrust Bearing Systems," Submitted to Tribology International.
10. Jiang, X., Cheng, H.S., Hua, D.Y., 2000, "A theoretical analysis of mixed lubrication by macro micro approach: Part I - Results in a gear surface contact," Tribology Transactions, Vol. 43 (4), pp. 689-699.
11. Wang, Q., Zhu, D., Cheng, H. S., Yu, T., Jiang, X, and Liu, S., "Mixed Lubrication Analyses through A Macro-Micro Approach and A Full-Scale Micro EHL," submitted to Journal of Tribology.
12. Dyson, A., 1976, "A Failure of EHD Lubrication of Circumferentially Ground Disk," Proceedings of IME, Vol. 190, pp. 52-76.
13. Polonsky, I. A. and Keer, L. M., 2000, "A fast and accurate method for numerical analysis of layered elastic contacts," Journal of Tribology, Vol. 122, pp. 30-35.
14. Polonsky, I. A. and Keer, L. M., 1999, "A numerical method for solving rough contact problems based on the multi-level multi-summation and conjugate gradient techniques," Wear, Vol. 231, pp. 206-219.
15. Greenwood, J. A. and Williamson, J. B. P., 1966, "Contact of nominally flat surfaces", Proceedings of the Royal Society of London, Vol. A295, pp. 300–319.
16. Weibull, W., 1939, "The phenomenon of rupture in solids," Ingeniors Vetenskaps Akademien, No. 153.
17. Zaretsky, E. V., Poplawski, J., V., and Peters, S., M., 1996, "Comparison of life theories for rolling-element bearings," Tribology Transactions, Vol. 39 (2), pp. 237-248.
18. Ioannides, E. and Harris T. A., 1985, "A new fatigue life model for rolling bearings," Journal of Tribology, Vol. 107, pp. 367–378.
19. Harris, T. A. and McCool, J. I., 1996, "On the accuracy of rolling bearing fatigue life prediction," Journal of Tribology, Vol. 118, pp. 297–310.
20. Ai, X. L., 1998, "Effect of three-dimensional random surface roughness on contact fatigue", ASME Journal of Tribology, Vol. 120, pp. 159-164.

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Progression of the stick / slip zones in a dry wheel-rail contact: Updating theories on the basis of tribological reality

Eric NICCOLINI*, Yves BERTHIER**

* Laboratoire Essais Mesures - RATP
1bis, rue des sablons – 94470 Boissy-Saint-Léger – FRANCE
Tel : 33 1 48 18 90 69 – E-mail : eric.niccolini@ratp.fr

** Laboratoire de Mécanique des Contacts – UMR CNRS-INSA 5514
INSA Lyon – 20, avenue A. Einstein – 69621 Villeurbanne Cedex – FRANCE
Tel : 33 4 72 43 84 52 – E-mail : yves.berthier@insa-lyon.fr

At present, optimizing control of maximum wheel-rail adhesion, from the anti-skidding systems piloting the sliding and the rolling speeds of locomotives, is done on the basis of the wheel-rail contact models available in the literature.

Improved accuracy of power train control has highlighted the limits of this optimization technique since the contact conditions of the models, i.e. the stick / slip zones, do not take into account the physical phenomena involved, although they are hardly known.

To attenuate this difficulty, the recent works presented in this paper, allow describing the tribological reality of a wheel-rail contact, i.e. the progression of the stick / slip zones. This reality is achieved thanks to monitoring the changes in the the characteristics of the third body present between wheel and rail.

The results of tests under initial condition show that there is a correlation between the progressions of the third body flows and that of the adhesion for the various phases of the adhesion-sliding curve. Therefore activation of the internal flow of third body is in fact the physical manifestation of the theoretical condition of stick. On the other hand, the activation of the internal source and recirculation flows of the third body are the physical manifestation of the theoretical condition of slip. This tribological reality therefore allows us to update the theoretical stick / slip conditions and by consequence, anti-skidding systems.

1. INTRODUCTION

The latest developments of dry wheel-rail contact theories are used in the anti-skidding systems of locomotives to optimize the control of maximum wheel-rail adhesion. Unfortunately, these theories are limited by misunderstanding of the contact conditions.

Very numerous attempts of parameter setting have been made in both on site laboratory tests [1]. Unfortunately, the multiplicity of operating conditions (tracks, rolling stocks and environment), the different instruments used, and numerous data processing techniques [2] lead to the building multiple adhesion-sliding curves that cannot be customized [3].

Identification and monitoring of the progression of the phenomena activated at the contact scale as a function of rolling and sliding speeds - which can be easily controlled by the locomotive constructor - thus remains the recurring question.

However, the joint works carried out by the SNCF (French Railway network) and the Laboratoire de Mécanique des Contacts highlight the existence of the third body layer (Figure 1) constituted by a mixture heavily compacted by wheel and rail elements and which has absorbed the particles of contaminants [4].

In this paper, the authors show that it is now possible to improve the definitions of stick / slip conditions on the basis of this third body.

The authors first propose a review of existing theories, showing the limits of knowledge related to contact conditions.

They then present a method that allows correlating the progression of the rolling-sliding conditions at locomotive start-up with the different activated third body flows. This correlation is also validated by laboratory tests [4].

Lastly, a discussion proposes a new orientation of modeling the progression of adhesion as a function of sliding speed on the basis of the different third body flows activated.

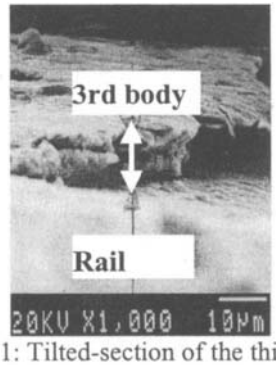


Figure 1: Tilted-section of the third body

2. REVIEW OF EXISTING THEORIES

A brief reminder of the foundations of the theories of wheel-rail contact in longitudinal sliding is explained to:

- define the adhesion-sliding curves used by the command laws at locomotive starting and braking,
- highlight the misunderstanding of the physical phenomena involved in a contact.

2.1 Foundations

Carter [5] developed the first theory of a one-dimensional elastic contact rolling with friction in longitudinal creep, very slow rolling speed (Va) and smooth surfaces. This author experimentally and theoretically proves the coexistence of stick / slip zones in the contact, the distribution of which varies as a function of longitudinal creep. The coexistence of zones is obtained by superimposing the tangential stresses (τ), established in the case of no-slip and complete slip respectively.

The integral of these stresses on the contact area represents the tangential force (T) transmitted by the contact. Carter uses a curve to reperesent the progression of this force as a function of longitudinal creep (ε) (Figure 2).

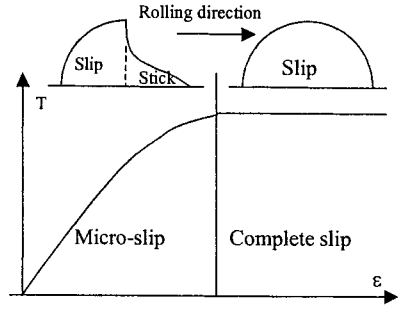


Figure 2 : Carter's curve [5]

This curve highlights the micro-slip and complete slip phase. These initial results allowed Haines and Ollerton [6] to deduct the distribution of the tangential stresses in an elliptical contact for the case of longitudinal creep. Nevertheless, the progression of Carter's curve remains valid, but in the knowledge that the distribution of the stick / slip zones is now elliptic (Figure 3).

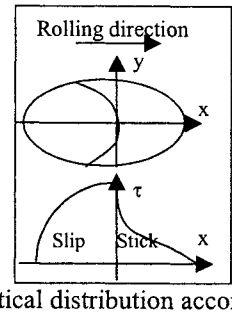


Figure 3: Elliptical distribution according to Haines-Ollerton [6]

Various theories have been developed on the basis of these works in order to process cases of lateral creep or spin, conformal contact and multi- or non-Hertzian contact [7].

In order to fit theoretical curves with those obtained on site, factors of influences observed on site or in the laboratory were introduced into the above theories. This re-adjustment was made according to several axes.

2.2 Axes of progressions

Various axes of progressions were studied whose main functions are friction, roughness and solid superficial layers.

2.2.1 Friction functions

Frederich established one of the first progressions [8]. The friction function which it defines results from on site tests. The friction is a function of the rolling speed and the normal force applied to a wheel. Unfortunately, the relations of macroscopic causes (mechanical parameters) in microscopic effects (variation of the friction) are not physically known. Therefore, Ohyama carried out laboratory experiments with layers injected into the contact where their shear stress were expressed as a function of the ratio of two coefficients: static friction μs and kinematic μd [9]. The first coefficient was defined in the stick zone while the second was in the slip zone (Figure 4). Thus the total tangential force was the integral of the tangential stresses $q_1(x)$ and $q_2(x)$ of each zone.

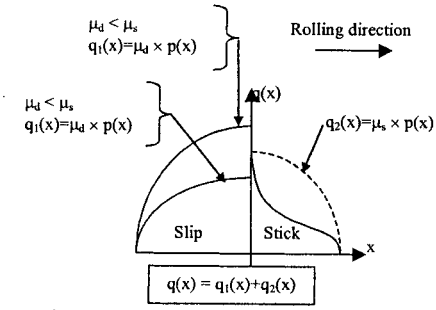


Figure 4 : Tangential stress distribution according to Ohyama [9]

However, the case $\mu s > \mu d$ is not a minimal solution for the tangential problem. To resolve this problem, Nielsen proposed a polynomial approach [10] allowing the introduction of Kronecker's peak on the edge of two zones.

What is more, thanks to this approach, Nielsen [10] defined a friction dependent local sliding speed of both bodies in contact, itself defined as a function of abscissa x in the one-dimensional contact (Figure 5). The problem that arises is resolved by defining the constitutive equation by a polynomial equation.

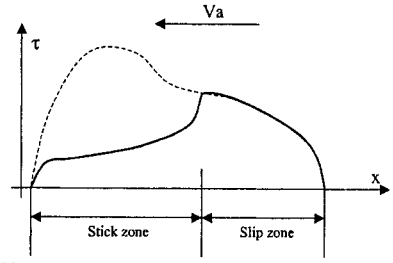
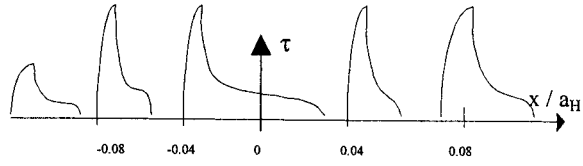


Figure 5 : Tangential stress distribution according to Nielsen [10]

2.2.2 Roughness

Bucher and al. [11] showed that the slope of the micro-slip depends on micro-roughness characteristics, by considering on the one hand the combined roughness of both surfaces and, on the other hand, several micro-contacts (Figure 6) in which Carter's stress distribution is applied. The phenomenon of plasticity engendered by the micro-roughness is assumed to be non-existent, contrary to the tests carried out in the laboratory by Krause [3].



a_H = Hertzian half-lenght

Figure 6 : Tangential stress distribution according to Bucher & al. [11]

The numerical study of a rough contact performed by Kalker and al. [12] demonstrated that the roughnesses do not change the slope of the micro-slip. These authors show that the evolution of the slope is due, rather, to the fluid interfacial layers.

2.2.3 Contaminants

Kalker [13] made a phenomenological mathematical description of a contact with three bodies based on the condition of an interfacial body that is weaker and more flexible than either of the first ones. The third body is modeled as an elastoplastic block of constant thickness, low stiffness, low yield limit, and strongly adhesive at the first body. The description of the third body is taken as a factor of empirical reduction of Kalker's coefficients Cij [7]. The actual creepage is then multiplied by a factor dependent on the flexibility of the two bodies (creepage and coefficients Cij) and the flexibility (thickness, elastic modulus, Poisson's ratio) of the third body.

On the contrary, Hou and al. [14] considered that the third body modifies the maximum of adhesion. These authors define a maximal coefficient (in micro-slip) in relation with the rheology of a layer of debris of constant thickness, the normal load and the relation between shear stress and sliding distance.

These progressions are only numeric permitting the empiric fitting of the the experimental adhesion-sliding curves with the theoretical ones. Thus the physical phenomena involved in the contact remains unknown and obviously not possible to model.

To attenuate this shortcoming, a method has been developed to highlight the tribological reality of the wheel-rail contact, and that of the rheological behavior of the third body evolving as a function of the rolling and sliding speeds and the contact environment. The ensuing result thus permits describing phenomenologically the progression of the stick / slip zones of a rolling - sliding contact in longitudinal direction.

3. METHOD

The followed method stems from the joint works of Goder [15], Berthier [16] and Niccolini [17]. It is based on a two scales of instrumentation :

- that of the locomotive, in order to establish the progression of adhesion as a function of the rolling and sliding speeds,
- that of the contact, in order to highlight the rheological characteristics of the third body as well as their progression.

3.1 Instrumentation of the locomotive

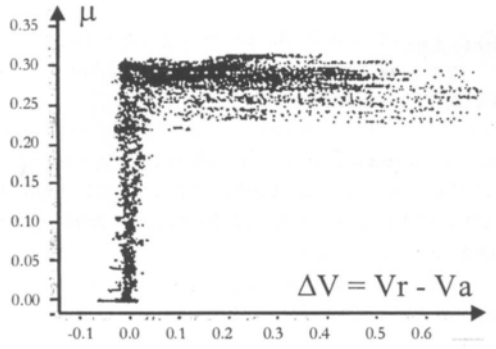
Adhesion (μ) is estimated on the first axle motorized by an engine of 84 tons locomotive by the following expression :

$$\mu = \frac{Cm/R}{N + \Delta N}$$

where :- Cm is the mechanical torque and estimated from strain gages fixed on the axle shaft,
- R is the average wheel rolling radius, equal to 625 mm,
- N is the static normal force, equal to a quarter of the weight of the locomotive, equal to 21 tons,
- ΔN represents the variations of normal force N due to mechanism effects and estimated on the basis of displacement of each primary suspension of the first axle.

The speed measurement of the wheel of a motorized and free running axle, estimates the linear speed of the wheel Vr. The translation speed measurement of the wheel of a non-motorized and free running axle, estimates the rolling speed of the locomotive Va. The difference of speed (Vr-Va) defining the sliding speed (ΔV) is calculated from these measurements. Furthermore, to better control the tests, both non-instrumented axles were motorized and braked to maintain rolling speed practically constant.

As a function of a given frequency of recording, the values taken by μ (from 0 to 0.35) and ΔV (from 0 to 0.8 m/s) allow describing the adhesion (μ) - sliding (ΔV) curve, with relatively constant rolling speed from 0.1 to 3 m/s (Figure 7)



3.2 Instrumentation of the contact

Instrumentation used for observation at the contact scale and more exactly at the third body scale consists of aluminum tape, polymeric replica, molding in resin, metallographic sections and a camera placed at the exit of the contact. The results obtained from these tribological tools placed on the rail and the wheel - detailed by Niccolini [4] - are used to identify the rheological characteristics of the third body that lead best to the creation and the progression of adhesion.

All the tools highlight the existence of two families of third body: those that are slightly adhesive, weakly adherent at the wheel or the rail, and those that are strongly adhesive, and very adherent. The definition of these families shows that the rolling action and sliding speeds and the contact environment create the material flow with different roles in the creation and progression of adhesion.

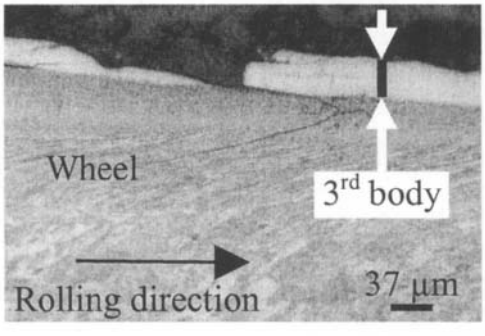


Figure 8: Metallographic section of a third body

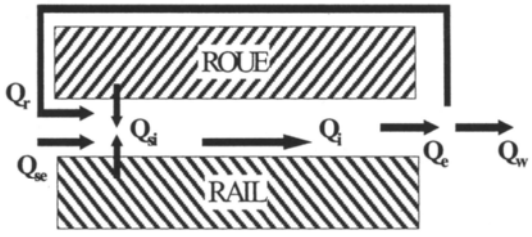
The geometrical (slenderness, area, surface density, cohesion, thickness), chemical, microstructural (texture, inclusion, crack) and micro-elastic (Young's micro-module, nano-hardness) particularities of each family of third body show very considerable shear, and even flow (Figure 8). These particularities observed on the surface or in section highlight the same consequences as that of the accommodation of the difference of speed between the wheel and the rail, and consequently the volume of material subject to adhesion.

Therefore, it is possible to correlate the characteristics of the third body observed in the surface with the results of the measurements made on the locomotive.

4. CORRELATION OF 3RD BODY / ADHESION / SLIDING

This correlation is based on the evaluation of the flow of third body in the contact described by the tribological circuit [18].

This circuit consists of six flows (Figure 9). Internal source flow Qsi corresponds to the detachment of particles that lead to the formation of natural third body in the contact. External source flow Qse corresponds to the introduction of artificial third bodies into the contact, for example, the debris created by environmental and operating conditions. Internal flow Qi is formed by third bodies trapped by the passage of the contact. External flow Qe is divided into a recirculation flow Qr and a wear flow Qw. Flow Qr is constituted by third bodies carried along the wheel and the rail and reintroduced into the contact. Flow Qw consists of third body, which is ejected and not reintroduced into the contact.



The activation of these flows, under rolling-sliding speeds, is directly correlated with :
- the creation of adhesion (linear phase of micro-slip),
- the progression of the adhesion (non-linear phase of micro-slip),
- the progression of maximal adhesion (phase of complete sliding).

4.1 Creation of adhesion

The internal flow (Qi) is highlighted by molding an angular portion of the tread of the wheel after a condition of pure rolling. The observations of the molding in resin reveal a set of localized flowing of strongly adhesive third body of the wheel oriented in rolling direction with a width from 10 to 50 µm (Figure 10).

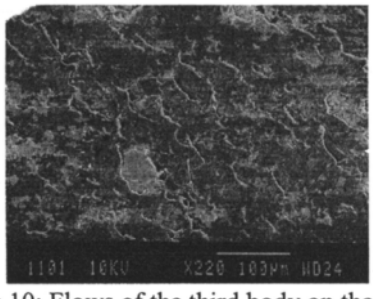


Figure 10: Flows of the third body on the wheel

4.2 Progression of adhesion

The internal source flow (Qsi) is highlighted by molding a sliding track formed on the rail after a rolling-sliding condition. The observations of the resin mold of this track reveal a set of third body flows in strips. These strips oriented in rolling direction have a thickness in the region of 10 µm, a length of about 1.5 mm and a width of about 100 µm. Each of these strips in turn highlight five strips with different features:
- strip B1 is characteristic of a slightly cohesive and slightly adhesive third body flow (Figure 11-a),
- strip B2 represents the flow of a very cohesive and slightly adhesive third body without flow direction (Figure 11-b).
- strip B3 is identified with a strongly adhesive third body flow, oriented in the rolling direction (Figure 11-b).
- strips B4 and B5 do not appear to have recognizable flows on the scale of observation. Contrary to strips B2 and B3, the authors presume that these strips B4 and B5 represents respectively the flow of slightly adhesive third body and the strongly adhesive flow of the wheel.

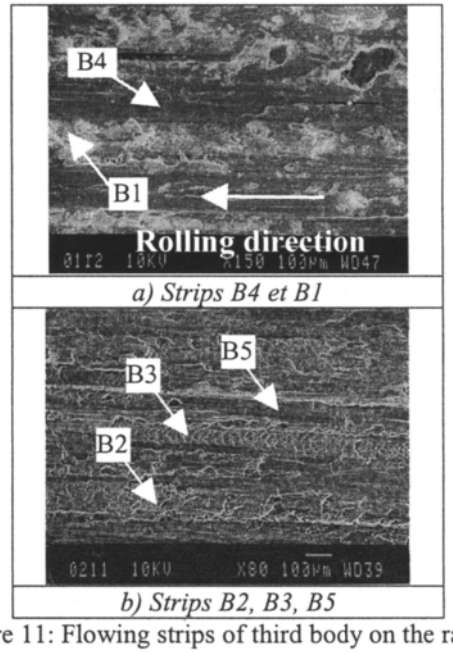


Figure 11: Flowing strips of third body on the rail

The presence of slightly adhesive third body shows the existence of an internal source flow (Qsi), produced by the flowing and detachment of strongly adhesive third body of the wheel and the rail (Cf. §4.1).

The considerable differences of these strips, due to the presence of slightly adhesive third body, show

that speed accommodation occurs at different depths. Moreover, the flow orientation of the third body in strips B3 (and B5), for which the depth is greater locally (> 10 µm) than for strips B2 (and B4), indicates substantial local sliding in the direction of the rolling speed. The tangential resistance proportional to the local adhesion is thus less when the strongly adhesive third body passes, compared with that produced by the flow of the slightly adhesive third body.

4.3 Progression of maximal adhesion

For different speed conditions, recirculation flow Qr and wear flow Qw are highlighted by analyzing the images taken by the camera placed at the exit of the contact as a function of the rolling direction. Revolution after revolution of wheel, recirculation flow Qr is highlighted by showing the third body passing on the tread of the wheel. When the fusion points on the latter are observed, a considerable flow Qr is activated. A wear flow Qw is identified due to the projections of third bodies outside the contact. This flow is ranges on average from 2 to 6 mm^3/s by considering the average elementary volume of an ejected third body equal to 0.6 mm^3 (2 mm^2 x 300 µm).

The activation of these flows depends on the speed conditions applied during the tests detailed by Niccolini [4]. The results are presented on a chart: rolling speed (Va) – sliding speed (ΔV). For this purpose, the mean values of these speeds for each test are represented in the chart by points (Figure 12).

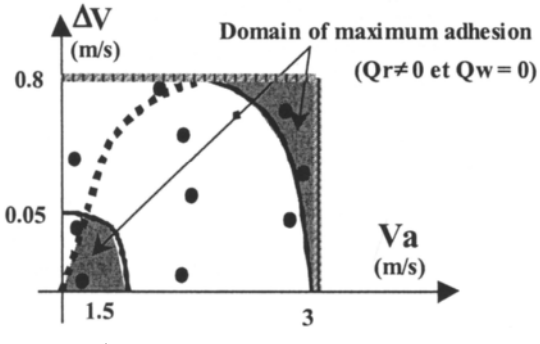


Figure 12 : External Flow (Qr, Qw)

For each test (a point on the graph) the lower (=0) and higher values (≠0) of each flow are represented on the chart. Consequently, it is possible to identify the weak (=0) and strong activation (≠0) domains for each flow.

The transposition of the adhesion values obtained for each test on the chart, shows that the domain of maximal adhesion (> 30 %) corresponds to the intersection of the domains of strong activation of flow Qr and the weak activation of flow Qw (Figure 11).

4.4 Synthesis of results

The previous results show that the third body flowing on the wheel and the rail contribute towards physically explaining the evolution of adhesion as a function of sliding speed.

These results can be summarized on the adhesion-sliding curve (Figure 13). The correlation between the activation of the third body flow and the various phases of the curve is shown below.

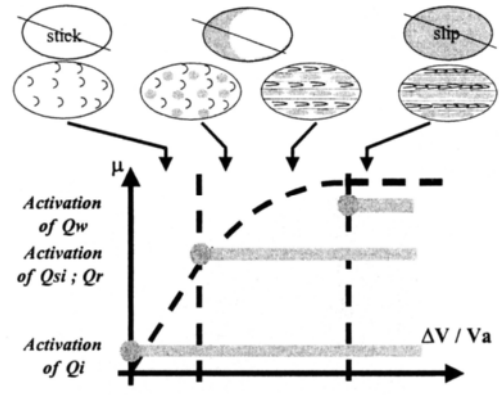


Figure 13: Correlation between third body flows and the adhesion-sliding curve

For the sliding speed of the linear part of the micro-slip phase, the flow of the strongly adhesive third strongly is activated – internal flow activated - according to §4.1. This activation physically represents the theoretical stick condition as defined by Kalker's models [7].

For the sliding speed of the non-linear part of the micro-slip phase, the slightly adhesive third body forms in the contact - internal source flow and

recirculation flow activated (Qsi and Qr) - according to §4.2 and §4.3. This creation of third body physically represents the initiation of the theoretical slip condition.

In the complete slip phase, the rolling speed will determine the level of adhesion by the activation or not of the wear flow (Qw) according to §4.3.

5. UPDATING THE LATEST THEORIES

Previous results contributed towards identifying the physical phenomena involved in the wheel-rail contact when starting. Up to now, some of the phenomena had above all been used as boundaries conditions and setting parameters for the different models used for anti-skid systems. The behaviours of the 3rd body have been highlighted for the following specific conditions:

- the contact is dry and there is no internal source flow of material (particles of sand or other artificial particles),
- rolling speed should remain lower than 3m/s (case of start-up),
- sliding speed is exclusively longitudinal and low enough to ignore the effects of temperature (< 0,8m/s),
- all the mechanism effects acting on normal load are ignored.

However, these specific conditions have permitted a) observation of the wheel surfaces that have simple tribological histories since they are limited to fractions of revolution, thus facilitating interpretations, b) visualization of 3rd body flows ejected of the contact and c) reconstitution of the flows activated in the contact as well as their changes (Figure 13).

Therefore, the following should be borne in mind for the test conditions:

- existence of two types of local ellipses and local strips. The former is defined for strongly adhesive third bodies for which a static friction coefficient exists. The latter is defined for slightly adhesive third bodies for which a kinematic friction coefficient exists.
- Kalker's theory of stress distribution is valid in local ellipses and local strips.
- the dimensions of the local ellipses and strips share the same width since the ellipses are formed in the same place as the strips as a function of sliding.

Concerning the changes:

- two types of local ellipses will be defined until reaching a transition sliding value to be defined, the former (flows of the strongly adhesive 3rd body) being activated in the linear part of the micro-slip while the latter (flows of the slightly adhesive 3rd body) being activated in the non-linear part of the micro-slip;
- two types of local strips will be defined from transition sliding value, the former representing the flow of the strongly adhesive 3rd body and the former that of the slightly adhesive 3rd body.

Although this works proposes the qualitative correlation between the morphology and rheological behavior of the 3rd body and the friction value, it is obvious that writing the equations of flows and measuring the rheology of the solid 3rd bodies remains a complicated task. This task is all the more complicated since instrumenting the wheel-rail contact in-vivo is extremely difficult. Thus it is necessary to continue getting round this difficulty by using the above results to refine existing models, so that they can be used as tools for analyzing experimental tests. Furthermore, it is necessary to continue characterizing the 3rd body in the corresponding range of sliding and rolling stresses, which is somewhat removed, for example from skidding and nano-indentation tests.

In industrial terms it has become relatively easy to use image analysis, for example, to collect data on the morphology of the 3rd body that could be incorporated in traction motor control algorithms so as to control locomotive forward speed and skidding, thereby obtaining maximal adherence.

REFERENCES

1. Boiteux M. – *Locomotive Adhesion under Aspects of their Builder and their Operation : Recent State of Wheel-Rail Adhesion Knowledge.* Utrecht : ERRI, 1978, 67p. Rapport No. B44/RP14.
Hill R.J., De La Vaissière J.F. – *A Fuzzy*

2. *Wheel-Rail Adhesion Model for Rail Traction.* Trondheim : EPE, 1997, pp. 416-421.
3. Krause H., Poll G. – *The Influence of Real Materials and System Properties on the Traction/Creep Relationships in Rolling Contact.* Proceedings of 1st International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Kalousek et al. Editors, UW press, 1983, pp. 353-371.
4. Niccolini E. – *Adhérence roue-rail en traction – Réalité tribologique en condition de démarrage.* Thèse de doctorat : Institut National des Sciences Appliquées de Lyon, 2001, 244p. (Jury : Kalker J.J., Kalousek J., Dubourg M.-C., Berthier Y., Jean M., Gras R., Barquins M., Alacoque J.-C., Girardi L.)
5. Carter F. W. – *On the Action of a Locomotive Driving Wheel.* Proceedings of the Royal Society of London, Serie A, 1926, pp. 151-157.
6. Haines D.J., Ollerton E. – *Contact Stress Distributions on Elliptical Contact Surfaces Subjected to Radial and Tangential Forces.* Proceedings of the Institution of Mechanical Engineers, 1963, Vol. 177, pp. 95-114.
7. Kalker J.J. – *Three- dimensional Elastic Bodies in Rolling Contact.* Dordrecht, Boston : Kluwer Academic Publishers, 1990, 314 p.
8. Boiteux M. – *Adhesion in braking and anti-skidding systems : Fundamentals laws of adhesion in braking.* Utrecht : ERRI, 1990, 77p. Rapport No. B164/RP2.
9. Ohyama T. – *Some Basic Studies on the Influence of Surface Contaminations on Adhesion Force Between Wheel and Rail at Higher Speeds.* Quarterly Report of Railway Technical Research Institute, 1989, Vol. 30, No. 3, pp. 127-135.
10. Nielsen J.B. – *New Developments in the Theory of Wheel-Rail Contact Mechanics.* Thèse de doctorat : Université Technique du Danemark, 1998, 211p.
11. Bucher F., Knothe K., Theiler A. – *Normal and Tangential Problem of Surfaces with Measured Roughness.* Proceedings of the 5th International Conference on Contact Mechancis and Wear of Rail/Wheel Systems, Tokyo, July 25th-28th, 2000, pp. 96-103.
12. Kalker J.J., Dekking F.M., Vollebregt E.A.H. – *Simulation of Rough Elastic Contacts.* Transaction of ASME, Journal of Applied Mechanics. 1997, Vol. 64, No. 2, pp. 361-368.
13. Kalker J.J. – *Modification of the Two-Body Contact Conditions to Account for the Third Body.* Proceedings of 19th Leeds-Lyon Symposium on Tribology , Leeds, September 8th-11th, 1992, pp. 183-189.
14. Hou K., Kalousek J., Magel E. – *Rheological Model of Solid Layer in Rolling Contact.* Wear, 1997, Vol. 211, pp. 134-140.
15. Godet M. – *Third body approach : a Mechanical View of Wear.* Wear, 1984, Vol. 149, pp. 325-340.
16. Berthier Y. – *Mécanismes et Tribologie.* Thèse ès Science : Institut National des Sciences Appliquées de Lyon, 1988, 71p.
17. Niccolini E., Berthier Y. – *Wheel-Rail Adhesion : What is the tribological reality ?.* Proceedings of the 5th International Conference on Contact Mechancis and Wear of Rail/Wheel Systems, Tokyo, July 25th-28th, 2000, pp. 219-225.
18 Berthier Y., *Background on Friction and Wear*, in Lemaitre Handbook of Materials Behavior Models. ISBN 0-12-443341-3, Academic Press, 2001, pp. 676 – 699.

This Page Intentionally Left Blank

Tribological Research and Design for Engineering Systems
D. Dowson et al. (Editors)
© 2003 Elsevier B.V. All rights reserved.

Tribology of Electron beam excited plasma nitrided tool steels with narrow gap features

Petros Abraha [a], Hiroaki Shoyama [b], Kazunari Taniguchi [c] and Tamio Hara [b]

[a] Meijo University, 1-501 Shiogamaguchi, Tempaku-ku, Nagoya 468-8502, Japan

[b] Toyota Technological Institute, 2-12-1 Hisakata, Tempaku-ku, Nagoya 468-8511, Japan

[c] Kyoto University of Education, 1 Fukakusa, Fujinomori-cho, Fushimi-ku, Kyoto 612-8522, Japan

The wear characteristics of narrow gap geometry nitrided tool steels formed by electron beam excited plasma (EBEP) is systematically investigated. Upon nitriding the samples for 1.5 hours, the wear and hardness of the treated and untreated regions of the tool steel specimen were measured. The measurements of the Vickers microhardness tests show the treated region to be up to 1.3 to 2 times as hard as the untreated region. The wear volume of the untreated region was measured to be 2 to 6 times more than the treated regions. These results indicate that the EBEP device is a suitable nitriding device for tool steels with narrow gap features.

1. INTRODUCTION

Nitriding has attracted considerable attention in material processes for producing films on surfaces and modifying their mechanical properties such as wear and hardness [1]. The primary plasma sources for the production of atomic nitrogen has been the conventional plasma sources, including electron cyclotron resonance (ECR) plasma, radio frequency (RF) plasma, dc glow discharge, cold cathode gun [2-4]. These plasma sources however, use small chamber volume and need a high pressure and discharge voltage[6] to produce enough flux of atomic nitrogen for plasma processing. On the other hand, atomic nitrogen fluxes from large volume chambers, with a pressure of under 2 mTorr, were reported to be below the limit of the experimental detection [5]. Thus, in large volume chambers, one can only have few electrons with energy that is enough to break the molecular bond of nitrogen gas. The cross sections for the dissociation of molecular gases depend on the impact electron energy. The maximum cross section for nitrogen dissociation is about 60-130 eV [6-7]. This means that a high dissociation rate of nitrogen can be achieved if the electron energy were around the maximum cross section. The electron-beam-excited plasma (EBEP) device used in this experiment was designed to produce electron beam energy of 50-150 eV [8].

2. EXPERIMENTAL APPARATUS

A schematic diagram of the (EBEP) device that is used for nitriding tool steels is shown in Figure 1. The system consists of three regions: the discharge plasma region (I), the electron acceleration region (II), and the plasma production region (III). Argon gas is introduced to region I, and the dc discharge is developed between the cathode (K) and the anodes (S_1, S_2). The cathode was a lanthanum hexaboride (LaB_6) plate heated by a tungsten filament. The discharge region is the source of the electrons for the EBEP system.

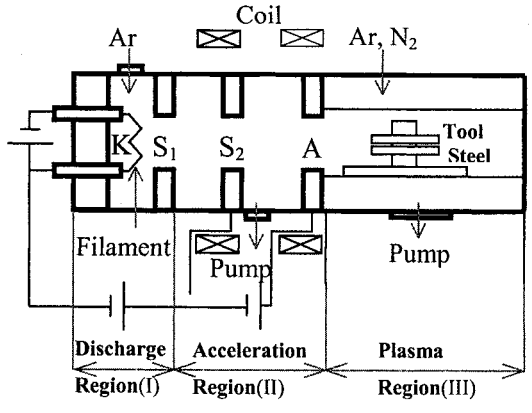


Figure 1. A schematic diagram of the electron-beam-excited plasma (EBEP) device

Upon applying the electron accelerating voltage V_a to the electrode (A), the electrons in region (I) are accelerated toward region (II) along the magnetic field (about 0.03 T on the central axis) and form an electron beam with an energy corresponding to V_a. The electron beam in region (III) collides with the process gases introduced through the mass flow controller to generate plasma within region (III). The carriergas used in this experiment is argon while the process gas is nitrogen.

Two pairs of turbo molecular pumps and rotary pumps were installed. The pressures of regions (I), (II) and (III) are 10^{-2} Torr, 10^{-4} Torr and 10^{-3} Torr, respectively. Typical plasma parameters are as follows; the electron density $n_e = 5 \cdot 10^{11}$ cm^{-3}, the electron temperature $T_e = 2$ eV.

The test specimen is placed inside glass pipe within region (III). The glass pipe prevents diffusion of the electron beam. This results in the formation of high-density plasma within the glass pipe where the experimental work piece is placed.

The beam current I_b can be controlled by the discharge current I_d. Figure 2 shows the relationship between the electron beam current and the discharge current at chamber pressures of 2.2 mTorr, 1.5 mTorr and 0.7 mTorr. The beam current increases linearly with the increase of the discharge current.

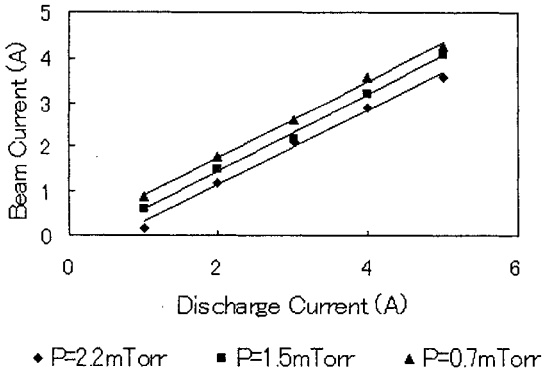


Figure 2. Relationship between electron beam current and discharge current

The applied acceleration voltage determines the total energy of the beam current. Figure 3 shows the relationship between the electron beam current and the acceleration voltage at discharge currents of 2 amperes and 3 amperes. The acceleration voltage was varied between 60 volts and 200 volts. The beam current increases with an increase of the acceleration voltage and remains constant after 60 volts. This means that one can increase the acceleration voltage while keeping the beam current constant. The acceleration voltage and beam current can thus be controlled independently, a unique feature of the EBEP device. In this experiment, nitrogen gas with a flow rate of 10 sccm and argon gas with a flow rate of 5 sccm were used. The treatment time was kept at 1.5 hours. The temperature within the chamber is relatively low, about 300-350 °C.

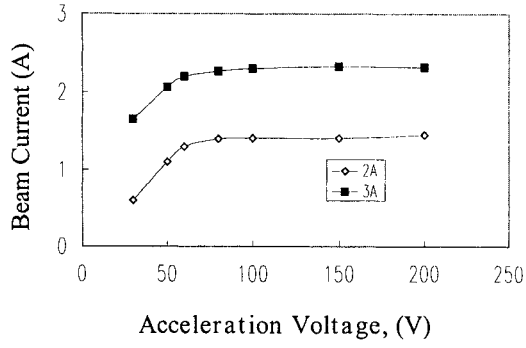


Figure 3. Relationship between the electron beam current and the acceleration voltage

3. EXPERIMENTAL METHOD

Figure 4 shows the schematics of the test specimen. Spacer plates in between two stainless disks were used to create narrow gaps. The thickness of the plates was 2 mm and the diameter was 20 mm. The gaps were varied by the thickness of the spacer plates. The diameter of the spacer was 10 mm. The two plates and the spacer are all fixed together by M5 bolt and nut.

In this experiment, wear and hardness tests were performed on the disc surface facing the spacer. The hardness was measured using Vickers microhardness tester. The indenter is a square shaped diamond pyramid with a surface angle of 136°. The applied load was 0.9807 N for 15 seconds. Two distinct regions, the treated and untreated regions, were

observed. The wear tests were evaluated from wear track tests that run on both the treated and untreated regions in circular paths.

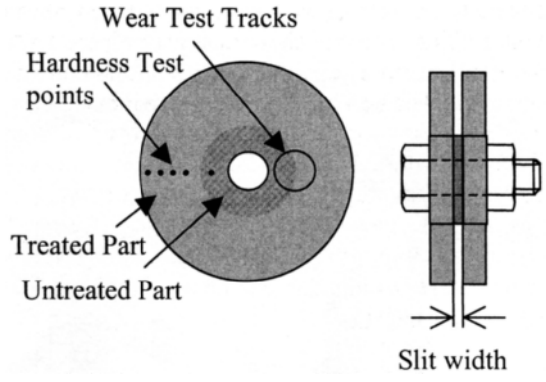


Figure 4 Schematics of the test specimen

The specimen were ground with 4800# SiC abrasive paper and polished with alumina paste to an average surface finish value of Ra=0.02-0.03 μm. The disk material was SKD 61 (5-Cr-1Mo-1V) stainless steel.

A pin on disk tribometer was used to evaluate the wear of the treated and untreated regions. The disks were set on the turntable driven at 12 mm/second for 10 minutes. A constant normal load of 50 grams was applied by JIS G3539SWCH10R pin material.

4. WEAR AND HARDNESS RESULTS

The nitrided case depth was measured for a sample with a slit width of 1.2 mm at a distance of 1mm from the outside edge of the specimen.

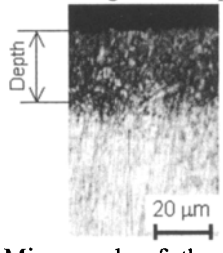


Figure 5 Micrograph of the cross section of the treated layer

The micrograph of the cross section of the treated layer is shown in Figure 5. The case depth was measured to be about 0.025 mm.

Figure 6 shows the X-ray diffraction spectra for the treated and untreated regions of the sample tool steel. The spectrum for the treated layer shows peaks of CrN that shows inclusion of nitrogen, whereas the untreated layer does not show any traces of nitrogen.

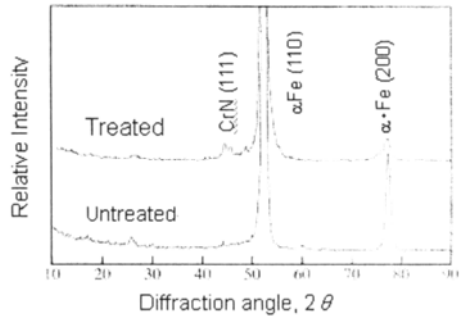


Figure 6 X-ray diffraction spectra for the treated and untreated regions of the test specimen

The measured wear depth for slit width of 0.4 mm, 0.8 mm and 1.2 mm are shown in Figure 7. The measurements for the wear depth were taken at a distance of 1 mm from the outside edge of the specimen surface. In comparing the wear depth, the treated region with a slit width of 0.4 mm shows only half the wear depth of the untreated specimen. The wear depth decreases further to about one fourth and one fifth at slit width of 0.8 mm and 1.2 mm, respectively.

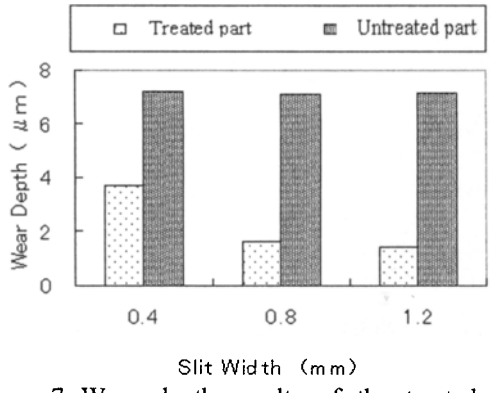


Figure 7 Wear depth results of the treated and untreated specimen

The measured wear track width for slit width of 0.4 mm, 0.8 mm and 1.2 mm are shown in Figure 8.

The measurements for the wear width were taken at a distance of 1 mm from the outside edge of the specimen. In comparing the wear width, the treated region with a slit width of 0.4 mm shows about four fifth the wear width of the untreated specimen. The wear decreases further only to a very small degree for slit width of both 0.8 mm and 1.2 mm.

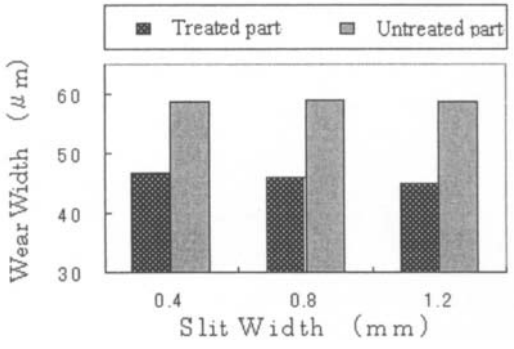


Figure 8 Wear width results of the treated and untreated specimen

The wear volumes calculated from the measurement of the wear width and wear depth of the wear tracks are shown in Figure 9. In comparing the wear volume, the treated region with a slit width of 0.4 mm shows about half the wear volume of the untreated specimen. The wear volume decreases further to about one fifth and one sixth for slit width of 0.8 mm and 1.2 mm, respectively. This demonstrates that the electron beam excited plasma device is suitable equipment for increasing the wear resistance of stainless steel materials with narrow gap features.

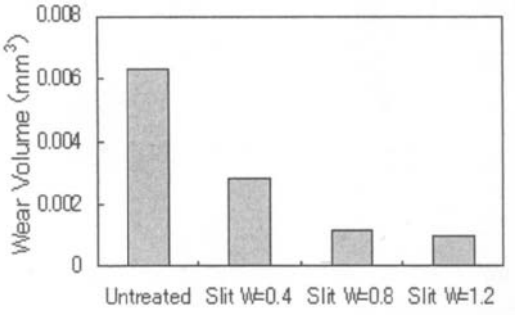


Figure 9 Wear volume of the treated (slit width = 0.4 mm, 0.8 mm, 1.2 mm) and untreated specimen

The Vickers microhardness test results of the treated and untreated specimen is shown in Figure 10. The hardness tests were measured at distances of 1 mm, 2 mm, 3 mm, 4 mm and 6 mm from the outside edge of the specimen disks. Here, a distance of 6 mm from the outside edge lies on the untreated region and the rest lie on the treated region.

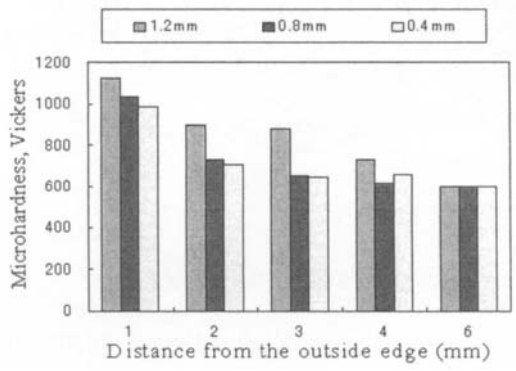


Figure 10 Vickers microhardness results of the treated and untreated specimen

The hardness values decrease with the distance from the outside edge. At a distance of 1 mm from the outside edge, the hardness values of a slit with a width of 1.2 mm and 0.4 mm are 2 times and 1.3 times that of the untreated region, respectively. This is a remarkable increase in hardness. Although the hardness values for slit width of 0.8 mm and 0.4 are relatively lower, nitriding at such narrow gap features has been successful. For example, the hardness value of a slit width of 0.4 mm at a distance of 1 mm from the edge was increased by about 1.6 times. On the other hand, at a distance of 4 mm from the outside edge, the hardness of the specimen with a slit width of 0.4 mm does not show a significant increase. The hardness values shown here are for only 1.5 hours of treatment time. Therefore, one can easily increase the hardness values by just increasing the treatment time of the specimen in accordance to the requirements.

The device is run at a relatively low temperature of 400°C, therefore, the structure of the treated specimen material is maintained with out any change. Moreover, the pre-treatment and surface finish of the specimen was maintained to the original one with no traces of compound layer on the surface.

5. CONCLUSION

The measurements of the Vickers hardness tests show the treated region to be 1.3 to 2 times as hard as the untreated region. On the other hand, the wear volume of the untreated region was measured to be 2 to 4 times more than the treated regions. The above experimental results demonstrate that the electron beam excited plasma device is suitable equipment for increasing the hardness and decreasing the wear volume of dies and tools that have intricate shapes with narrow gaps between the features.

ACKNOWLEDGEMENTS

This research work was partially supported by the Nitto Foundation.

REFERENCES

1. M. Berg, Surf. Coat. Technol. **124** (2000) 25.
2. S.C. Mishra, B. B. Nayak, B. C. Mohanty, Surf. Coat. Technol. **145** (2001) 24.
3. J. M. Priest, M. J. Baldwin, M. P. Fewell, Surf. Coat. Technol. **145** (2001) 152.
4. H. Michel, T. Czerwiec, M. Gantois, D. Ablitzer, A. Richard, Surf. Coat. Technol. **72** (1995) 103
5. Y. Wang, R. J. Van Brunt and J. K. Olthoff, J. Appl. Phys. **83** (1998) 703.
6. M. Hamagaki and T. Hara, Jpn. J. Appl. Phys. **33** (1994) L383.
7. K.Taniguchi, M. Sugimoto, S. Masuko, T. Kobayashi, M. Hamagaki, P. Abraha and T. Hara, Jpn. J. Appl. Phys. **39** (2000) L999.
8. S. Tada, S. Takashima, M. Ito, M. Hori, and T. Goto, Y. Sakamoto, J. Appl. Phys. **88** (2000) 1756.

This Page Intentionally Left Blank

WRITTEN DISCUSSION

This Page Intentionally Left Blank

Written Discussion-Contributions

DISCUSSION

SESSION I – KEYNOTE ADDRESS

Paper I (i) Keynote Address "Tribology in Design" by M J Neale (Neale Consulting Engineers,UK).

Dr H Van Leeuwen (Eindhoven University of Technology). Developments in the practice of mechanical engineering deserve a moment of reflection. I wonder why the last edition of the Tribology Handbook shows different boundaries for operational limits of seals, but not on bearings, when compared to the former edition? Did not the bearings have any development since the early 1970's?

Reply by the Author The data in the Tribology Handbook does not show the extreme limits of component performance, but is intended to provide safe overall guidance for designers. Bearings have advanced since the early 1970s but not enough to justify any major changes in their overall guidance limits. Mechanical seals on the other hand have experienced new technology injections under pressure from major demands for improved performance with less leakage. This has resulted in the introduction of new technologies such as gas film lubricated face seals, giving substantial improvements in their performance at high rotational speeds. It is probably because of its overall guidance role, that twenty two years elapsed between the first and second editions of the book.

Professor D Dowson (Institute of Trbology, The University of Leeds, UK). Your paper nicely places developments in triboloy in the spectrum of progress in machine design, particularly in relation to design and reliability. You have also illustrated most effectively the way in which studies of the performance of machinery in the field can illuminate our understanding of fundamental aspects of tribology. You have also uttered a plea for more reference works relating the sensitivity of machinery to the reduced maintenance of critical tribological components which has been evident in recent years, sometimes with disastrous consequences.

Such developments need to be supported by a much wider and deeper understanding of the basic features of tribology. by all engineers, yet there seems to be less and less time in courses and training programmes for our subject. It is easy to forget the significant impact of generic subjects, like tribology, upon the whole field of machinery design, manufacture, and operation. My concern is that there seems to be a decline in the attention given to basic features of machine operation in many programmes of education and training for mechanical engineers around the world. It appears that in the U.K. we have to be reminded of these things every quarter of a century or so! For example, the stirring developments in understanding of basic features of lubrication and the development of rational bearing design in the first two decades of the 20th Century were followed in 1920 by a remarkable Department of Scientific and Industrial Research (DSIR) inquiry which focussed attention upon the significance of *"Lubricants and Lubrication"*. In October 1937, the Institution of Mechanical Engineers organized a three-day General Discussion on *"Lubricants and Lubrication"* attended by no less than 600 delegates. The Proceedings of this General Discussion now represent a collection of classical papers on many aspects of our subject. Then came the work of the Lubrication Working Party (1964-1966) established at the behest of the Rt. Hon. the Lord Bowden, Minister of State for Education and Science, to consider the ... *"present position of lubrication education and research in this country"*. The Report by this Working Party, headed by. H.Peter Jost, was entitled..."*Lubrication (Tribology) Education and Research-A Report on the Present Position and Industry's Needs"*. It was published in 1966.

I am not sure whether other nations experience these historical fluctuations in the fortunes of *tribology* that appear to afflict the U.K. I would be

interested to hear from our speaker and distinguished designer, whether he recognizes this effect and, if so, how he would promote an effective response to the present perceived tribology trough! Perhaps we need to promote a further national inquiry into the subject (Son of *'Jost'*?).

Reply by the Author Tribology does seem to be in a bit of a trough at the moment, and I think that this is probably because it is essentially an engineering subject, which is concerned with what happens at the moving surfaces of actual machines. The problem is a symptom of the lack of public regard for engineering, which is seen as something to do with dirty hands, rather than the basic reason why we no longer live in caves. Engineers and tribologists need to stand up for the performance and validity of their work as a major input to the knowledge and living standards of the world. They need to do this continuously, rather than at about twenty year intervals after downwards drift into a new trough.

We need to be emphasising the value of knowledge, as distinct from the listing of large numbers of basic scientific investigations in the research literature. By knowledge, I mean the understanding of how things work and how to make them better. It is interesting that the product knowledge of say a motor manufacturer and a fertiliser manufacturer is a major factor in determining the value of each company and the differences between them, but it never appears on the balance sheet.

My own experience is that I have had some success in solving problems with machinery, all over the world, and in recommending improvements to its design. I am sure that an important contribution to this was that I was a shop floor engineering apprentice from the age of 16. The key part of this was the opportunity it provided to become close to machinery and to identify from the feel and appearance of its component parts, as to how it had been operating, together with any problems which it had. This provided an intimate knowledge of machinery. I also studied part time for a university degree, and this provided some useful fundamental flesh to put on the bones of the basic knowledge structure. The priorities were not the other way round.

My advice to tribologists who wish to advance their subject, and not allow it to fall into troughs, is that they should do more work on understanding the key factors in the design and operation of all kinds of machines and systems. They should then widely publicise their success as a useful contribution to the public good, so that the value of their work is more widely appreciated. There is a great opportunity here for a university to take a lead in the study of multivariable practical problem areas, and to resist any pressure from their pure science colleagues to only do laboratory tests with single controlled variables, which make only a very narrow contribution, and then only to the scientific literature. It is interesting that multi-variable research in the field is acceptable in the social sciences, presumably because controlled experiments on people are not ethically acceptable.

Another area in which tribologists could make their mark is in optimising the application of computerised methods to the design of machines. While these are totally valid and extremely useful for analysing and improving the detail design of machines, they are not really appropriate for producing the design as such. This requires a major input from past operating experience, to incorporate a large number of complex interacting variables. Problems are beginning to arise in the field with products designed on computers by confident highly qualified young design teams, who do not have a wide range of background experience with the particular product. If the problems arise with the tribological components there can be a substantial reduction in reliability.

In summary, tribologists need to maintain a continuing public respect for their subject by showing how it can advance the essential knowledge of machinery and its operation, and enable reliable and successful products to be created for the benefit of the public, and the prosperity of industry and the country.

SESSION II – CHEMICAL ACTIONS in TRIBOLOGY

Paper II (i) Invited Paper – "Tribochemistry: Friction-induced Lamellar Solids from Lubricant Additives" by Dr R Ladavière, Professor Jean-Michel Martin, Dr T le Mogne,B Vacher, Mr B Constans and Mr S Iovine (Ecole Centrale de Lyon, France).

Professor R C Coy (Institute of Trbology, The University of Leeds, UK) Have you checked that MoS_2, BN and Graphite added to a lubricant give the same effect as these materials formed through tribochemistry?

Reply by the Author We tried for MoS_2 and BN as a powder dispersed in a base oil. Results are generally not as good as tribo-formed species. We believe that the reason is that added powder is partially oxidized due to contact with the environment prior to friction. In the triboformed material, oxidation has practically no time to take place in the contact area (partly depleted in air).

Dr J A Williams (Cambridge University, UK) You commented that the effect of h-BN was principally on the coefficient of static friction rather than the coefficient of dynamic or kinematic friction. Is it possible to suggest why this might be so?

Reply by the Author We believe that the smoothing effect of h-BN on friction is due to its specific mechanical properties. However we have no clear explanation at the moment.

Dr A Oldfield (Uniquema Lubricants, UK) Please comment on the reaction kinetics for forming crystalline tribofilms versus amorphous films. Consider the reaction between separate MoDTC molecules to form large scale crystalline MoS_2 versus amorphous ZDDP glassy films. Consider orientation effects of reacting moieties, time for reaction pressure, temperature, and mass flow effects in the high viscosity in the contacts. Also consider the low concentration of the reacting moieties.

Reply by the Author So far, the kinetics of formation of tribochemical products has not been investigated in detail. We think that MoS_2 is formed after the amorphous phosphate film and that the phosphate prevents MoS_2 from oxidation in the tribofilm material. The next investigation in our lab will be a study of the genesis of such films by using friction tests and a combinatorial approach (see Nic Spencer's approach in Zurich for example).

Professor S Chandrasekar (Purdue University, Indiana, U.S.A.) Have you seen any evidence of formation of cubic boron nitride by tribochemical reaction?

Reply by the Author No, because XANES signature of c-BN is completely different from that of h-BN that we detected. The absence of $1s/\pi^*$ transition in XANES of c-BN is one of the differences.

Professor H Spikes (Imperial College, London, UK) Do you have any actual evidence that graphite formed from detergents influences the physical properties of the overbased material?

Perhaps you should carry out similar work using a non-aromatic dispersant.

Reply by the Author Turbostratic graphitic carbon is thought to form from the aromatic parts of the detergent molecules. Please note that this can be different from pure graphite. When using non-aromatic amphiphilic molecules, graphitic carbon does not form and large calcite crystals grow in the tribofilm material. The role of this graphitic carbon could be to prevent further crystallisation of colloidal calcite nanoparticles and to help plastic flow of tribofilm material like in superplastic behaviour.

Paper II (ii) "Assessing Friction, Wear and Film Formation Characteristics in Formulated Lubricants in Severe to Moderate Boundary Lubrication Conditions" by Mr A Morina[1], Professor A Neville[1], Mr J H Green[2], and Dr M Priest[2] (1-Heriot-Watt University, Edingurgh, UK; 2-Institute of Trbology, The University of Leeds, UK)

Dr P A Dearnley (Institute of Tribology, The University of Leeds, UK). How thick are the tribological films and how important is film thickness on friction behaviour?

Reply by the Authors The thickness of the film has been the topic of much debate amongst researchers. Because of its layered structure it is often debatable where the film itself starts and ends. However, it is generally agreed that on a steel/steel couple in an oil containing anti-wear additives, detergents, friction modifiers etc. it is in the order of a few hundred nanometres. Importantly the nature of the film and its thickness determines the friction response. It is known that films formed with ZDDP generally increase friction in contrast to Mo-containing additives.

Paper II (iii) "Approaches to Sensitising Engine Valve Train Models to Lubricant Formulation Characteristics" by Mr J H Green[1] and Dr M Priest[1], Mr A Morina[2] and Professor A Neville[2] (1- Institute of Tribology, The University of Leeds, UK; 2 - Heriot-Watt University, Edingurgh, UK)

Dr J A Williams (Cambridge University, UK) In your block and ring test, which simulates the cam and which the follower? How important is the fact that the cam and follower is a dynamic situation?

Reply by the Authors The stationary block represents the follower and the rotating ring the cam. As Dr Williams rightly points out this is a simplification of the real situation where operating conditions change continuously in terms of radius of curvature, load, speed, temperature and heat transfer into the surfaces. Most notably in relation to surface temperature and chemistry, the stationary contact point on the block is unrepresentative of the follower where the contact point reciprocates across the surface.

The block on ring tests replicate the contact conditions at particular points in the cam / follower cycle. As such they are the first stage of an experimental programme to investigate the correlations between surface chemistry and tribology in such interfaces. An important second phase will progress to more realistic simulations.

Paper II (iv) "Electrochemical Control of Wear" by S Mischler (EPFL-STI-IMX-LMCH, Lausanne, Switzerland)

Professor N Spencer, (ETH-Zurich, Switzerland). How are we to think about the potential on the surface of the wear particles? Does it not depend on the effectiveness of the electrical contact to the substrate?

Reply by the Author The effective potential of wear particles depends on the electrical contact to the substrate as well as on ohmic potential drops caused by the finite distance between metal plate and reference electrode.

Usually, passive films covering iron are reasonably good electrical conductors allowing, for example, the electrochemical oxidation of water to take place at very high rates on passive iron electrodes. So metal particles pressed against the metal surface can be considered to be in electrical contact and thus to experience the same potential as the substrate. In case of thicker third bodies the electrical contact could be less efficient leading possibly to electrically insulated particles. However, in this case the particle would be protected against corrosion by the surrounding particles and in any case its oxidation would not contribute to the overall electrochemical current (chemical corrosion).

Probably more relevant, at least in the systems we encountered so far, are the ohmic drop effects. They were already discussed by Stemp et al in a previous Leeds Lyon meeting (27) for the case of tribo-corrosion experiments, where they are exacerbated by the shielding effect of the counter piece. In the case of electrolytes of relatively low conductivity, large counter pieces, small displacements and high currents the ohmic drop effect may become large so that the effective electrochemical potential in the contact zone may significantly differ from the overall applied potential. For example, a situation of active dissolution could be established on the contacting surface of a steel pin held at an imposed passive potential and rubbing against an alumina plate simply by reducing the solution conductivity.

Certainly all these effects must be considered in the interpretation of tribo-electro-chemical experiments. Suitable model systems should be selected for mechanistic interpretation in order to

minimise deviation of the actual potential from the applied potential.

SESSION III ELASTOHYDRODYNAMIC LUBRICATION (1)- Dynamic Effects

Paper (III) (i) Invited Paper "Dynamic Effects in EHL Contacts" by Dr C J Hooke, (University of Birmingham, UK)

Dr J B Medley (University of Waterloo, Canada). It seems to me that local squeeze effects would damp out some of the rapid local changes in surface shape, and clearance. Perhaps I am being misled by differences in horizontal and vertical scales that make the changes look more dramatic, but is local squeeze velocity included fully in the model?

Reply by the Author Squeeze effects are fully included in the model. However, there is a significant difference between the behaviour of hard contacts which we are modelling here and the soft contacts with which you are, probably, more familiar. In Reynolds' equation:

$$\frac{\partial}{\partial x}\left(\frac{\sigma h^3}{12\eta}\frac{\partial P}{\partial x}\right) = u\frac{\partial \sigma h}{\partial x} + \frac{\partial \sigma h}{\partial t}$$

when the viscosity becomes large, as it does under piezoviscous contacts in pure rolling, the term on the left of the equation becomes small and the equation degenerates into:

$$\sigma h = f(x\text{-}ut).$$

That is, clearances are transported unchanged through the conjunction. Thus any clearance variations generated in the inlet are carried through the contact and produce the large changes in pressures and stress seen there. Under soft EHL conditions (and when there is significant shear thinning) the left hand side of Reynolds' equation cannot be ignored and squeeze film effects modify the clearances as they pass under the contact, generally reducing the magnitude of the associated pressure and stress perturbations.

Professor H Spikes (Imperial College, London, UK) In all of the work you present, you assume the validity of Reynolds' equation. Do you think that any of the conditions you model are likely to mean that the assumptions used in deriving Reynolds equation may be invalid-such as negligible fluid inertia, no visco-elasticity, no slip at the solids etc.?

Reply by the Author Any EHL model is only as good as the assumptions made in the analysis. I believe, however, that the assumptions of no fluid inertia, no visco-elastic effects, and no wall slip are reasonably accurate for the pure rolling conditions examined here.

Your own work has shown that these assumptions give accurate values of film thickness for smooth contacts under steady conditions provided allowance is made for the presence of surface monolayers.

Our experimental results examining the stresses under rough EHL conjunctions under near rolling conditions gives close co-relation between theory and experiment ignoring these effects (although we do include shear thinning). This suggests that the assumptions are also valid under near rolling conditions where there are large perturbations in pressure and high pressure and clearance gradients – the type of behaviour we are predicting under some dynamic conditions.

It would, however, be extremely valuable to have some experimental results in which film thickness is accurately co-related to dynamic changes in load and velocity.

Dr A Jackson (Exxon Mobil R & E Co., New Jersey, USA). Pure rolling may not truly represent what happens when speed is changed, since one component will be decelerated or accelerated as a result of velocity changes in the other component. This will result in significant variations in tangential stress due to traction forces. Would the addition of these tangential stresses change your conclusion that stress variations under speed changes are not significant?.

Reply by the Author The traction forces will produce changes in the stress distribution and we do include them under sliding conditions.

However, our experience is that they produce only minor changes in stress. Far more significant when slip is present is the reduction of viscosity due to thermal and shear thinning. This reduction in viscosity tends to reduce the magnitude of the pressure perturbations since the clearance changes are smoothed by squeeze action. Thus we believe that our comment that the change in maximum stress under speed changes will be small will still apply under conditions of relative sliding as well as pure rolling. However, it is worth noting, also, that there will be an increase of stress at the surface and, although this will generally still be lower than the Hertz stress, it may be significant if there are high surface residual stresses present.

Professor A Combescure (LMC, INSA, Lyon, France)..What are the stress equations solved in your model?-static or wave equations?. In some of the computations you have stress (von Mises) levels of about 1GPa. If plasticity is present, do you take that into account? What is the expected effect?

Reply by the Author We used a static, plane-stress elasticity model. Dynamic effects are likely to be significant if the period is comparable to, or smaller than, the product of the Hertz width times the velocity of sound in the contacting materials (around 5 km/s for steel). We believe that, for the cases we presented, the periods are likely to be large enough to avoid having to include inertia effects in our stress model.

Only elastic behaviour has been considered. Plastic deformation may be present as part of an initial shakedown process (particularly for rough contacts) but our evidence is that this occurs within a few thousand cycles and that, thereafter, the contact operates entirely elastically. Melan's theorem indicates that the build up of residual stress during shakedown can be investigated using the elastic stress distribution and we see little reason, therefore, to include the additional complexity of plastic deformation in the analysis.

Mr H van Leeuwrn (Eindhoven University of Technology, The Netherlands). The author has given a nice presentation on film dynamics under EHL conditions. The case of sinusoidally varying entrainment velocity has invoked a question on the practical application of the results in real car cam mechanism lubrication. It is assumed that the entrainment action drops to zero twice per cycle. If the period of the sine is very brief, the film thickness will never come close to the quasi-steady state value: the fluid flow is reversed before it comes close to a stationary solution.

I wonder (a) if this was observed, (b) when does it happen, and (c) if it can happen in practice.

Reply by the Author The sinusoidally varying velocity we examined was for a velocity varying as:

$$u = u_0 [\, 1 + 0.5\sin(\omega t)]$$

Thus we did not encounter entrainment reversal in this example.

The entrainment reversal example that was presented does cover the situation you discussed and the analysis shows that the film thickness remains finite. A detailed discussion is given in [6].

Experimental evidence by Spikes [8] also found that the film thickness remains finite during entrainment reversal.

Paper III (ii) "Comparison of Transient EHL Calculations with Start-up Experiments" by Dr M J A Holmes, Dr H P Evans and Professor R W Snidle (Cardiff School of Engineering, Cardiff University Cardiff, UK)

Professor H A Spikes (Imperial College, London, UK). We went to some lengths in our experiments to eliminate possible machine vibrational response but can, of course, never be sure to wholly eliminate this. One controversial possibility that I think we suggest in our paper is that on very rapid start-up, lubricant sticks to the steel but slips against the stationary SiO_2 surface. Thus initially it has an entrainment velocity of 0.4 m/s which then changes to 0.2 m/s as stick between the SiO_2 and lubricant occurs.

Reply by the Authors This is an interesting suggestion and would mean that Reynolds' equation did not apply in the form used in our analysis. The effective entrainment would then be

as suggested by the questioner. However, to produce the observed variation in central film thickness the effective entrainment velocity would still have to vary as is shown in Figure 11. Clearly, if the disk is clamped and the angular acceleration of the ball is constant, as intended, the mechanism suggested by the question would have to alternate between slip and stick at the glass surface several times during the experiment. For the highest acceleration, slip stick cycles would have to continue well after the 4 ms taken to establish the steady state velocity of the ball. For the 20 m/s^2 case slip stick cycles would have to occur during the acceleration phase only. So an initial slip behaviour will not be sufficient to explain the film thickness variations observed.
A further experiment is suggested at the end of Section 5 where the contact would have an initial very small steady entrainment velocity that would establish a full lubricant film prior to the speed increase. This might prove helpful in establishing which of the various explanations for the experimental observations is the most plausible.

Paper III (iii) "Comparison of Transient EHL. Calculations with Shut-down Experiments" by Dr M J A Holmes, Dr H P Evans and Professor R W Snidle (Cardiff School of Engineering, Cardiff University, Cardiff, UK)

Dr P Vergne (LMC, INSA, Lyon, France). The authors had to increase the lubricant temperature to obtain satisfactory starting conditions when comparing their numerical results with experimental data. Thus, they admit that shear heating occurs in the contact inlet.

Do the authors think that it would be realistic to implement a thermal model in their numerical simulations?

Reply by the Authors The cases analysed are in simple sliding as is made clear in the paper, and consequently inlet shear heating will be a factor in determining the initial steady state film thickness. It would be entirely realistic to include a thermal analysis as suggested. This would make the temperature adjustment a calculated value rather than a constant value adjusted to obtain the experimental film thickness. The work reported in the paper was carried out as part of a PhD thesis, and a transient thermal analysis was not developed during that study. A steady state thermal analysis of the initial conditions was carried out as reported in Section 3. The difference between the calculated steady state temperature rise from this thermal analysis and the temperature rise assumed in order to obtain the same film thickness values as are reported in the experiment was small. This gives confidence that the adjustment made was reasonable.

SESSION IV – TRIBOLOGICAL COATINGS- Advances in Surface Engineering (1)

Paper IV (i) "Deformability in the Near Surface Layer of the First Body" by Professor A Eleod[a], Dr L Baillet[b], Dr Y Berthier[b] and T Törköly[a] ((a) Budapest University of Technology and Economics, (b) LMC, INSA de Lyon, France)

Dr M F Workel (Corus Research Development & Technology, Ijmuiden, The Netherlands) How do the flow curves at p=0 for the C45 and aluminium compare to 'traditional' tensile test results, according to the normal B S standards?

Reply by the Authors The flow curves, which were obtained under hydrostatic pressure, remain below the 'traditional' ones, which were determined also by upsetting tests. That means that the deformation resistance of the metals examined decreased under the influence of the superimposed hydrostatic pressure. The reason for this phenomenon is not yet clear. One might perhaps assume that under the influence of the hydrostatic pressure the distance between the atoms decreases slightly and the energy necessary for their removal also decreases. To clarify this question, we intend to repeat the measurement using magnesium specimens, which have a hexagonal structure with only one closely-packed plane.

Paper IV (ii) "Thickness Dependency of the Lubricity of Ag Thin Films Deposited on Si

(111)7x7 Surface" by Mr M Goto, Mr T Yamamoto and Professor F Honda (Toyota Technological Institute, Hisakata, Japan)

Dr P A Dearnley (Institute of Tribology, The University of Leeds, U.K) Did you determine the hardness of the Ag films?

Reply by the Authors No, I didn't. The hardness shown in my presentation is bulk property. To measure and/or to define Vickers hardness is hardly possible for nanometer thick films. But I believe that the fact that the wider worn track width on the 1.5 nm film compared to the width on 170 nm film is responsible for the softening of the film hardness in my experiments.

Paper IV (iii) "Sliding Wear Behaviour of Nano-Crystalline Cr_3C_2-25(Ni20Cr) Coatings at Elevated Temperature" by Mr M Roy, Dr A Pauschitz and Professor F Franek (Vienna University of Technology, Vienna, Austria)

Dr P A Dearnley (Institute of Tribology, The University of Leeds, UK) What are the intended applications for this new material? How is the powder made?

Reply by the Authors Cr_3C_2-25(Ni20Cr) layer is used as wear resistance coating at elevated temperature. It is to be used in gas cooled reactors [1,2], gas turbines [3], fuel rod mandrels and forming tools [4] and brake discs of high-speed trains [5]

Coatings with nano-crystalline grains have certain advantages over coatings with conventional grains. Nano-crystalline materials exhibit improved combination of strength and ductility [6]. Nano-crystalline coatings have reduced friction coefficient [7]. They show favourable oxidation kinetics [8,9]. Some investigators also reported that nano-crystalline coatings are expected to show better thermal resistance and reduced thermal conductivity [10]. Thus it is expected that this coating with nano-crystalline grains will exhibit superior performances compared to similar coatings with conventional grains.

Our own study shows that Cr_3C_2-25(Ni20Cr) layer with nano-crystalline grains has a friction coefficient which is around 20% lower than the friction coefficient of similar coatings with conventional grains at room temperature for a range of test conditions. Against a ferrous base material (100Cr6 steel), nano-crystalline coatings shows comparable performances at elevated temperature. Work is in progress on elevated temperature wear of nano-crystalline coatings against a Ni base high temperature material. An enhanced wear resistance of nano-crystaline coating is realised in this case.

Cr_3C_2-25(Ni20Cr) powder is manufactured by agglomeration through spray drying followed by plasma densification. In the case of nano-crystalline powder an additional step of severe quenching is involved. Further clarification on the manufacturing process can be sought from the supplier. via the authors

References :

1) G.Y. Lai, Thin Solid Films, 64 (1979) 271.
2) C.C. Li, Thin Solid Films, 73 (1980) 59.
3) E. Lugscheider, P. Remer, C. Herbst, K. Yushchenko, Y. Borisov, A. Chernishov, P. Vitiaz, A. Verstak, B. Wielage and S. Steinhauser, Proc. Of ITSC'95, Kobe, May, 1995, 235.
4) J. Mateos, J.M. Cuetos, R. Vivande and E. Fernandez, Tribol. Intl. 34 (2001) 345.
5) M. Watremez, J.P. Bricout, B. Marguet, J. Qudin, J. of Tribology, 118 (1996) 457.
6) H.Hahn, Nanostruct. Mater. 2 81993) 251.
7) Y. Zhu, K. Yukimura, C. Ding and P. Zhang, Thin Solid Films, 388 (2001) 277.
8) G. Chen and H. Lou, Oxidation of Metals, 54 (2000) 155.
9) L. Ajdelsztajn, J. A. Picas, G.E. Kim and F.L. Bastian and J. Schoenung, Mater. Sci. and Engg. A338 (2002) 33
10) M. Cell, Mater. Sci. and Engg. A204 (1995) 246.

SESSION V- TRIBOLOGICAL STUDIES of DRY and LUBRICATED SURFACES (1)- Dry.

Paper V (i) "Modelling of Variable Friction in Contact with Fretting" by Dr M Cheikh, Dr S Quilici and Dr G Gailletaud (Centre des Matériaux P-M Fourt, Ecole de Mines de Paris, France)

Mr T R Hyde (Rolls-Royce plc., Derby, UK) Is there a contact geometry effect on constants in the fretting equation?

Reply by the Authors The contact geometry should not have an effect on the constants of the equation ie. only material / fretting variable effects.

Dr A Kapoor (Sheffield University, UK) During the experiment a wear scar develops and leads to changes in the contact geometry. At the ends of the stroke, the slider pushes against the edges of the wear scar leading to an increase in the lateral force and friction coefficient, which in fretting experiments reaches very high values. Does your model account for this effect?

Paper V (ii) "Micro-abrasive Wear of Semi-crystalline Polymers" by Dr Y J Mergler[a,b] and Dr A J Huis In 't Veld[b] ((a) Dutch Polymer Institute, Eindhoven, The Netherlands, (b) TNO Industrial Technology, Eindhoven, The Netherlands)

Mr A F Beke (University of Sheffield Department of Mechanical Engineering, UK). Polymers have characteristic wear rates at different temperatures. Have you carried out, or considered, tests at different temperatures?

Reply by the Authors Temperature is known to have an effect on wear and friction of polymers. Especially at around transitions, like the glass temperature, wear and friction can be affected. This is an interesting research area for semi-crystalline polymers used in a temperature range where they can undergo a transition, like for instance PA6.6 and PBT with a glass transition temperature of 50°C and 70°C, respectively. Unfortunately, it is not possible yet to carry out micro-abrasion tests at different temperatures with our apparatus.

Professor Dr L Deters (Otto-von-Guericke-Universitaet, Magdeburg Institut fur Maschinenkonstruktion Magdeburg, Germany). Where do you see a practical application for your results and will hard materials other than SiC show the same behaviour?

Reply by the Authors Our results showed that in a three-body wear regime polymers with a relatively low density show a high wear resistance. This knowledge can be used as a criterion in the selection of polymers for specific applications.

Shipway (Wear, **237**, (2000), 90) showed the same wear behaviour when Al_2O_3 was used as abrasive particles compared to SiC, although the absolute value of the wear factor of the measured material was different.

Dr A Kapoor (Sheffield University, U.K.) Your experiments show a switch from 3-body abrasion to 2-body abrasion as the load increased. Was it because of thinning of the water (fluid) film, not letting the abrasive particles go into the contact, or the abrasives getting blunted? Why does this switch (change from 3-body to 2-body abrasion) occur?

Reply by the Authors Fresh slurry is used during the experiments and, therefore, it is unlikely that the switch from two- to three-body wear occurs due to blunting of the abrasive particles. It is unlikely that particles break or blunt as the polymer and the 100Cr6 hardened steel ball are softer than the SiC particles. We expect that at higher load or lower slurry concentration fewer particles are able to enter the contact area, resulting in primarily two-body wear.

Paper V (iii) "Wear Testing of Acetal for the Prediction of Polymer Rolling Bearing Durability" by Mr A F Beke, Dr R S Dwyer-Joyce, Professor F R Jones and Mr J Gregory (The University of Sheffield, Department of Mechanical Engineering, Sheffield, UK).

Dr C Jacq (Snecma Moteurs, Noissy Gamayel, France). What is the influence of the rotation of the rolling bearing on the wear rate due to the evacuation of wear debris ?

Reply by the Authors The influence of rotation of the rolling bearing on wear rate due to wear debris evacuation has not yet been quantified. However, it is likely that due to the rolling nature of the bearing some wear debris will be expelled. Furthermore, due to flash temperatures expected at the ball and raceway contact, and the relative low melting point of POM, it is likely that some wear debris will melt and reattach to the raceway upon cooling. Therefore until the true effects of temperature at the point of contact have been investigated further, the extent of influence that rotation has on wear rate due to wear debris evacuation can not be quantified.

Professor Dr L Deters (Otto-von-Guericke-Universitaet, Magdeburg Institut fur Maschinenkonstruktion Magdeburg, Germany) Is there also an influence on wear by amplitude and frequency, especially of higher frequencies and very small amplitudes?

Reply by the Authors It is expected that there will be an effect on wear by amplitude and frequency, especially at higher frequencies and very small amplitudes. The tests to date have only been done at three frequencies (with 40% variation between minimum and maximum frequency) and at a fixed amplitude. Therefore, until further testing over a larger range of frequencies and amplitudes has been carried out, the influence on wear can not be determined.

Dr A Kapoor (Sheffield University, Sheffield, UK) Could you please explain how the fatigue cracks in your experiments initiate and propogate?

Reply by the Authors In these experiments the fatigue cracks appear to be controlled by stress rather than velocity. Stress being the dominating factor because the initiation and propagation of fatigue cracks were observed along the whole wear track length and not just at regions of maximum velocity. It is also believed that due to frictional heating induced by sliding and the thermal insulating properties of POM, the surfaces of the POM specimens soften and by repetitive cycle loading fatigue cracks are initiated and propagated.

Paper V (iv)."Self-lubricating Dry Rubbing Bearings-Fundamentals and Methods of Calculation" by Professor L Deters[a],Mr F. Mueller[a] and Dr M Berger[b]. ((a) Otto-von-Guericke-University, Magdeburg, Machine Elements and Tribology, Magdeburg, Germany (b) Glacier Garlock Bearings, Heilbronn, Germany)

Mr T Hyde (Rolls-Royce plc, Derby, UK) Is there an upper limit of temperature for this bearing configuration?

Reply by the Authors The upper limit of temperature of self-lubricating dry rubbing compound bearings made of a steel backing with a porous bronze layer filled with a PTFE/solid lubricant mixture corresponds with the lowest upper temperature limit of one of the components of the bearings (bronze, PTFE, MoS_2, or lead etc).

The upper temperature limit of the bearing types, which are discussed here, is approximately 300°C. The critical upper temperature is also indirectly included in the design rule that the shear stress ratio τ_f^* must be less than 1 ($\tau_f^* \leq 1$). After exceeding the limit value $\tau_f^*=1$ friction, wear and temperature increase very fast and the bearing is going to fail very quickly.

SESSION VI – TRIBOLOGICAL STUDIES of DRY and LUBRICATED SURFACES (2) Dry and Lubricated

Paper VI (i) Invited Paper "Tribological Design Criteria and Standards for Spacecraft Mechanisms" by Dr R A Rowntree (Engineering Solutions, AEA Technology, Warrington, UK)

Dr B Roylance (University of Wales, Swansea, Department of Mechanical Engineering, Swansea, UK). Can you provide some indications of the type of condition monitoring sensors that assist in monitoring the tribological performance while operating in space?

Reply by the Author The general approach is to prove tribological component performance and life by test before launch, with the appropriate testing margins as given in the paper. In practice it is relatively uncommon to be able to switch on some lubricant re-supply system and full redundancy in many mechanisms is not possible. But knowledge if a component is failing is clearly of use, particularly if mission critical and where redundancy is available. This is usually surmised from measurements of motor currents, local temperatures and mechanism performance (eg loss in pointing accuracy or scan angle). One exception to this was the PAX experiment on Olympus that was in essence a micro-accelerometer, originally placed on board to measure spacecraft disturbances. When this spacecraft was in recovery after an anomaly, the PAX instrument was used to diagnose bearing condition on the various mechanisms employed. These measurements were of considerable value and since Olympus many spacecraft have flown with micro-accelerometers as part of their payload/instrumentation. However the author is not aware of further applications of this type of instrument for indicating tribological performance.

Mr A F Beke (University of Sheffield, Department of Mechanical Engineering, Sheffield, UK). With longer and longer storage times of spacecraft does "shelflife" have to be specified and taken into account when the specifications are drawn up? If so, what issues are of importance?

Reply by the Author Mechanism and spacecraft specifications invariably state on-ground storage periods that must be tolerable without degradation to any component therein. These storage periods have increased over the years, for example the author tested flight parts for the Hubble Space Telescope some 10 years before it's actual launch. The author is also aware that prior to the launch of Envisat reviews were taken to ensure no tribo-component deterioration while awaiting launch. Issues of importance include corrosion (not all materials employed are fully corrosion resistant), time-dependent creep of polymeric parts under high load, gravity creep of fluid lubricants, increased moisture absorbance or oxidation of dry lubricant films, and too much ground testing which can lead to premature lubricant degradation or particle generation.

Paper VI (ii) "A Simple Numerical Procedure to Calculate the Input Data of Greenwood-Williamson Model of Asperity Contact for Actual Engineering Surfaces" by Dr E Tomanik, Mr H Chacon and Mr G Teixeira. (Mahle-Tech Center, S. Paulo, Brazil).

Professor Dr L Deters (Otto-von-Guericke-Universitaet, Magdeburg Institut fur Maschinenkonstruktion Magdeburg, Germany). How does filtering of the original roughness trace influence the calculation results of the G & W parameters?

Reply by the Authors With the filtering we have tried to reproduced the roughness profile, i.e. to omit form and waviness. Contact between two surfaces is obviously affected by their shape, being questionable to filter surface form and waviness. In the simulation code used in this work, the form is considered separately (e.g. the ring contact face was assumed to be a parabola with 10 μm barrel drop). We have not yet considered the profile waviness influence on the contact.

Dr Q Zhou (Jaguar Cars Ltd., Coventry, UK) Can you consider the effect of cross hatch honing angle on film thickness?

Reply by the Authors I don't think so. Although different honing cross angles may produce different calculated surface parameters, the oil film thickness calculation, per se, does not consider the cross honing angle.

Professor Dr Ing G Poll (Institut fur Maschinenelemente, Konstruktionatechnik, University of Hannover, Germany). In the paper, roughness values for the four different surfaces were compared with computed film thicknesses. How was the computation done? I understand, it is

based on Reynolds' equation. But how was the roughness influence accounted for? Obviously, between the different surfaces there is no difference in macro-geometry, so the differences in computed film-thickness must be related to micro-geometry.

Reply by the Authors The calculation of oil film thickness is started considering the lubrication regime as hydrodynamic. The oil film thickness is calculated by the Reynolds equation. If the oil film thickness is equal or higher than three times the combined roughness, the calculated oil film thickness can be accepted. Otherwise, part of the load will be supported by the contact between asperities. Oil thickness will be higher than that calculated by the Reynolds equation. The solution is reached interactively. The hydrodynamic pressures (P_H) and the minimum lube oil thickness (H_m) are calculated. If H_m is lower than three times the combined ring/bore roughness, the asperity contact pressures are calculated. The algorithm is repeated until the sum of hydrodynamic and asperity contact pressures equals the applied load.

Paper VI (iv) "Behaviours of Ultra-Thin Lubricant Films Between Solid Surfaces Having Atomic Scale Roughness" by Mr T Okumura, Dr J Sugimura and Professor Y Yamamoto. (Department of Mechanical Engineering, Kyushu University, Fukuoka, Japan).

Dr B Bou-Said (LMC, INSA, Lyon, France). Please explain your definition of traction coefficient for molecular dynamics.

Reply by the Authors Traction coefficient is defined as the ratio F/N, where N is the normal load and F is the x-direction force acting on the fixed atoms of the upper surface averaged over time.

Professor F Honda (Toyota Technological Institute, Nagoya, Japan). You concluded the molecular form and surface atomic structure strongly affect the traction coefficient. It looks reasonable if you have started your calculation on the assumption that both factors are effective. Please make clear that your conclusions came from your initial assumptions or from your calculation results.

Reply by the Authors From the calculation results. We had expected, as you suspect, that surface roughness of atomic scale would affect traction in some way, but it was only after we obtained MD results that we knew for certain that the molecular interactions induced by the presence of the atomic steps influenced adsorption, slip, aligning and bridging of lubricant molecules, which in turn caused changes in the traction coefficient.

SESSION VII
BIO-TRIBOLOGY (1) – Hip and Knee Joints

Paper VII (i) Invited Paper "Fluid-Film Lubrication in Artificial Hip Joints" by Dr Z-M Jin[a], Dr J B Medley[b] and Professor D Dowson[c] ((a) University of Bradford, Bradford, UK, (b) University of Waterloo, Waterloo, Ontario, Canada, (c) Institute of Tribology, University of Leeds, UK).

Dr C J Hooke (University of Birmingham, Birmingham, U.K.). You found that the film thickness could be estimated adequately on the basis of a steady state analysis. Is this particular to the hip joint design that you were investigating or is it a general result that applies to all designs?

Reply by the Authors It is generally true that the transient minimum film thicknesses in both natural synovial joints, such as the ankle (Dowson and Jin, 1987) and the hip (Jin et al, 1993), and artificial hip joints, such as UHMWPE-on-metal (Jalali-Vahid and Jin, 2002) and metal-on-metal (Chan et al 1998), are relatively constant during steady-state walking cycles, and are very close to the quasi-static predictions based on the average load and the average speed.

Dr S L Smith (Centre for Design Research, Northumbria University, Newcastle-upon-Tyne). What cup deformations were calculated for the thinnest cups-and would the deformation compromise the longevity of the cement liner?

Reply by the Authors The cup deformation depends on the bearing materials and structural support and fixation. For the UHMWPE-on-metal hip joints, the deformation of the cup is mainly on the bearing surface, and is unlikely to have any significant effect on the longevity of the cement liner, particularly for the metal-backed cup. For metal-on-metal hip joints, the deformation of the cup consists of both the surface and the structural deformation, which could introduce a relatively large strain within the cement liner.

Professor J F Booker (Cornell University, Ithaca, NY, USA). Please comment on the relationships of predicted minimum film thicknesses and surface roughnesses for the cases considered.

Please also comment on the theoretical treatment of lubricant cavitation (if any).

Reply by the Authors For UHMWPE-on-metal hip joints, the composite surface roughness (mainly from the UHMWPE bearing surface) is often quoted as 1 μm (Jin et al 1997). Although the UHMWPE bearing surface is generally smoothened by either wear or creep, the average roughness is still greater than 0.05 μm (a detailed discussion was given on Page 806, In 'Thinning Films and Tribological Interfaces', Eds D Dowson et al, Elsevier, 2000). Therefore, the predicted minimum film thickness for UHMWPE-on-metal hip joints is considerably smaller than the surface roughness of the UHMWPE bearing surface and a mixed lubrication regime is therefore predicted, particularly if a smaller, and more realistic viscosity value for synovial fluid is used. For metal-on-metal hip joints, the predicted film thickness is generally reduced, compared with UHMWPE-on-metal hip implants, but the composite surface roughness of the metallic bearing surfaces is significantly reduced to as low as 0.014 μm. Therefore, it is more likely that a fluid film lubrication regime exists in metal-on-metal hip joints than in UHMWPE-on-metal ones.

The theoretical treatment of the lubricant cavitation was based on the Swift-Steiber-Reynolds cavitation boundary condition adopted on the outlet boundary.

Additional Reference:

Dowson D and Jin Z M (1987) An analysis of micro-elastohydrodynamic lubrication in synovial joints considering cyclic loading and entraining velocities, In Fluid Film Lubrication - Osborne Reynolds Centenary (Eds D Dowson, C M Taylor, M Godet and D Berthe) (Amsterdam: Elsevier), pp.375-386.

Paper VII (ii) **"Wear of Alumina / Zirconia Nano-Composite for Ceramic / Ceramic Joint Prostheses"** by Professor K Ikeuchi[a], Dr Y Morita[a], Dr K Nakat[a], Dr Y H Kim[b], Dr T Sekino[b] and Dr K Niihara[b] ((a) Institute for Frontier Medical Sciences, Kyoto, Japan (b) The Institute of Industrial and Scientific Research, Osaka, Japan)

Professor N Spencer (Laboratory for Surface Science and Technology, ETH-Zurich, Switzerland). Could you please tell us the thermal conductivity of the alumina and zirconia?

Reply by the authors The thermal conductivity of the alumina is 30 *w/(m.k)* and that of zirconia 3 *w/(m.k)*. The thermal conductivity of the composite ceramics are in between them.

Dr J B Medley (University of Waterloo, Waterloo, Ontario, Canada). Do you intend to perform simulator tests under various adverse conditions, such as with abrasive particles, longer test times, larger zones subject to impact etc?

Reply by the Authors Yes, we intend to practice simulator tests for ceramic/ceramic hip prostheses. On the other hand, we are performing edge on flat wear tests simulating joint motion under dislocation. The test includes collision during sliding.

Mr P Paré (University of Waterloo, Waterloo, Ontario, Canada). Can you describe the process of tribo-chemical lubrication of nano-composite ceramics?. Which molecules are implicated? How do these molecules adhere to the surfaces?

Reply by the Authors When alumina slides in water, tribochemical films appear due to the reaction with water [a1, a2]. For nano-composite ceramics, it is likely that the alumina phase reacts with water producing alumina hydrates.

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK). Some of the measured wear rates are very low. Did you detect any contribution of chemical wear, with either Alumina or Zirconia in the presence of distilled water or 30% serum?

Reply by the Authors It is reported that alumina hydrates appear on alumina surface during sliding in water [a1, a2]. The authors think that adhesion and third body abrasion are prevented by the tribo-chemical reaction because the wear products are softer than alumina. On the other hand, zirconia is not reactive with water and severe adhesion and third body abrasion were observed on the zirconia surface.

Additional References

a1. R S Gates, S M Hsu and E E Klaus, Tribology Trans., 32, 3 (1989), 357-363.
a2. M G Gee and N M Jennett, Wear, 193 (1995) 133-145.

Paper VII (iii) "Surface Characteristics and Tribological Behaviour of New Zirconia Ceramics for Cementless Knee Joint Prostheses" by (Dr Y Sawae[a], Dr T Murakami[a], Dr T Sawano[a], Dr I Noda[b] and Dr T Shimotoso[b]) (Kyushu University, Japan)

Dr J B Medley (University of Waterloo, Waterloo, Ontario, Canada). Have you performed film thickness calculations for your novel knee implants to support your suggestion that fluid films may be providing some protection to the polyethylene surface?

Reply by the authors In this study, the friction and wear characteristics of the uhmwpe-on-htoz sliding pair were examined experimentally and compared with other material combinations for the sliding surface of joint prostheses, such as the uhmwpe-on-tzp. The results of these experiments indicated that the friction and wear characteristics of the uhmwpe-on-htoz sliding pair were potentially superior to that of the uhmwpe-on-tzp sliding pair. Although we have not calculated yet the fluid film thickness in our pin-on-disk and pin-on-plate tests which have a flat-on-flat configuration and are not suitable to use ehl theory, we postulate from the results of the friction measurement that hydrodynamic lubrication has been relatively effective for the uhmwpe-on-htoz sliding pair compared with uhmwpe-on-tzp and possibly reduced the wear amount of uhmwpe.

The UHMWPE-on-HTOZ sliding pair has successfully showed preferable tribological behaviours in the screening friction and wear tests. In the next phase, the tribological characteristics of our novel knee implants made from UHMWPE-on-HTOZ will be examined in a knee joint simulator test. The film thickness calculation will also be performed with the design parameters of our knee joint prostheses to examine the effect of the surface characteristics of HTOZ on the friction and wear characteristics under the physiological walking condition in detail.

Paper VII (iv) "The Relationship Between Steady-state Wear Rate and Theoretical Film Thickness in Metal-on-Metal Total Replacement Hip Joints" by Professor D Dowson, (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK).

Professor N Spencer (Laboratory for Surface Science and Technology, ETH-Zurich, Switzerland). Is the (λ) parameter really useful? After all, it says nothing about the amount of the roughness frequency component that lies in the range of the contact area. Surely this is what is important?

Reply by the Author Professor Spencer raises an important question. The concept of the Lambda Ration is useful; the difficulty lies in its evaluation and application. Since many , *'hard-on-hard'* total replacement joints appear to operate in the mixed lubrication regime, both the wear and the friction of such components will depend strongly upon the sharing of the applied load between asperity contact and fluid-film action. Both friction and wear can vary by some two orders of magnitude in the transition from full boundary to fluid film conditions. Furthermore, this transition can be achieved with relatively small changes in operating conditions, as demonstrated by the results shown in the present paper. The Lambda Ratio reflects the

severity of boundary lubrication conditions and hence it is, in principle, an important indicator of low ($\Lambda > 3$) or high ($\Lambda \approx 1$) wear.

Most of the problems encountered in the determination of the Lambda ratio are related to the estimation of the composite surface roughness, since the effective film thickness can be calculated with fair accuracy. If the solid bearing surfaces were isotropic and they presented Gaussion distributions of heights, the Lambda ratio would be meaningful. However, the running-in process modifies the initially manufactured surface topography and this needs to be considered in any evaluation of the effective Lambda ratio. The running surface may be neither isotropic nor Gaussian.

Professor Spencer seems to be concerned about the surface characteristics within the contact zone, but for hard-on-hard components this is unlikely to be very different to the surface over a wider region. Micro-elastohydrodynamic action, which is so powerful in natural (soft) joints, is unlikely to be significant in hard - on - hard joints.

One way forward in the development of design guides would appear to be to work with a *'modified'* Lambda Ratio taking account of non-isotropic and non-Gaussian surface topography. Such work would have to be supported by extensive studies of the surface topography of hip replacements in both simulators and in vivo, as functions of time. In the meantime, it has been shown in the present work that film thickness itself can be a good indicator of the tribological performance of hard-on-hard total replacement joints.

Paper VII (v) "On The Role of the Lambda Parameter in Simulator Wear of Metal-on-Metal Hip Implants" by Mr P Paré, Professor J B Medley, Dr F W Chan and Mr S Young, (University of Waterloo, Waterloo, Ontario, Canada).

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK). Great care has to be exercised in relating wear rates to the lambda ratio . The concept is clear; the ratio of effective film thickness between ideally smooth bearing surfaces to some acceptable measure of asperity heights must indicate the extent to which lubrication can be effective in separating the surfaces and controlling wear. This concept first came to prominence in the 1960's in relation to studies of gear lubrication, at a time when the mode of lubrication in gears was still under active investigation.

It was found that the ratio of calculated elastohydrodynamic film thickness to composite surface roughness correlated exceedingly well with the fatigue life and wear of gears. A value of lambda ratio of about 3 resulted in near zero wear and maximum life of gear sets and this value soon appeared in design procedures as the desired target for designers; the principal disposable variables being the effective viscosity of the lubricant and the surface finish of the gears.

It is interesting to note that when we introduced the concept of the lambda ratio into studies of the wear of total replacement joints, the mode of lubrication in such bearings was also ill-understood. In the previous paper the point presented was that steady state wear in metal-on-metal joints appeared to correlate quite well with the calculated elastohydrodynamic film thickness, but not with available information on the lambda ratio. The difficulty appears to arise mainly in relation to the specification of the latter quantity.

A number of key requirements thus emerge in the use of the lambda ratio;-

(1) The measurement of '*roughness*' of the exceptionally smooth femoral and acetabular components must be made with great care.
(2) Any deviation from non-Gaussian distributions of heights on the surface profiles should be noted. If the profiles are non-Gaussian some additional factors used in metrology should be adopted in the calculation of the effective roughness and hence lambda ratio.
(3) The roughness should be recorded in a period appropriate to the wear measurements-eg. late in the test for steady-state wear and early in the test for running-in wear.

It would be helpful if the authors of this interesting paper could confirm that their roughness

measurements were made at times strictly corresponding to the recorded wear rates.

Was there evidence of any non-Gaussan distribution of heights, associated with scratches or isolated peaks on the surface traces?

Reply by the Authors As proposed in the present discussion of Professor Dowson, it would be of great interest to have frequent roughness measurements for the initial high wear rates followed by less frequent periodical measurements during the steady state regime of wear. Unfortunately, the implant roughness was only measured at 0, 1, 1.9, 3 and 4 million cycles. In retrospect, more frequent roughness measurements should have been made and this is planned for future testing.

The surface micro-geometry was modified by the presence of superficial scratches and some deep scratches that formed during the wear process. Using the average and standard deviation of the measured traces, a Gaussian distribution was superimposed on the frequency density of the surface height measurements. The initial distribution of surface height was close to Gaussian (Fig. A1).

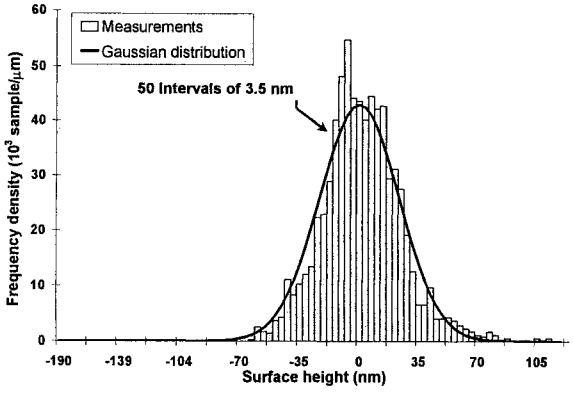


Fig. A1: Frequency of asperity height (Ultamet #7 - head 90° at 0 Mc).

However, the wear of the asperity tips and the numerous superficial scratches (depth ≤ 0.1 μm) made the surface roughness non-Gaussian (Fig. A2).

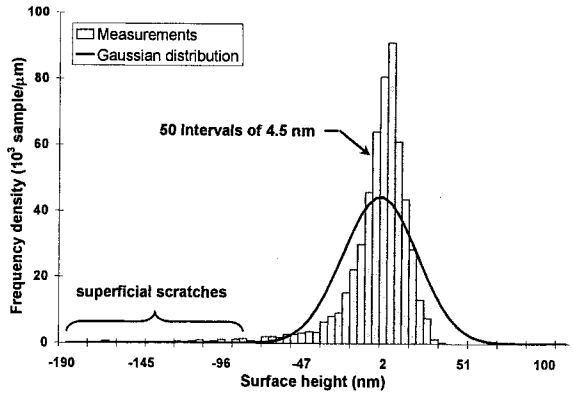


Fig. A2: Frequency of asperity height (Ultamet #1 - cup 23° at 3 Mc)

The less frequently occurring deep scratches were excluded from the roughness measurements used in calculating the lambda values as discussed towards the end of the "Materials and Methods" section. In the literature, many papers have dealt with the relationship between roughness and film thickness. For example, Masen et al [R1] results suggest that for low entrainment velocity, the lubricant film is thicker for rougher surfaces. However, as the entrainment velocity increases, the smoother surfaces are protected by thicker films. Unfortunately, a simple procedure to account for non-Gaussian distribution of surface heights was not found. A correction factor to convert non-Gaussian to effective Gaussian distributions for the purposes of determining an effective lambda value would indeed be useful, as suggested by Professor Dowson. Such an approach would be almost certain to give a more precise and accurate determination for the extent of elastohydrodynamic lubrication protection.

Reference

R1. **Masen, M A, Venner, C H, Lugt, PM, Tripp, J H** Effects of surface micro-geometry on the lift-off speed of an ehl contact. *Tribology Transactions*, 2002, Vol 45, No. 1, 21-30.

Paper VII (vi) " Advances in Simulator Testing of Orthopaedic Joint Prostheses" by Dr T D Stewart, Dr S Williams, Dr J L Tipper, Professor E Ingham, Mr M H Stone and Professor J Fisher.

Professor J B Medley (University of Waterloo, Waterloo, Ontario, Canada). Can you discuss how widespread micro-separation is in clinical practice, whether it is likely to reduce with time and whether surgical techniques can be introduced to prevent it from occurring?

Reply by the Authors. Nevelos et al observed stripe wear associated with microseparation on the majority of retrieved Mittelmeier prostheses manufactured from the first generation of alumina. Current designs, however, are manufactured from hot isostatic pressed alumina with a refined grain size. There are also many design changes in these new modular devices. We are working with Bill Walter an Australian surgeon who has observed microseparation stripe wear in approximately 40 percent of retrievals from these new designs, however, clinical data is limited. Fluroscopy studies by Dennis and Komistek have shown joint separation in-vivo in well fixed UHMWPE hip and knee prostheses. Joint separation and stripe wear will most certainly be effected by surgical technique and this is a current research area we are most interested in pursuing.

SESSION VIII TRIBOLOGICAL COATINGS- Advances in Surface Engineering (2)

Paper VIII (i) "Rolling Contact Fatigue of Nitrided 32CrMoV13 Steel" by Dr C Jacq[a], Dr D Nelias[b], Dr G. Lormand[c], Dr A Vincent[c] and Dr G Dudragne[d] ((a) Direction Technique Snecma Moteurs, Moissy-Cramayel, France (b) LMC, INSA de Lyon, France (c) INSA de Lyon, GEMPPM, France (d) SNR Roulements, Annecy, France)

Mr T Hyde (Rolls-Royce plc, Derby, UK). Does the increased yield stress of the nitrided layer cause stress concentration due to substrate yield, and if so, does this detrimentally affect the fatigue performance?

Reply by the Authors This particular point has not been studied in details by the authors. However, if the substrate yields and the nitrided layer does not, stress concentration might effectively occur because of the non-uniform plastic strains at the interface between the plastic and non plastic domains. The stress concentration effect should strongly depend upon the severity of the transition between elastic and plastic domains. In the case of nitrided 32CrMoV13 steel, two points must be considered. First, the evolution of the micro-yield stress is almost linear between the surface and the nitrided depth. Second, since the shape of the hardening law does not change in the nitrided layer (parameters C and n of the Swift law are constant in the whole material), no distortion is generated as plastic strains are created. Consequently, the transition between elastic and plastic domains might be smooth, and the stress concentration limited.

This question points out the importance of the choice of the hardening treatment depending upon the aimed application. Hence it has been shown by Vincent et al. [1] (figure 1) that, for a nitrided steel, the rolling contact fatigue endurance limit of the material strongly depends upon the equivalent radius of contacting bodies because of the evolution of the maximum stress depth compared with a fixed micro-yield stress profile, linked to the material.

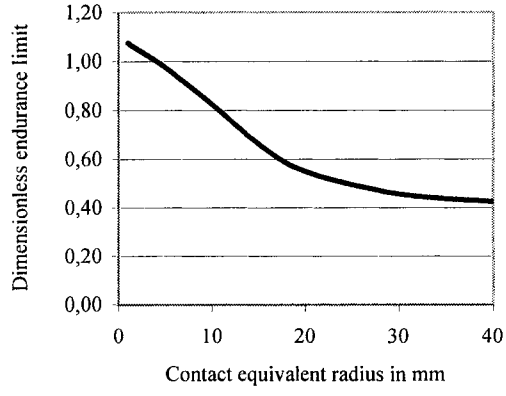


Figure 1. Evolution of the rolling contact fatigue endurance limit H1 with the equivalent radius of curvature for a spherical point contact [1].

[1] A Vincent, R Fougères, G Lormand, G Dudragne, D Girodin. « A Physically Based Endurance Limit Model for Through Hardened and Surface Hardened and Surface Hardened Bearing Steels », *Bearing Steel Technology*, ASTM STP 1419, J M Beswick, Ed. American Society for Testing Materials, West Conshohocken, PA, 2002, pp.459-473.

Paper VIII (ii) " Effect of Hf Addition on the Tribological Performance of ZrN Coating" by Dr E Atar[a], Dr H Mindivan[b], Professor Dr H Cimenoglu[b] and Dr E S Kayali[b] ((a) Bebze Institute of Technology, Gebze Kocaeli, Turkey (b) Istanbul Technical University, Istanbul, Turkey).

Professor A T Alpas (University of Windsor, Ontario, Canada). Hf addition in ZrN increases wear resistance according to your scratch test results. Could you please discuss the physical mechanism by which this increase is achieved. (what is the role of Hf in increasing the wear resistance? What is the structure of ZrHN?)

Reply by the Authors A scratch test carried out under constant load revealed that the cracks developed inside the scratches were thicker and wider for ZrN coating compared to Zr(Hf,N) coating. The cracks formed on the Zr(Hf,N) coating were more frequent and closer. This qualitative observation indicates that the bonding strength between the coating and the substrate increases by alloying of the ZrN coating with Hf. On the other hand, we have measured the coefficient of friction values during reciprocating wear testing. The Zr(Hf,N) coating exhibited relatively lower coefficient of friction values than ZrN coating against both Al_2O_3 and steel balls. This result indicates that Hf addition encouraged decrease of frictional force between the coating and the ball under dry sliding conditions. In the literature the crystal structure of ZrN and HfN are given as cubic. Alloying of Zr with Hf forms substitutional solid solution.

Paper VIII (iii) "Tribological Properties of TiAlSiN Film Coated by AIP Equipment onto SKD61 Substrate" by Professor R Murakami, Mr Y Matsuda and Dr D Yonekura (The University of Tokushima, Tokushima, Japan).

Dr P A Dearnley (Institute of Tribology, School of Mechanical Engineering,University of Leeds, UK). The coating material you have investigated (TiAlSiN) is very unusual,-please explain why you selected this for your experiments.

Reply by the Authors TiAlN film has a high wear resistance against a steel, and this film is used for various kinds of cutting tools. However, TiAlN film has also the bad adhesive property between the material and the film. The recent study topics is that such a behaviour should be improved. To protect the adhesion of the steel to the TiAlN film, we selected Si as a doping element in the TiAlN film. The TiAlSiN film has high fracture toughness and adhesion resistance against the steel In the present study, TiAlSiN film has protected the adhesive behaviour of the steel to the film and then showed the excellent wear behaviour.

Paper VIII (iv) " Wear Mechanisms in Thermal Spray Steel Coatings" by Dr A Edrisy[a], Dr T Perry[b], Dr Y T Cheng[b] and Professor A T Alpas[a] (((a) University of Windsor, Ontario, Canada (b) General Motors Research and Development Centre, Warren, MI, USA)

Dr. G Repphun (Robert Bosch GmbH, Stuttgart, Germany) What is the influence of humidity on wear rate and the wear mechanism?

Reply by the Authors At high humidity conditions the wear rates decreased. The decrease in wear rates started at higher relative humidity levels as the load and sliding velocity increased. The Fig. 6 in the paper summarized the effect of humidity on the wear rates of the coating for different loading conditions.

The effect of humidity on the wear rates and wear mechanisms of low carbon steel coatings produced using the PTWA process has been investigated in more details by the authors [4].

Professor Dr H Cimenoglu (Istanbul Technical University, Istanbul, Turkey). Have you examined the effect of carbon content on the wear performance of the coating?

Reply by the Authors The coating was sprayed using low carbon steel with a nominal 1020 composition as wire feed stock. In all the wear tests the same coating composition was used. The wire with this composition is of significant interest for the automotive industry because of cost considerations, and good ductility.

SESSION IX – FRICTION, WEAR and LIFE OF LINE CONTACTS IN MACHINE ELEMENTS (1) – Life Predictions

Paper IX (i) Invited Paper "The Role Wear in Enhancing Rail Life" by Dr A Kapoor, Dr D I Fletcher and Dr F J Franklin (The University of Sheffield, Department of Mechanical Engineering, Sheffield, UK).

Professor Dr L Deters (Otto-von-Guericke-Universitaet, Magdeburg Institut fur Maschinenkonstruktion Magdeburg, Germany). Is it possible to include inclusions like MnS in the model as these are very important for crack propogation?

How is the wear rate or wear coefficient determined?

Reply by the Authors The brick model was designed with the idea of modelling microstructure. Brick size is one of the model's parameters and can be chosen to be, for example, similar to the grain size of steel. The initial hardness, strain-hardening characteristics and critical plastic shear strain for failure can, if desired, be specified on a brick-by-brick basis. With an appropriate choice of brick size, therefore, it should be possible to model MnS inclusions. However, the brick model does not (currently) model crack propagation, although it can and has been used to investigate crack initiation.
Bricks are removed from the surface under a set of specified conditions and the wear rate is simply the depth of material removed per cycle (usually averaged over several hundreds of cycles).

Professor Dr Ing. G Poll (Institut fur Maschinenelemente, Konstruktionatechnik, University of Hannover, Germany). I like the conclusions of the paper, but I would like to suggest to look into another mechanism as an explanation. The highest stresses most of the time exist beneath the rail surface. If a crack penetrates there, its propogation or growth rate will be high. When removing material from the surface, the most highly stressed zone moves deeper into the material together with the surface. That way, cracks may be prevented from reaching this zone altogether, when the wear rate is large enough. This mechanism should be considered when continuing the research.

Reply by the Authors The mechanism by which crack growth rate for contact stress driven cracks is related to crack length or depth is dealt with in Section 3 of the paper. Surface wear is automatically considered in the calculations because the crack length used is measured from the surface after taking account of wear. It is correct that as the surface is worn away the stress field from the surface contact also moves into deeper material, and it may therefore be possible to prevent cracks reaching this high stress region if the wear rate is high enough. However, current results indicate that cracks driven by ratchetting almost always reach the high sub-surface stresses, even in high wear situations.

Dr A Olver (Imperial College, Tribology Section, London, UK) The 'brick' wear model assumed material properties which varied by a few percent. However, pearlitic rail steels differ by hundreds of percent. How might this affect the wear predictions?

Reply by the Authors With pearlitic rail steels there is variation within individual grains. The grains are composed of laminae of cementite and ferrite, materials which differ greatly, the cementite being far harder than the ferrite and also very brittle in comparison. The brick model is capable of modelling such a structure, but is a ratchetting model and cementite will not fail by ratchetting – a more sophisticated model is required.

At a larger scale, where variation within grains is no longer a factor, microhardness measurements of unused rail (Grade A, unused) give means of about 210Hv with standard deviations of about 15Hv. Typically the brick model is used with a standard deviation between 4% and 10% of the mean.

Early results indicate that large variations (or, technically, a large proportion of bricks with shear yield stress exceeding the orthogonal shear stress near the surface) have the effect of reducing the wear rate because clusters of hard bricks act

protectively and prevent the large amounts of damaged material from exposure to the surface.

Paper IX (ii) "Laboratory-based Wear Life Testing of Sliding and Rolling Contacts-The First Step Towards Technology Transfer' by Dr T P Sperring and Dr B J Roylance (University of Wales, Swansea, Department of Mechanical Engineering, Swansea, UK).

Dr J A Williams (The University of Cambridge, Engineering Department,Cambridge, UK). In using observations on the morphology of retrieved wear debris particles to indicate potential component failure I imagine the skill and diligence of the operator is critical. To what extent is it possible to automate this process?

Reply by the Authors The Syclops system currently in use by the RAF is a manually operated system in order to comply with their monitoring procedures. There is no reason in principle why it cannot be adapted to function as an automated device and research is continuing to establish how best this can be achieved within the prescribed limits of operation. The ultimate goal of this investigation is to create an automated analysis tool because of operator subjectivity. Additional research into the automatic input of the shape and surface characteristics is on-going, using a variety of artificial intelligence-based techniques. LaserNet, produced by the US Naval Research Laboratory already extracts size and silhouette features of wear debris in oil and can also assess machine condition using information on particle size distribution.

Mr. A F Beke (University of Sheffield, Department of Mechanical Engineering, Sheffield, UK). How difficult is it to obtain wear debris from real working planes of the RAF?

Reply by the Authors The Royal Air Force routinely takes wear debris samples from the oil systems of their aircraft which have been captured using a magnetic detector plug (MDP) in accordance with AP119, the RAF standard procedure for Early Failure Detection Centres. Once removed from the oil-line, the MDP are washed with solvent and inspected under an optical microscope. The debris is then wiped onto a piece of adhesive tape and stored in an archive which also contains maintenance action information. Once the archives are taken to the laboratory, it is a simple operation to remove the debris from the tape and obtain images suitable for inclusion in the Syclops database.

Paper IX (iii) "Life Prediction for Surface Crack Initiated Contact Fatigue of Silicon Nitride Bearing Balls" by Dr Y Wang and Professor M Hadfield (Bournemouth University, School of Design, Engineering and Computing, Bournemouth, UK).

Professor T H C Childs (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK). Your last slide claims that you can predict fatigue life. But you presented some theory and some experimental results. I could not see where you brought the two together. What can you predict?- and how accurately?

Reply by the Authors Prediction of fatigue life in this paper is based on 3D crack propagation analysis using fracture mechanics. If the stress intensity values along the crack front are known, the possibility of crack growth can be predicted through comparing SIF values to threshold value ΔK_{th}. The distribution of the stress intensities along the crack front at various contact stresses are calculated and shown in Figure 8. The crack propagation threshold value for the silicon nitride ceramics ranges from 2-3 MPa. According to the calculated results (Figure 8), if the contact stress is less than 5.0 GPa, the spalling fatigue will not happen because the fatigue crack front is hard to propagate at that loading condition. This prediction is based on the friction coefficient $f = 0.05$. The influence of the friction coefficient on the stress intensity factors is also calculated as in Figure 10. Friction coefficients have a significant effect on the stress intensity factor value. Increasing traction force increases the magnitude of the stress intensity values. The fatigue life decreases as tangential traction increases. The results (Figure 12) from experimental tests are in good agreement with the above prediction. Therefore, the SIF values along the crack front can provide the necessary information to determine whether the ceramic ball would survive or fail.

Paper IX (iv) "Experimental Investigations on Rolling Contact Fatigue of Dented Surfaces Using Artificial Defects: Sub-surface Analysis" by Dr Y Diab, Dr S Coulon, Dr F Ville and Professor L Flamand (LMC, INSA de Lyon, Lyon, France).

Dr T Sperring (University of Wales, Swansea, Department of Mechanical Engineering, Swansea, UK) Have you observed sub-surface cracks that have no associated surface wear?

Reply by the Authors Firstly, the tests performed did not induce wear but only rolling contact fatigue. Secondly, you could observe on the figure 4.a2 that the crack named C2 (parallel to the surface) is not yet propagated. But owing to C1 we could suppose that the crack C2 is the first step before micro spalling. To summarise, during the tests only rolling contact fatigue was observed and all the cracks seem to induce micro spalling

Paper IX (v) "Incorporating a Railway Wheel Wear Model into Multi-body Simulations of Wheelset Dynamics" by Mr A Ward, Dr R Lewis and Dr R S Dwyer-Joyce (The University of Sheffield, Department of Mechanical Engineering, Sheffield, UK).

Professor Dr. L Deters (Otto-von-Guericke-Universitaet, Magdeburg Institut fur Maschinenkonstruktion Magdeburg, Germany). Why is the lubricated case not taken into account as it is very common?

How is the traction coefficient determined which is necessary for the wear calculations?

Reply by the Authors The authors are aware that on many severe curves lubrication is provided to reduce the severity of wear to both wheel and rail.

It would be possible to run lubricated twin disc tests to establish the effect on the wear coefficients and friction and input this to the model.

It should be noted, however, that the wheel would only pass through the lubricant for a fraction of a second and as such it would probably have little impact on the overall wear, the lubrication primarily being to protect the rail which continually experiences the severe contact conditions.

Global traction coefficients for the wheel/rail contact are determined within ADAMS/Rail. Local tractions for the discretised contacts are determined in the local contact analysis, which is part of the wear code.

SESSION X
COMPONENT TRIBOLOGY (1)- Plain Bearings

Paper X (i) Invited Paper "Tilting Pad Thrust Bearings" by Dr S Glavatskih (Lulea University of Technology, Lulea, Sweden).

Dr. R S Dwyer-Joyce (The University of Sheffield, Department of Mechanical Engineering, Sheffield, UK). Is there a possibility that the thickness of PTFE on a Bearing Pad varies during bearing operation by creep or elastic / plastic deformation?

Reply by the Author Yes, it does vary during bearing operation due to resilient nature of the PTFE coating.

Dr H Xu (Glacier Vandervell Technology Centre, Rugby, UK). You stated in your remarks on the future trends of tilting pad bearings that the utilisation of micro-dimpled bearing or collar surfaces would be beneficial. It appeared that there were some contradicting views on this subject.

Reply by the Author Publications on this subject, referred to in the paper, show that introduction of textured surfaces can be beneficial for bearing operation. This beneficial effect depends on bearing operating conditions and geometry of the micro-indentations. Further research is thus required, as stated in the paper, to fully understand the nature of this phenomenon and to develop an optimisation procedure.

Dr J B Medley (University of Waterloo, Waterloo, Ontario, Canada). Can you give some details regarding the problems that Russian power

generating stations encountered that motivated their development of metal-fibre reinforced PTFE surface layers?

Reply by the Author In 1974 Kuibyshev Aviation Institute helped a hydroelectric station on the Volga River to solve a thrust bearing problem. The bearing, 4.3 meters in diameter, had two rows of pads. The problem encountered was due to high waviness, 0.4 mm, of the collar. During normal operation pulsations caused by collar waviness were compensated for by the flexible support structure of the bearing pads. However, during start-ups such waviness led to very high contact stresses, oil film rupture and bearing failure. This malfunction was avoided by the introduction of PTFE/wire mesh coating. Successful application of the PTFE composite coating led to its wide use in other hydroelectric stations not only in the former USSR but also in China.

Mr P Harper (The University of Sheffield, Department of Mechanical Engineering, Sheffield, UK) Does the wire mesh that the PTFE is mounted on act like a spring? To what extent is this utilised in Industry?

Reply by the Author Yes, the wire mesh acts like a spring. Some bearing manufactures utilise this effect while others prefer to fill the wire mesh with a solder.

Paper X(ii) "Study on the Development of a Thin-film Sensor for Measuring Oil-film Pressure in Engine Bearing and Piston-boss" by Dr Y Mihara[a], Mr K Sekine[a] and Professor T Someya[b] ((a) Musashi Institute of Technology, Tokyo, Japan (b) University of Tokyo, Tokyo, Japan).

Professor J F Booker (Cornell University, Ithaca, USA). Theoretical studies commonly assume the engine cylinder pressure and hence bearing film pressure, is perfectly periodic. In reality, of course, it is not. Could you please tell us how much cyclic variation you encountered in film pressure?

Reply by the Authors When the engine operating condition was made constant, the cyclic variation of the oil-film pressure in the engine main bearing was changed about 5% by cyclic variation of cylinder pressure.

Dr H Xu (Glacier Vandervell Technology Centre, Rugby, UK). The authors presented a very interesting piece of work. The achievement is commendable. Can the authors provide some information on the following two points?

(1) What was the life of the sensor?
(2) What was the time sensitivity of the sensor? ie. what was the engine speed for which the sensor was successfully used?

Reply by the Authors (1) and (2). The durability of the sensor was a little different by each sensor. In the case of the test diesel engine, the oil-film pressure in the engine main bearing could be measured in the condition of from 1000 rpm to 2800 rpm under no-load, half-load and full-load by one thin-film sensor. The time to take the measurement data was about 2hours.

In a past experiment, a gasoline engine with the displacement of 1600cc was used as the test engine. The maximum operating condition was 4500 rpm under full-load.

Dr I Iordanoff (LMC, INSA de Lyon, France). The thickness of the sensor is about 3-5 μm. What is the effect of this thickness when the oil film thickness is about the same order of magnitude?

Reply by the Authors The diametral clearance of the test main bearing was 0.022 to 0.045mm. In this case, the thickness of a thin-film sensor of 3 to 5 micro-meters did not influence the measured oil-film pressure. But, in the piston-pin boss, which has a diametral clearance of 7 to 10 micro-meters, the measured oil-film pressure was influenced by the circumferential shape of the piston pin after the sensor film deposited on the piston-pin surface.

Therefore, the sensor film was deposited on the piston pin surface along the circumferential shape.

Dr Q Zhou (Jaguar Cars Ltd., Engineering Centre, Coventry, UK). Does this sensor require very clean oil? Will the debris affect the reading?

Reply by the Authors When experimenting by using the thin-film sensor, new oil was always used

to prevent scratch damage to the sensor film. However, the yield of the experiment might improve though the sensor is not necessarily broken by metal debris or something like that.

If the protection film of the sensor was broken by direct contact of crank journal or by the metal debris in the lubricating oil, the measured oil-film pressure generates an abnormal out-put voltage like the electric noise.

Paper X (iii) "Optimisation of Static Behaviour and Stability Level of a Multilobe Journal Gas Bearing" by Mr G Grau, Dr I Iordanoff, Dr B Bou Saïd and Dr Y Berthier (LMC, INSA de Lyon, Lyon, France)

Dr D S Mehenny (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK). What assumptions were made about the thermal properties of the gas?

Reply by the Authors The fluid is considered as a perfect gas. Thus the Reynolds equation is modified with the use of the perfect gas relation. The flow is assumed to be isothermal. This simplifying assumption is justified because it leads to a minimum pressure field and hence a minimum load capacity . Temperature (and then viscosity) are then considered constant in the bearing.

SESSION XI
BIO-TRIBOLOGY (2) – Biological Systems

Paper XI (i) – Invited Paper "Bio-Tribological Approaches to the Lubrication of Engineering Systems" by Professor N D Spencer, Professor S S Perry, Dr S Lee, Mr M Müller., Mr S Pasche, Mr S M De Paul, Mr M Textor Mr X Yan and Mr M S Lim (Laboratory for Surface Science and Technology, ETH-Zurich, Zurich, Switzerland and Department of Chemistry, University of Houston, Texas,USA).

Professor K Ikeuchi (Kyoto University, Kyoto, Japan). You clearly showed the importance of bio-tribology in human motion and prevention of biological adhesion.

Do you think that repulsion and separation of tissues or organs are essential in making a human body?
If it is true, biotribology will be a major field in studies in the post-genome project.

Reply by the Authors We think it is clear that many internal organs depend on lubrication/anti-adhesion mechanisms for their proper functioning. In addition to obvious cases, such as joints, the lungs, for example, can only function if certain natural surfactants are present, inhibiting adhesion between the oxygen-permeable surfaces.

The understanding of the human genome is exciting, in that the molecules that are responsible for these various tribological mechanisms may finally be conclusively identified, meaning that they can be scrutinised in isolation and lessons learned that could provide us with useful aqueous lubrication technology for applications outside the medical field.

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). Your appraisal of the role of lipids in synovial joint lubrication would be of interest.

Reply by the Authors This is a topic that concerns me greatly and is currently the subject of a new project within our group. Frankly the literature in this area is very confusing, with totally different conclusions having been reached by different workers in the field. We hope we will be in a better position to answer this important question once our research is properly underway.

Paper XI (ii) " Eyelid / Lens Contact During a 100 Metres Sprint. A Tribological Approach " by Dr B Bou-Said (LMC, INSA de Lyon, Lyon, France).

Professor N Spencer (Laboratory for Surface Science and Technology, ETH-Zurich, Zurich,

Switzerland). What is the composition of tears? Is boundary lubrication involved?

Reply by the Author The tear is mainly water-like. I did not know exactly its composition. I am not sure that boundary lubrication exists because the eye is usually well lubricated

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, The University of Leeds, Leeds, LS2 9JT, U.K.) Have you analysed other situations, in addition to sprinters, where the body is vibrated (eg on vehicles or when operating machinery)?

Reply by the Author Yes and we are working on this with a laboratory more involved experimentally situated in Saint Etiennne.

Professor R Murakami (The University of Tokushima, Tokushima, Japan). Is there enough lubricant in the tear?-or does starvation occur?

Reply by the Author The question is very difficult to answer. They are people who have what we call ' *wet eyes'* and for them no lubricant starvation exists. There are other people who have ' *dry eyes* ' and for them the contact may not be fully lubricated and starvation might exist . That is why the choice of lens depends on the wettability of the person who wants to wear lenses and the views of the doctor.

Paper XI (iii) "Biological Lubrication of Hydrated Surface Layer in Small Intestine" by Mr H Yoshida, Dr Y Morita and Professor K Ikeuchi (Institute for Frontier Medical Sciences, Kyoto University, Kyoto, Japan).

Mr P Paré (University of Waterloo, Ontario, Canada). Is the water absorbed or adsorbed on the intestine surface?

Reply by the Authors In the model of hydration lubrication, the hydrophilic macromolecules are adsorbed by the internal surface of the small intestine. The water is absorbed by the surface layer because it is attracted by the hydrophilic macromolecules.

Paper XI (iv) " Tribology of a Snail (Terrestrial Gastropod) by Professor A Kobayshi, Mr I Yamamoto and Dr T Aoyama. (Meijo University, Nagoya, Japan).

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). The viscosity of the mucus appears to fall to about 40 or 50 mPa.s at shear rates of only a few hundred reciprocal seconds. Were you able to estimate the shear rate during locomotion?

How do the viscosity values compare to those of other human body fluids associated with bio-tribological functions?

Reply by the Authors The locomotion of a snail is done by pedal waves of the ventral surface. Therefore, the part of the ventral surface in which the pedal wave does not pass through contacts to the ground without slip namely statically. Consequently, the mucus sandwiched between that part of the ventral surface and the ground is also in a static state. As a result, it is considered that the viscosity of the mucus of the contact part of the ventral surface has the value of the shear rate of 0 1 /s in Fig.2.

We do not compare the viscosity value to those of other human body fluids in this study.

SESSION XII TRIBOLOGICAL STUDIES of DRY and LUBRICATED SURFACES (3) – Lubricated

Paper XII (i) Invited Paper "The Role of Tribology Research in the Development of Advanced Lubricants" by Dr A Jackson and Dr M N Webster (Exxon Mobil Research and Engineering Company, Annandale, NJ, USA).

Professor D. Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK). I enjoyed your excellent, comprehensive overview of lubricant development.

One of your conclusions was that equipment and lubricant designers need to work together. I seem to remember that this was an important finding from the D.E.S. Jost Report in the mid-1960's. Are we getting the message through? -or is no one listening? In tribology and fluid mechanics courses we should emphasise the point that the lubricant is an engineering material.

I wonder if you could give an example or two to illustrate your point that lubricant developments can '*enable*' new designs?

Reply by the Authors The authors thank Professor Dowson for his kind comments about their paper. With regard to our conclusion that equipment and lubricant designers need to work together, perhaps for many cases it would have been better stated as *continue to* work together. We believe the message in the Jost Report did get through and people certainly listened, but that was over 35 years ago. Many of the resulting gains are now taken for granted and there is a new generation of employees in both industries. There are also new lubricant technologies and new demands on equipment and a little reminder about what might be possible by working together is perhaps not out of place.

The clearest example of lubricant developments enabling engineering design is in aviation turbines, where the temperatures in the central shaft bearings exceed the capability of mineral oils. Only synthetic lubricants are used in these applications. A number of "high-end" car manufacturers are using synthetic lubricants as initial fill and are tailoring their designs, operational envelopes and maintenance procedures to the capabilities of synthetic oils. Car manuals and even filler caps indicate the need to continue to use synthetics during service. Gearbox power ratings, particularly for worm gears, have been increased through the use of synthetic lubricants. Several service manuals dictate that power ratings should be decreased if mineral oils are used due to reduced oil life from the combined effect of lower oxidative stability and higher temperatures due to higher traction coefficients.

More often than not, synthetics are used in equipment designed to work with mineral oils. While synthetics bring a number of benefits, as listed in Part 7 of our paper, the systems are not optimised. Part of the message about equipment and lubricant designers working together is aimed at facilitating system optimisation by considering the engineering possibilities of a variety of new lubricants.

Dr. J B Medley (University of Waterloo, Waterloo, Ontario, Canada). Do you think that all cars will eventually run on synthetic oil?

Reply by the Authors We think that all cars would *benefit* from the use of synthetic oils, but in the short term there is not enough synthetic base stock available to fill the sumps of all cars in current service. In the next ten years or so, a new generation of base stock, synthesised from natural gas and known as GTL (gas-to-liquids), is likely to become available and this will greatly expand the supply. Pressure for high vehicle fuel efficiency, low emissions and low maintenance costs will increase demand for high quality lubricants and in the end, these requirements are likely only to be met by formulating engine oils using base stocks such as GTL.

Paper XII (ii) "Ultrasonic Characterisation of an Interference Fit", by Dr R Lewis, Mr M B Marshall and Dr R S Dwyer-Joyce. (The University of Sheffield, Department of Mechanical Engineering, Sheffield, UK).

Dr. H Xu (Glacier Vandervell Technology Centre, Rugby, UK). The authors have presented a very interesting research project. How are they going to cope with the situation of different surface finishes?

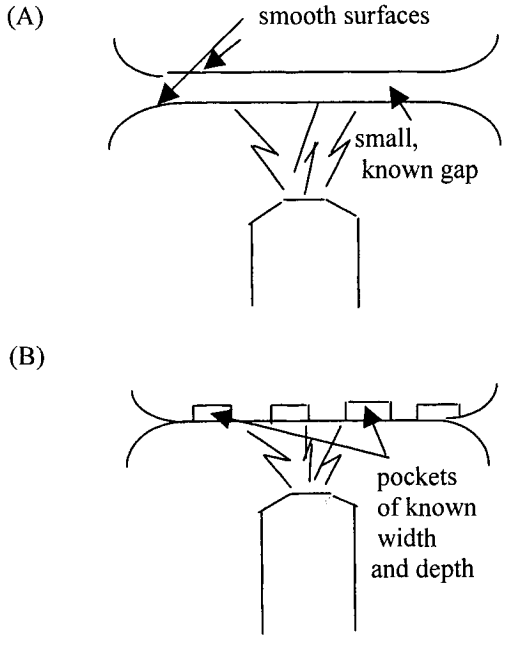
The results of lower contact pressure at the centre of interference fit, for the case of shrink fit, are very interesting. Is it possible that they were caused by the entrapment of air in the central region of contact, since both ends of the contact would close early, at tighter than average pressure. The tighter fit is likely to provide good sealing effects.

Reply by the Authors In order to handle different surface finishes we would have to perform a new calibration procedure for each pair. This is a major limitation of the approach. It is not possible to

deduce contact pressure directly from ultrasonic measurements of contact stiffness because stiffness is a function of the number, size, and distribution of asperity contact regions. Our previous work where we have tried to assume an asperity distribution and thus achieve a direct relation ship between pressure and stiffness has shown limited success. So unfortunately, we still require some form of experimental calibration.

Yes, it is possible that trapped air might be responsible for the shrink fit pressure features. The edges of the shrink fit would cool quicker and so make contact before the central region.

Dr J B Medley (University of Waterloo, Waterloo, Ontario, Canada). Have fundamental calibrations been conducted on the relationship between gap geometry and reflected signal characteristics?



Reply by the Authors Yes, some work has been carried out by Kendall and Tabor (*Proc. R. Soc. Lond.* A 323, 321-340). But only to investigate how contact stiffness varies with the size and distribution of contact regions. If we try to machine features of known size and shape on the surface (e.g. the pockets you describe) they inevitably tend to be large compared with the ultrasonic wavelength. They thus act to scatter the ultrasound and our analysis based on a low frequency regime spring model becomes invalid.

Professor A A Lubrecht (LMC, INSA de Lyon, Lyon, France). How do you distinguish between edge effects and edge stresses?

Reply by the Authors The transducer has a finite, and quite large (~1mm) spatial resolution; although within that spot their is a central weighting. As the transducer approaches the edge of the specimen a slope change is observed in the reflected signal which corresponds to the measuring interval extending past the specimen edge.

Dr C J Hooke (University of Birmingham, U.K.). You showed that there was a difference between the behaviour of shrink fit and pressed contacts. How important is it to match the actual details of the fitting process to obtain an accurate pressure-transmission calibration curve?

Reply by the Authors Results from both the press fitted and the shrink fitted components have been calibrated using the same normal loading contact. Clearly this is only strictly representative of the shrink fit case. The push fitting will induce shear stresses in the surface that may lead to a higher contact area than would be reproduced by a normally loaded calibration test. However, measurements of the roughness after joint disassembly did not show large differences. At this stage, without a more complicated calibration, we have to be satisfied with this assumption.

SESSION XIII
COMPONENT TRIBOLOGY (2)
-Bearings and Mechanisms

Paper XIII (i) "Tribologically-based Design of Precise Positioning Stage" by Dr K Adachi, Mr T. Yamaguchi, Professor K Kato, (Tohoku University, Sendai, Japan) and Mr Y Ishimine (Kyocera Co., Kokubu, Japan)

Professor A Kobayashi (Meijo University, Japan). The authors discuss the tribological

characteristics of driving parts by means of an ultra-sonic driver and a ceramic slide. However, the driven stage has other friction parts (maybe a linear guide). I would appreciate it if the authors could tell me how the other friction parts influence the tribological characteristics of the driving parts and whether they have investigated this.

Reply by the Authors Friction resistance of the linear guide for the linear stage is one of the important factors, which influence the tribological characteristics of the driving parts. At the same time, it has direct effects upon PID gains for the precise positioning of the stage. Although we have not investigated the direct effect of friction resistance of the linear guide on wear of the driving part, the wear characteristics measured in this paper can be evaluated by the proposed parameter named fluctuation of the operating voltage E_f, which is determined by total system response including many effects such as PID gains, pre-load, stage mass, stage speed, friction resistance of the linear guide and so on.

Paper XIII (ii) "Journal Bearing Oil Film Measurement Using Ultrasonic Reflection" by Mr B Harper[a], Mr B Hollingsworth[b] Dr R S Dwyer-Joyce[a] and Dr B W Drinkwater [c] ((a) Department of Mechanical Engineering, University of Sheffield. UK (b) Department of Mechanical Engineering, Brown University, Providence, Rhode Island, USA (c) Department of Mechanical Engineering, University of Bristol, UK).

Dr T Sperring (University of Wales, Swansea, UK). What is the resolution of the sensor and how does the bearing geometry affect the results?

Reply by the Authors The resolution of the system is dependent on the geometry and frequency of the transducer. A planar transducer was used in these experiments. The emitted signal is not focused and travels perpendicular to the front face of the probe. This provides a resolution of approximately 15mm. However, it is possible to use focused high frequency transducers that would provide resolution down to around 100μm. Bearing geometry will affect the refraction of the sound wave, but this only becomes significant for small geometries where there is a significant refraction effect.

Dr S Glavatskih (Luleå University of Technology, Luleå, Sweden).

1. What are the main metrological characteristics of the system (such as total uncertainty).

2. How was the calibration performed?

3. What is the effect of shear-rate on system performance?

4. What is the effect of temperature on system performance? How was the film temperature measured?

5. What is the effect of oil degradation?

Reply by the Authors The paper goes someway to describing the metrological characteristics but it is not intended as a full and thorough study. We leave this for future development.

1. Calibration was performed using two glass plates separated by a known thickness shim at one end and a known, but different, shim thickness at the other end. The change in the thickness of the separation with distance was then measured using ultrasonic reflection. The results were compared to the expected geometrical change in thickness with distance.

2. The only parameter that will affect results is a change in the acoustic properties of the lubricant layer. In our investigations so far this has not proved to be a problem. Only if the shear rate affects the density or bulk modulus significantly will this have an effect.

3. Film temperature only affects measurements due to the fact that the speed of sound in oil changes with temperature. The effect is about 20% over a temperature increase of 80K. Oil film measurements were made using a K Type thermocouple monitoring the temperature of the oil as it left the bearing. This will give an approximate guide to temperature and so further reduce and inaccuracies introduced if there were no temperature compensation.

4. The crucial factor in measuring the thickness of the oil film is in knowing the speed of sound in the oil. This can be monitored on line and over a period of time monitor the change in speed of sound caused by a degradation of the lubricant layer.

Professor J F Booker (Cornell University, Ithaca, N.Y., USA) The work reported has very exciting prospects for equipment monitoring. Less clear are its prospects for use as a research tool.

Could you please comment on the following?

1 How quickly can measurements be captured That is, what sort of dynamic phenomena can potentially be monitored?

2 Can a transducer be mounted internally on the shaft and thus be used to scan the circumference of the scan?

3 What will be the effect when cavitation is encountered?

Reply by the Authors

1. To capture a transient event the ultrasonic pulse must be bounded by that event. The duration of the pulse is typically 1.2μs (i.e. 3λ, where λ is the wave length of the sound travelling through the media) at 10MHz. Provided the pulse can be triggered then potentially phenomena of this duration can be measured. Continuous pulsing takes place with pulse separations of typically 2.0 μs. This means that dynamic phenomena of this duration can be continuously recorded without any triggering.

2. It should be possible to mount a transducer in the rotating shaft, and work has been done by others in the instrumentation field in mounting thermocouples and other monitoring equipment on rotating shafts. The only problem is getting the data out of the spinning shaft through slip rings without noise generation.

We have done some work on the effect of air in the oil. Measurements can be taken even with a high amount of small air bubbles in the oil. The only requirement to be able to measure the thickness of the film is that there is oil present, i.e. ultrasound will not be transmitted across an air gap. It is not clear to us yet whether the bubbles in a cavitated film are large enough to permit measurement.

Paper XIII (iii) "THD Lubrication of Mechanical Face Seals During Transient Period After Start-up: 2D Modelisation" by Dr B Tournerie, Dr J C Danos and Professor J Fréne, (Université de Poitiers, Laboratoire de Mécanique des Solides, France)

Dr B Bou-Said (LMC, INSA de Lyon, Lyon, France) Is the variation of (μ) with (T) correct?

Why has the convective term in the circumferential direction in the energy equation been neglected ?

Reply by the Authors 1 - The law

$$\mu(T) = \mu(T_0)\, e^{-B(T-T_0)}$$

of the viscosity variation μ with the temperature T is well adapted for oil when temperature variations are limited to some 10°C. But for variations equal or greater than 50°C it does not fit as well the actual oil viscosity variation. Other laws have to be determined. The temperature variation in our study is at the limit of these two ranges.

2 – There is no convective term in the circumferential direction in the energy equation because our model is axisymetric and the governing equations are only two dimensional.

SESSION XIV. -ENGINE TRIBOLOGY (1)

Paper XIV (i)-Invited Paper "Modelling Lubricant Related Fuel Economy in Heavy Duty Diesel Engines" by Dr R Castle and Dr S Arrowsmith (Infineum U.K.,Ltd., Abingdon, Oxon, UK).

Dr R I Taylor (Shell Global Solutions (UK), Chester, UK). Your results suggest that surface active additive packages are particularly effective at

reducing friction under transient conditions. Could the author comment on this?

Reply by the Authors The data certainly suggest that the effect of the additive package changes are more pronounced under transient conditions than steady state conditions. Since these additives are known to operate more effectively under boundary and mixed lubrication than full film lubrication conditions, the data imply that the proportion of boundary lubrication is higher under transient conditions. This might be expected since under the transient conditions, the lowest speeds in the cycle could result in more boundary like conditions.

Dr F Qureshi (The Lubrizol Corporation, Wikcliffe, OH, USA). What is the impact of using minimum base oil viscosity fluid with highest polymer loading on film thickness and component durability?

How did the three polymers used in the study differ in their shear thinning behaviour?

Reply by the Authors Minimising the Newtonian viscosity of the fluid was found to decrease film thickness under EHD rolling conditions, and could be expected to be a debit on the field for component durability. In this case, a trade off results between fuel economy and wear protection. However, the most encouraging results from this study showed that by maximising the Newtonian viscosity and adding surface active additives to the package, significant fuel economy benefits could be achieved without compromising the film thickness or wear protection. The difference in the shear thinning behaviour of the polymers is believed to account for the difference in FE observed for the formulations including the different polymers. However the shear thinning behaviour was not characterised directly for each of the polymers.

Dr M Priest (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). A customer driver for heavy duty diesel engine lubricants is undoubtedly extended oil drain intervals.

1. What are typical current targets for oil drain intervals in heavy duty diesels?
2. How does this influence formulation?
3. Do you evaluate the tribological and fuel economy performance of what we might call aged oils, characteristic of service fills after significant mileage?

Reply by the Authors Oil drain interval expectations differ significantly by geographical region. In Europe they also depend on the performance level of the lubricant, but current targets range from 30,000 to 150,000 km. Future EU requirements may go up to 250,000 km for top tier products. In the US, oil drain intervals are typically lower than in Europe and are currently at the 30,000 – 40,000 km level. The effect of these timelines on formulation is immense – the key areas of concern are oxidation, TBN depletion and soot generated wear. Consequently the formulation requires excellent antioxidancy (which usually means higher cost), high TBN (which usually means higher ash), boosted antiwear (which is a challenge given the impending limits on P level in the oil) and improved dispersancy which affects low temperature viscometrics, and can be a debit in elastomer seal compatibility. The tribological behaviour of used oils (either from engine tests or field samples) can be evaluated using most of the standard tests. Artificially ageing the oil in a laboratory environment can provide some insight into the performance of used oils, but generally does not represent the impact of engine use, particularly in the case of mechanical shear of the oil.

Paper XIV (ii) "A New Tribometer for Piston Ring Friction Measurements" by Mr P Andersson (VTT Technical Research Centre of Finland, Espoo, Finland).

Dr R Castle (Infineum UK,Ltd., Abingdon, Oxon, UK). Will you use the tribometer to study the impact of cylinder liner honing patterns on friction and wear?

Reply by the Author Yes, studies on the influence of the surface pattern of the cylinder liner are forecasted.

Dr R I Taylor (Shell Global Solutions (UK), Chester, UK) In a real piston assembly the top ring is usually 'starved' of oil due to the action of the

lower piston rings and this has a great effect on both film thickness and friction.

Are you planning to look at this by limiting the supply of oil in your rig?

Reply by the Author Yes, and for enabling controlled dosing of the lubricating oil to the test samples, oil film thickness measurements will be included in the tests.

Paper XIV (iii) "The Wear Response of Ceramic Matrix Nano-composite Coatings in Simulated Piston-ring / Cylinder Wall Test Environments" by Dr K L Dahm Mr K Panagopoulos and Dr P A Dearnley (Institute of Tribology, School of Mechanical Emgineering, The University of Leeds, Leeds, UK).

Dr. R Castle (Infineum U.K.,Ltd., Abingdon, Oxon, UK). Did you make any fracture toughness measurements of the nano-composite coatings?

You might want to look at achieving different shaped SiC particles in the coating which are more effective at crack bridging than spheres.

Reply by the Authors While microhardness indentations were made on polished cross-sections of the plasma sprayed coatings, we made no direct crack-length measurements in order to determine K_{IC}. In general the hardness indentations produced cracks that ran along the inter-lamella boundaries within the coating. This was due to the low cohesive strength of the coatings. Fracture toughness (K_{IC}) measurements will be made on the next generation of nano-composite coatings which have a much higher density.

The method used to produce the nano-composite powders does not allow direct control over the size or shape of the SiC particles. Additionally, the mechanism(s) by which fracture toughness may be enhanced in such nano-composites are not well understood (see refs. 5 and 6 for example).Therefore, crack-bridging may not be as important in increasing the fracture toughness of ceramic nano-composites as it is for micro-scale ceramic matrix composites.

Paper XIV (iv) " The Lubricant Condition in the Ring Zone Under Cold Start Conditions" by Dr. A Datoo and Dr. M F Fox (Lubricant Research Centre, De Montfort University, Leicester, U.K.)

Dr. E Tomanik (Mahle Technical Centre, S Paulo, Brazil) Do you think that it is possible to artificially reproduce the degraded oil conditions in order to obtain larger amounts of oil samples and then to do bench tests (eg wear tests)?

Reply by the Authors It would be helpful to be able to artificially reproduce larger volumes of the degraded lubricant condition. But there are difficulties concerning the extent of the underlying lubricant degradation, the extent of fuel and water dilution, the state of the fuel and the mixing of lubricant, fuel and water.
The lubricant samples are produced from the routine degradation of the lubricant as the engine test proceeds, the fuel is partially degraded within the engine to an extent which is difficult to reproduce and the lubricant, partially degraded fuel and water are mixed by a severe shearing process unique to the piston ring zone which is not well understood. Therefore, the synthesis of samples to reproduce the 'cold start' condition samples described is not meaningful.

There is a false analogy to this, that of adding carbon black to a lubricant to simulate a soot loaded lubricant, which never gives comparable samples, even for the same apparent soot analyses.

SESSION XV ELASTOHYDRODYNAMIC LUBRICATION (2) – EHL-Mixed Lubrication

Paper XV (i) "Slip at the Wall-Evidence and Tribological Implications" by Professor H A Spikes, (Imperial College, Tribology Section, London, UK).

Dr H Xu (Glacier Vandervell Technology Centre, Rugby, UK). The subject of a "half-wetted bearing" is very interesting. It is fascinating to be

told that such a system can have much reduced frictional force while still maintaining a significant hydrodynamic capacity (a half of the normal non-slip lubrication capacity). If this is the case, does that mean that we can develop a product with virtually zero frictional force if a lyophobic surface with $\tau_c = 0$ is used?

Reply by the author In principle yes. But in practice it is too early to say. Firstly, the type of slip discussed in my paper requires very smooth surfaces, of the order of root mean square roughness one nanometer. This is relatively easy to achieve in very small components by chemical-mechanical polishing, but is currently not practical for macro-scale surfaces. That said, it has also been shown that very rough lyophobic surfaces can also promote slip, probably by entrapping air within the surface cavities, and thereby reduce energy losses during pipe and other fluid flow (1)

Second, we still have very little quantitative information about this form of slip. It is really only in the last three years that physicists and engineers have started to work seriously on the phenomenon. There is currently a very large amount of research taking place since the potential applications are enormous – e.g. fluid pumping, windage losses, microfluidic devices, filtration as well as lubrication – but our understanding is still limited. Thus we do not yet know how low a value of τ_c can be reliably obtained and maintained nor, even, what is the fundamental mechanism of slip.

But it is certainly a very exciting field, with many possibilities. I am quite sure that slip at the wall as described in my paper will be applied in some devices, if not yet in bearing technology, by the end of the decade. Currently most work in the field is taking place in Japan and the US. Let us hope that some of this application originates from the UK.

[1] Watanabe, K, et al., "Drag reduction of Newtonian fluid in a circular pipe with a highly water-repellent wall." J. Fluid Mech. (1999) 381, pp. 225-238.

Paper XV (ii) "Calculation of EHL Contacts in Mixed Lubrication Regime" by Dr A C Redlich, Dr D Bartel and Professor L Deters. (Otto-von-Guericke-Universitaet, Magdeburg Institut fur Maschinenkonstruktion Magdeburg, Germany).

Dr J A Williams (The University of Cambridge, Engineering Department, Cambridge, UK). Plastic deformation of asperities makes a crucial contribution to friction within your islands of 'real' contact. Most tribological contacts are loaded many times so that the initial topography 'shakes down' or 'runs-in' to essentially elastic conditions. Under such conditions how do you then estimate interface stresses or energy losses at the points of dry contact?

Reply by the Authors In the case of pure elastic conditions we assume that a certain part of the elastic deformation – the hysteresis – gets lost and therefore must be taken into account for friction calculation. The authors know that this assumption is not sufficient because additional mechanisms must be taken into consideration as well. The most important point for future work will be the temperature calculation at these contact spots, to check whether the molecular thin film between the asperities still exists or not. If there is no more oil film available in the micro contact area, adhesive forces are assumed between the two surfaces. The authors believe that the temperature calculation will play a decisive role in the friction calculation.

Dr. F Ville (LMC, Insa de Lyon, Lyon, France) What is the influence of the mesh size on the results? The pressure field and the film thickness seem to be due to numerical features. So what is the proportion between the numerical precision and the real results?

Reply by the Authors This is an important question and we have studied this point carefully. You must notice that the new pressure p(num) depends on the mesh size but one can show, that on coarse grids the sum of both p+p(num) is a good approximation to p on finer grids (the real result!), where p(num) is no longer necessary because no negative film thicknesses occur.

Paper XV (iv) "On the EHL Film Breakdown" by Dr D Zhu, (Eaton Corporation, Southfield, MI, /USA).

Dr H P Evans (Cardiff School of Engineering, Cardiff University Cardiff, UK). Have you calculated flow vectors within the contact area?- because the liquid entrained into the inlet part of the Hertz area must principally pass through the contact in the entrainment direction.

Reply by the Author The current model considers that when the film thickness is sufficiently small, the left-hand side pressure driven flows in the Reynolds equation become negligible, because they are related to the third power of the film thickness. It will yield a zero Couette flow in steady-state solutions. When this happens, the majority of the fluid flows around the contact. There may still be some lubricant passing through the very thin gap under the no-slip boundary condition. However, such flow through the nearly zero gap becomes trivial.

It should be mentioned that this treatment for the terms that vanish under certain conditions is proper. It is similar to the treatment of neglecting the mass of the fluid when solving fluid dynamic problems for low-Reynolds number flows, where mass exists, but just too small to have any significant influence.

Dr C J Hooke (School of Engineering, University of Birmingham, UK). What mesh size was used in the analysis? Did increasing the number of nodes affect the results?

Reply by the Author The mesh sizes are $\Delta X=0.0117$ and $\Delta Y=0.0125$, which are sufficiently small for the current smooth point contact EHL problems. These sizes are about the same as, or smaller than, those used in some previous studies by Venner et al (1992, 1994), Chevalier and Lubrecht (1998), Ai and Cheng (1996), Xu and Sadeghi (1996), and others. We have tried to reduce the mesh size by half and obtained nearly the same results, although the computing time was greatly increased.

Dr H Xu (Glacier Vandervell Technology Centre, Rugby, UK). The author has made a great effort to carry out analysis of EHL contact for wide ranges of speed and load parameters. Can the author provide an explanation of the whereabouts of the lubricant at the time of oil film breakdown? From the view point of shear flow in the centre of the lubricated conjunction, a certain amount of lubricant is dragged along in the direction of surface motion. When the oil film drops from a finite value to zero (breakdown), the condition for the oil flow to be zero in the direction of surface motion is an infinite pressure gradient. Can the author comment on this?

Reply by the Author By the term of "EHL film breakdown", the author meant to explore the mechanical parameters that may reduce the hydrodynamic influence and, as a result, reduce the EHL film thickness to a practical zero at which the other films (such as boundary film) may become dominant in supporting the load and protecting the surfaces. The "practical zero" for an EHL film can be a sufficiently small value at which the hydrodynamic action becomes negligible. It may not necessarily mean an absolute zero gap. Even though the film thickness could be zero, it does not necessarily cause an infinite pressure gradient, as in 3-D point contacts the fluid will always find its way to go around the contact. This is different from the situation in 2-D line contacts.

The current model does not invoke an infinite pressure gradient at the solid-fluid interface because the shear flow into the solid contact (if any) is negligibly small. It simply indicates that at the extremely small gap of a contact spot, saying at the nano scale, both pressure driven and shear flows vanish. No assumption is made about the pressure gradient in the model. Mathematically, the Reynolds equation is balanced both sides at a solid contact spot because on the left-hand side the terms with the third power of film thickness vanish and on the right-hand side the gap gradient vanishes, assuming steady state. Physically, most of the fluid is repelled to flow around a contact spot, and within a control volume the flow continuity is satisfied.

Professor H A Spikes (Imperial College, Tribology Section, London, UK). Your computed results predict an enhanced reduction of film thickness at very low U* values. This has not been predicted in previous numerical work. Nor has it been found in experimental work over the same U* range.

Previous experimental work on base fluids (without boundary film-forming additives) shows linear log

(film thickness) versus log (speed) behaviour within the speed range of interest-eg. Guangteng & Spikes, WEAR 200 (1996), p 336, Figure 4.

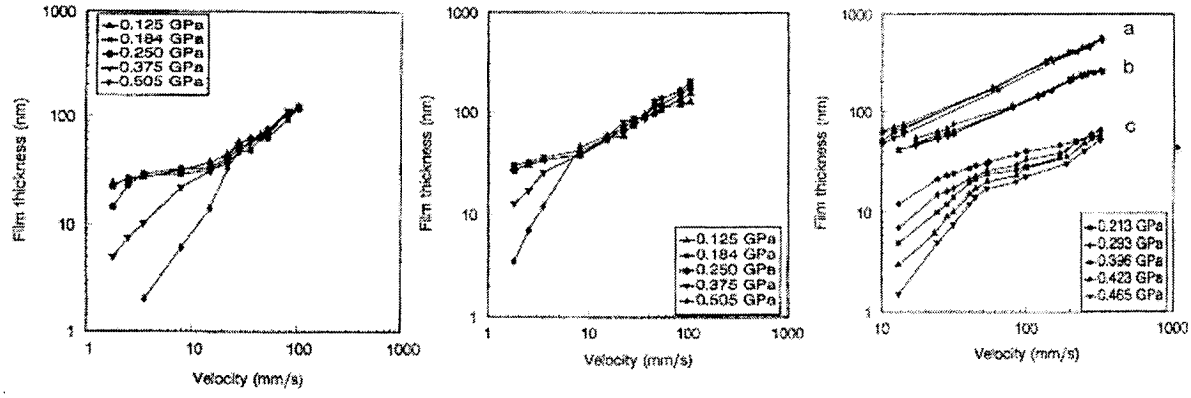
Could you comment on this discrepancy between experiment and your numerical predictions?

Reply by the Author The author had read the paper referenced by Professor Spikes and noticed that the current results reasonably agree with many of the measurement data for base stocks. Based on the Young's modulus of 73 GPa for glass (neglect the coating effect, although the chromium coating has a different Young's modulus), 220 GPa for 52100 steel, 0.019 m of radius for the steel ball, the author found that the maximum Hertzian pressure due to the load of 20 N in this paper is about 0.33 GPa.

To the speed range down to U* ~10^{-14}, the film thickness down to Hc ~ 5.3 10^{-8} (1 nm), and load W* up to ~ 10^{-6}, the model used in the present paper does not predict accelerated EHL film reduction. Actually, corresponding to this maximum Hertzian pressure at the given fixed load, the current model well agrees with the Hamrock-Dowson formula that was used by Dr. Gao and Professor Spikes in their paper. At this Hertzian pressure, the model predicts a near straight-line relationship between the film thickness and U* in a double logarithm scale (see Figure 1 in the present paper), almost the same as the straight-line part for the base oils, especially in Figure 2, in the paper mentioned by Professor Spikes. In these ranges of data, the current model slightly deviates from the Hamrock-Dowson formula at Hc ~ 10^{-7} (about 1 nm or 2 nm), where the measured data exhibit notable scattering (Figure 2-4 in the paper mentioned by Professor Spikes). There, the rough surface asperity effect was neglected, although the ball surface roughness is much larger than the measured film thickness.

The author has realized that in the paper referenced by Prof. Spikes, no film collapse is observed as the speed continuously decreases. Instead, the film thickness reduction is, in many cases, slower than that predicted by the Hamrock-Dowson formula, and the curve is often levelled when the film becomes very thin, in the range of several nano-meters or a few tens of nano-meters. This appears to be a phenomenon caused by the boundary film, which is the focal point of the referred paper, instead of the EHL film. Moreover, base fluids without additives do not necessarily mean that there is no boundary film. Actually, the measured film in the order of a few nanometers can most unlikely be the hydrodynamic film. However, what the present paper is trying to discuss is still the hydrodynamic film that can be modeled and predicted numerically with continuum mechanics. The breakdown discussed in the paper is also strictly for the EHL film.

The author is aware of the fact that not all experimental results previously obtained/published demonstrate the straight-line relationship. The figure given below, for example, shows the work published in 1999 by Luo and Wen et al [1], where the slope of the curves for the film thickness



vs. speed is not constant. The heavier the load is, the greater the negative slope can be observed. There could be a quick decline as the speed decreases, especially when the load is heavy. This is qualitatively in good agreement with the conclusions in the present paper. Such results The speed, for example, typically covered a couple of orders of magnitude or less. The load range for most optical interferometric measurements, similarly, is also relatively narrow and the maximum Hertzian pressure is still relatively light.

However, the present paper extends the speed range to cover nearly 10 orders of magnitude, and the maximum Hertzian pressure reaches 6-7 GPa. The conclusions drawn in the present paper are based on the extended ranges. Further efforts are certainly needed for satisfactory experimental verification.

Additionjal References

[1] Jianbin Luo, Linmao Qian, Shan Wen, Liu Wen and Shizhu Wen, "The Failure of Fluid Film at Nano-Scale," Tribology Transactions, Vol. 42, (1999), pp.912-916.

[2] Koye, K.A., and Winer, W.O., "An Experimental Evaluation of the Hamrock and Dowson Minimum Film Thickness Equations for Fully Flooded EHD Point Contacts," ASME Journal of Lubrication Technology, Vol.103, 1981, pp.284-294.

[3] Greewood, "A re-examnination of EHL film-thickness measurements," Wear, Vol. 15, (1970), 281-289.

[4] Coy, J.J., Zaretsky, E.V., "Some Limitations in Applying Classical EHD Film Thickness Formulas to a High Speed Bearing," ASME Journal of Lubrication Technology, 1981,Vol. 103, pp.295-304

SESSION XVI COMPONENT TRIBOLOGY (2) – Transmissions

showing a non-straight line relationship can also be found from other experiments [2,3,4].

It is also important to note that most experimental results obtained in the past were within relatively narrow parameter ranges under different conditions.

Paper XVI (i) Invited Paper "Gearing Tribology" by Professor R W Snidle, Dr H P Evans and Mr M P Alanou (Cardiff School of Engineering, Cardiff University Cardiff, UK).

Dr J B Medley (University of Waterloo, Waterloo, Ontario, Canada). In metal-on-metal hip implants, smooth, super-finished surfaces operate under starting (λ) values of about 1.5 and are subject to mild wear followed by abrasive damage due to exposure of carbides and varying amounts of micro-pitting. Thus, the micro-ehl modelling described in your paper may be very applicable and I would like to ask the following question.

If superfinished surfaces are placed under severe conditions such that (λ) < 3, does a micro-ehl analysis show any fundamental differences in terms of "co-operation" vs "independence" of asperity contacts and the incidence of cavitation within the contact? Do you have any advice regarding the possibility of applying your analysis to metal-on-metal hip implants?

Reply by the Authors We have not examined extreme cases where superfinished surfaces are in micro asperity contact such that (λ) < 3. In gear situations superfinished surfaces generate what are effectively full films with the corresponding benefits described in the paper.

The (λ) < 3 conditions we have analysed are for ground surfaces with a finish that is essentially transverse to the entrainment and sliding direction. The overall entrainment action draws oil into the contact which is typically pressurised in valley features of the surface. Asperities have their own individual film forming action but are not independent of the main EHL entrainment action in that the pressurised oil surrounding them has enhanced viscosity.
In principle there is no impediment to applying the solution method to metal on metal hip joint configurations, but clearly the properties of the lubricant differs considerably from conventional

engineering lubricants. Other differences arise from the much higher geometric conformity and the much lower pressures generated in such contacts.

Dr. R I Taylor (Shell Global Solutions (U.K.), Chester, UK). Have you compared your micro-ehl calculations with simpler dry contact models where the lubricant effectively acts as a "spacer layer" and each asperity is treated by Hertz theory?

Are there differences between the two approaches?

Reply by the Authors Yes. There are significant and substantial differences as discussed by the authors in a previous contribution to the 25th Symposium (pp 163-174) where line contact between smooth and rough surfaces were considered. In a dry contact analysis load is carried solely on the contacting asperity tips. With rough surfaces this leads to a relatively high number of asperity contacts which develop very high pressures over very limited areas. The introduction of a lubricant changes the situation significantly. The number of asperity contacts where high pressures are developed is considerably reduced, the elevated pressures developed are lower and act over larger areas. The load support contributed by pressurised oil in the valley features also contributes significantly. Overall, the effect of the lubricant is to reduce the severity of the pressure peaks and the consequent sub-surface shear stresses.

Paper XVI (ii) Invited Paper "Tribology of Spline Couplings" by Dr A V Olver and Mr S Medina (Tribology Section, Imperial College of Science Technology and Medicine, London, UK)

Mr T Hyde (Rolls-Royce, Derby, UK) Has the effect of friction coefficient been considered?

Reply by the Authors Some details on the effect of friction on the slip path are included in the paper (Figure 9). For more details the discusser is referred to our earlier paper [7]. Appreciable regions of stick
are encountered at coefficients of friction of above about 0.7 (for $E^*=3.3$, $b^*=11$, $z=18$, $\theta=0.1°$).

Dr H P Evans (Cardiff School of Engineering, Cardiff University, Cardiff, UK). Is the contact analysis carried out with a commercial Boundary Element Method package?

Reply by the Author Yes. Details are given in the paper by Adey, Baynham and Taylor [13].

Mr. T Hyde (Rolls-Royce, Derby, UK). Have comparisons been made between predicted slip and experimentally observed slip?

Reply by the Authors No but we have correlated our predictions of bending moment with experimental measurements using the three tooth spline rig described in [3]. A difficulty is that the coefficient of friction is unknown and often appears to vary during extended testing.

We have also compared wear patterns on spline teeth with the spatial distribution of the wear depth parameter. Qualitative agreement is obtained with a plausible value of the coefficient of friction. This is reported in [9]. A further problem is that quantitative experimental wear depths are difficult to obtain because of the local effects of debris, which cannot of course be modelled by the present two-body approach.

Paper XVI (iii) "Prediction of the Friction Coefficient Between Gear Teeth Using a Mixed Film Lubrication Model" by Dr J Castro[a], Dr A Sottomayor[a], Dr A Campos[a] and Professor J Seabra[b] ((a) Departamento de Engenharia Mecânica, Instituto Politécnico do Porto, Portugal, (b) Departamento de Engenharia Mecânica e Gestão, Universidade do Porto, Portugal)

Dr. A V Olver (Tribology Section, Imperial College of Science Technology and Medicine, London, UK). Why is the coefficient of friction very much higher near the pitch line of the gear?

Reply by the Authors The authors would like to thank Dr A. V. Olver for his question about this work.

There are two major factors contributing to the high values of the friction coefficient near the pitch point in gears.

One is the low sliding speed that occurs in that zone of the gear meshing line, implying reduced heat dissipation, lower lubricant temperature and higher lubricant viscosity. Thus, the tangential stresses

remain at a similar level because the lower slide-to-roll ratio is compensated by the higher viscosity.

The other factor is that the local contact pressure is higher near to the pitch point (in this type of gears), because the contacting load is supported by a single pair of teeth, that naturally contributes to a higher friction coefficient.

Only at the pitch point (or very close to it), where $Ve \approx 0$, the tangential shear stresses and the global friction coefficient are very small, but their influence on the average friction coefficient over the meshing line are very limited.

Paper XVI (iv) "Influence of Temperature Distribution on the Tribological Performance of Automotive Synchronisers" by Professor Dr-Ing. G Poll and Dr-Ing. M Spreckels (University of Hannover, Hannover, Germany).

Dr H P Evans (Cardiff School of Engineering, Cardiff University Cardiff, UK). Can you comment on the factors influencing the critical temperature of 150°C that you describe?

Reply by the Authors Most probably, the transitions in tribological behaviour can be attributed to changes in the chemical and physical properties of the lubricant itself in the contact region and of boundary layers formed by adsorption as well as chemical reactions. As to the lubricant, there is certainly an influence of temperature on viscosity which may lead to a breakdown of partial micro-EHD-films when temperature rises. As to adsorbed layers, they will detach from the surface at a critical temperature. For reaction layers formed by additive-surface-interaction, it is known that threshold temperatures do exist.

The exact chain of interactions that lead to the observed transitions could not yet be established and is subject of further research at the Technical University of Munich in parallel to our phenomenological studies.

SESSION XVII
FRICTION, WEAR AND LIFE OF CONTACTS IN MACHINE ELEMENTS (2)-
Wear

Paper XVII (i) "Feasibility of Using Electrostatic Monitoring for Oil Lubricated Ceramic to Steel Sliding Contacts" by Dr L Wang, Dr R J K Wood, Dr T J Harvey, Mr S Morris, Dr H E G Powrie and Mr I Care (University of Southampton, School of Engineering, Southampton, UK).

Dr S Mischler (EPTF-STI-IMX-LMCH, MX-C Ecublens, Switzerland) Did you observe discharge effects due to the electrical conductivity of oil, which is a function of temperature?

Reply by the Authors There are four possible mechanisms which could lead to charge generation as a consequence of frictional work and/or wear [1]. Only one of these mechanisms, tribocharging, has been investigated regarding the influence of temperature and discharge effects. A linear relationship between the tribocharge and temperature was found between 25 and 80 °C for a mineral base oil [2]. The tribocharge level decreased when the temperature went above 80 °C and discharge effects become apparent. Therefore, there will be an influence on the charge levels, as a result of oil temperature changes. The influence of discharge effects is probably not that apparent in the current tests because there is not a significant variation in the oil temperature which was between 50 °C to 75 °C. However, in realistic monitoring applications, signal analysis methods have been developed which do not depend solely on the overall magnitude of the charge signal [4].

Professor T H C Childs (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). I wonder what fraction of wear debris you feel are charged? Is this a condition monitoring process that needs calibration in the exact conditions of its application for its interpretation to be quantitative?

Reply by the Authors Practical experiences suggests that the majority of debris acquires some detectable

level of charge when it is produced. The pivotal factor is, how much charge is retained by the debris once it is transported from the wear site? For all types of debris, there are mechanisms by which charge can be diluted (oil conductivity may cause charge leakage) and for metallic debris, charge may be lost through contact with neighbouring earthed surfaces. The latter happens when the debris and oil flow conditions are such that the debris is not suspended in the flow. However debris, once generated, can also acquire charge through tribocharging. Practical experience has shown that it is possible to monitor charged metallic debris at significant distances downstream of the wear-site [3]. At this remote location, any debris charge loss or gain is sufficiently balanced for monitoring to be effective without the need for system calibration. It should again be noted that signal analysis methods have been developed which do not rely on a fixed relationship between the overall magnitude of the charge signal and the size of the debris.

Dr J Sugimura (Kyushu University, Department of Mechanical Engineering, Fukuoka, Japan). You mentioned that the signal from the electrostatic sensor is not absolute but relative. Does that mean that you take the differential of the output signal, which I think is helpful in detecting the occurrence of wear?

Reply by the Authors The electrostatic sensor is designed to detect changes in charge levels. The reasons for the electrostatic sensor system measuring relative charge levels are: The signal conditioning does not have a d.c. level response, (it measures down to 1Hz) and so the system is insensitive to d.c. or background charge levels. (Sensitivity to background charge levels would preclude simple measurement of discrete particle charges). Also, for practical purposes, the wear-site sensor is a button sensor and therefore it only detects a fraction of the debris charge (the amount of charge detected being related to QA/x^2, where Q is the particle charge, A is the sensor area and x is the distance between the charged particle and the sensor face [5]). Thus for this system, the size of the detected charge does not necessarily relate to the size or quantity of the debris produced and so non-amplitude dependent analysis methods have been developed to determine this information, as already discussed.

Dr. Karl Dahm (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). In the friction and wear illustration given in the synopsis and presentation-there does not seem to be any correlation between the slope of the wear curve and the specific wear rate curve. Could the authors explain the calculation of the specific wear rate?

Reply by the Authors The paper shows five on-line monitoring results namely; temperature, coefficient of friction, electrostatic sensor response, laser displacement probe response and specific wear rates. The wear displacement (wear curve) responses are from a laser probe that monitored the disc wear track.

However, a LVDT probe was also installed on the rig that monitors the wear of both pin and disc. The specific wear rates shown in the paper are calculated according to the LVDT readings, i.e. the total loss of both pin and disc. Hence, these two graphs show separate information.

References:

[1] Morris S ., Wood R J K , Harvey T J, Powrie H E G, Use of electrostatic charge monitoring for early detection of adhesive wear in oil lubricated contacts, Transaction of the ASME, Vol. 124, 2002, p288-296

[2] Harvey T J, Wood R J K , Denuault G, Powrie H E G , Effect of oil quality on electrostatic charge generation and transport, Journal of Electrostatic, 55 (2002) p1-23

[3] Powrie, H E G, Fisher, C E, Tasbaz, O D and Wood, R J K, Performance of an electrostatic oil monitoring system during an FZG gear scuffing test, International Conference on Condition Monitoring, University of Wales Swansea, 12-15th April 1999, p155-174

[4] Powrie, H E G Wood, R J K. Harvey, T J, & Morris, S, Re-analysis of electrostatic wear-site sensor data from FZG gear scuffing tests, International Conference on Condition Monitoring, 25th-27th June 2001 also Condition Monitor, ISSN 0268/8050, Sept 2001 6-12

Powrie, H E G, Wood, R J K, Harvey, T J, Wang, L, & Morris, S, Electrostatic Charge Generation Associated with Machinery Component Deterioaration, 2002 IEEE Aerospace Conference Proceedings, ISBN 0-7803-7232-8

Paper XVII (ii) "High Temperature Abrasion Behaviour of Al/SiCp Composites" by Mr H. Ahlatci[a], Professor E Candan[b] and Professor H

Cimenoglu[a] ((a) Istanbul Technical University, Department of Metallurgy and Materials Engineering, Istanbul, Turkey (b) Zonguldak Karaelmas University, Technical Education Faculty, Turkey).

Dr Karl Dahm (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). What environment (gases, humidity) were the wear tests performed in and were any chemical effects (oxidation etc.) observed?

Reply by the Authors Wear tests were carried out at various temperatures up to 500 C under normal atmospheric conditions. The relative humidity of the environment was about 50 %. During SEM examinations no chemical effects, such as oxidation, were observed on the worn surfaces.

Professor T H C Childs (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). I wonder if you have reduced the wear observations to wear coefficient form,K, where

$$\text{Wear Volume} = KWs/H$$

I wonder how much your observations are in line with well accepted abrasive wear theory, where K may be 10^{-2} to 10^{-4} if debris are mechanically created like chips and so on.

Reply by the Authors We were not able to measure high temperature hardness (bulk or matrix) of the examined composite. Therefore we did not make any attempt to determine the K value. Abrasion tests were carried out at different temperatures under a constant test load on abrasive grains having a certain grain size. These testing conditions gave appropriate results to compare the effect of test temperature and composition on the abrasion resistance of the examined composites.

Dr S Mischler (EPTF-STI-IMX-LMCH, MX-C Ecublens, Switzerland) Did you observe the formation of an interfacial Mg=Si phase? What is the formation mechanism of the Mg-Si intermetallic phase?

Reply by the Authors EDS analysis showed that, in the matrix of Mg alloyed composites eutectic Si and Mg_2Si intermetallic are present. SiC reacts with Al at temperatures beyond 650 °C and form Al_4C_3 and Si, which react with Mg to form Mg_2Si intermetallics. In pure aluminium matrix composites we observed Si particles but not Mg_2Si.intermetallics.

Dr. P A Dearnley (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). Please comment on the effect of reinforcement particle size of the composite on abrasive wear rates.

Reply by the Authors The effect of reinforcement particle size on the abrasion rates of the examined composites is related with the size of the abrasive grains being rubbed. On fine abrasive grains, wear rate of the composites decreased with increasing reinforcement particle size. Contrarily, on coarse abrasive grains the wear rate of the composites decreased with decreasing reinforcement particle size.

Paper XVII (iii) "Friction and Wear of Spheroidal Graphite Cast Iron Under Severe Sliding Conditions" by Mr K Hirasata, Professor K Hayashi and Mr H Matsunami (Osaka Sangyo University, Department of Mechanical Engineering, Osaka, Japan)

Professor H Cimenoglu (Istanbul Technical University, Department of Metallurgy and Materials Engineering, Istanbul, Turkey). Why did you use a mild steel as a counterface which is softer than the spheroidal cast iron whose wear behaviour you studied ?

If you use a counterface material which is harder than the spheroidal graphite cast iron, the equations between μ and V will change. What changes do you expect in your derived equations?

Reply by the Authors The reason why the authors used a mild steel as a counterface comes from the point of practical application. The authors aim at a emergency brake of high speed elevator as a practical application. In this application, a mild steel is usually used as a counterface to a brake shoe. Therefore, it is important to choose a brake shoe material fitting the purpose. The authors have chosen a spheroidal graphite cast iron as an applicant for a brake shoe material here.

Meanwhile, the authors think of the changes in the equations between l and V in the case using the harder counterface material as followed. In the present case, the wear rate of the cast iron pin in length, l, in each region is given by Eq.(8). In the first wear region in which the hardness of the cast

iron pin is higher than that of the disk surface, l increases with the increment of V, but it goes contrary in the third wear region where the relation of hardnesses of the pin and the disk is reversed. From these results, the authors expect that, if a counterface material harder than a spherical graphite cast iron is used, l may be almost indepent of V for a while after the start of sliding or decreases with the increment of V from the early stage of sliding.

Paper XVII (iv)) "Wear Particles Produced in Repeated Contact of Two Asperities" by Dr J Sugimura[a], Dr W Kang[b], Ms S Tanaka[a] and Professor Y Yamamoto[a] ((a) Department of Mechanical Engineering, Kyushu University, Fukuoka, Japan. (b) Tsukuba Technical Laboratory, NOK Corporation, Ibaraki, Japan).

Dr. J A Williams (Cambridge University, Engineering Department, Cambridge, UK). I wonder if you have done any modelling of the repeated sliding situation which might lead to you being able to make predictions of the form of the wear maps showing the conditions giving rise to each form of the wear particles-and if so how these compare with some of the models of similar situations in the literature.

Reply by the Authors No, we have not done such modelling yet. We are in the stage of trying to observe how surface geometries and contact conditions change with repeated sliding for future development of some theoretical model which enables us to predict types and the amount of wear as you suggest.

Dr R J W Wood (University of Southampton, School of Engineering, Southampton, UK). Is there an influence of contact friction (ie diamond to steel) on chip formation or wear rate?

Reply by the Authors We think, and existing theoretical models tell, that there is some influence. We have started measuring friction on the rig, but we have not come to know how wear is changed by friction.

Dr P A Dearnley (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK).
(1) How were the wear particles collected and what were the smallest particles that could be found?
(2) Are the amounts of fine particles (< 1 μm) significant?

Reply by the Authors We collected the particles by rinsing the specimens after each test with heptane, and filtered the heptane containing oil and the particles with membrane filters of pore size of 0.45 μm. Considering the optical resolution of the microscope, the smallest size detectable may be half of a micrometer. In the present tests, particles of larger representative diameters than several micrometers dominate total wear amount, and submicron particles should not be significant.

Dr. Karl Dahm (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). Do the authors have any results regarding the change in particle type as the wear tests progress?

Reply by the Authors We have results showing that the wear rate changes as tests progress, but we do not have clear data which show changes in particle type, because our tests so far have been conducted for small number of contact cycles, and also because it is difficult to collect and detect particles at shorter intervals.

SESSION XVIII ELASTODRODYNAMIC LUBRICATION (3) – Experimental and Theoretical

Paper XVIII (i) "In-situ Pressure Measurements in an Elastohydrodynamically Lubricated Point Contact Using Raman Microspectroscopy. Comparison with Numerical Calculations" by Dr I Jubault[a], Dr J L Mansot[a], Dr P Vergne[b], Professor A A Lubrecht[b] and Dr J Molimard[c] ((a) Groupe de Technologie des Surfaces et Interfaces, Université des Antilles et de la Guyane, Guadeloupe (b) LMC, Insa de Lyon, Lyon, France (c) Centre SMS, Ecole des Mines de Saint Etienne, Saint Etienne, France)

Professor R C Coy (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). What is the energy input from the Raman beam?

Is this sufficiently low to cause no appreciable heating in the lubricant?

Reply by the Authors The argon-ion laser input energy is approximately 15 mW in the contact. This low value together with the continuous contact feeding allows us to neglect heating caused by the laser beam. This point is confirmed when considering the very good agreement found between some experimental results and numerical isothermal simulations, see for instance figure 4 when $u_e \leq 0.62$ m/s.

Dr A Jackson (Exxon Mobil R & E Co., Annandale, NJ, USA). What is the effect of the thermal conductivity of the contacting materials on the temperature rise in the inlet due to shear heating? Could you use different materials with different thermal conductivity to validate your results?

Reply by the Authors Even if heating occurs within the lubricant flow in the convergent region, the thermal properties of the solids influence energy dissipation. Thus, for given contact conditions, the thermal behaviour of the contact results from an equilibrium involving both the lubricant and the solids properties.

It appears almost impossible to run similar experiments involving the same techniques with different materials. Raman spectroscopy imposes a transparent specimen and the contact conditions require materials with high elastic properties. The conjunction of these two requirements leads inevitably to sapphire selection.

Finally, it should be mentioned that inlet shear heating has been numerically simulated by modifying the lubricant rheological properties only.

Dr M J A Holmes (Cardiff School of Engineering, Cardiff University, Cardiff, UK). Have you considered many rheological models in the numerical analysis?

Reply by the Authors No, we only consider a Newtonian piezo-viscous model in the numerical approach. This is justified because this work is focussed on pressure and film thickness determinations, mainly determined by phenomena occurring in the contact inlet (low shear stress region).

Dr S Oldfield (Uniquena Lubricants, Wilton Centre, Redcar, UK).

(1). What problems are seen with fluorescence from the lubricant or the sapphire?

(2) How do you calibrate pressure and temperature shift of the Raman band? Does the pressure change affect the temperature shift behaviour of the Raman line of the lubricant and can they be de-coupled?

Reply by the Authors

(1) The 5P4E synthetic nature guarantees a low level of fluorescent impurities. In the same way, the high quality of the sapphire used in this work limits the fluorescence background from the disk.

(2) We use a previously published calibration of 5P4E with pressure, reported by Laplant et al. [24]. Concerning temperature influence, we showed [18] that the frequency of the 1001 cm^{-1} Raman band was not temperature dependent, in the range 20-100°C.

Paper XVIII (ii) "Numerical Modelling of Thermal Effects in Elastohydrodynamic Lubrication Solvers" by Dr R Fairlie[a], Dr. C E Goodyer[a], Professor M Berzins[a] and Professor L E Scales[b] ((a) Computational PDE's Unit, School of Computing, University of Leeds, U.K. (b) Shell Global Solutions, Chester, UK).

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK). Specification of the mean fluid temperature may be particularly critical in the inlet region, where the minimum film thickness in the elastohydrodynamic contact is determined. Could the authors please comment on the relationship between the mean film temperature and the cross-film temperature distribution in the region of reverse flow in the inlet region?

Reply by the Authors We make the assumption that the parabolic temperature approximation is adequate in regions where reverse (Poisseuille dominated) flow takes place, even though it is strictly only true for planar Couette flow. To do anything better would also mean abandoning the mean cross-film viscosity assumption, which would make the modelling much more complex. The mean film temperature therefore bears the same relationship to the contact surface temperatures as elsewhere. Compression effects and increasing shear in the inlet region lead to a small rise

in the mean film temperature there, which in turn modifies the film thickness slightly. We assume that any errors in our assumption of a parabolic profile do not greatly influence this. Shear is introduced into the energy equation via the absolute difference in the surface velocities divided by the local film thickness. Although this is the form of the shear rate for Couette dominated flow, it is also in fact the mean cross-film shear rate for any combination of Couette and Poisseuille terms in the fluid velocity (under the mean cross-film viscosity assumption). As such, we use it everywhere, and it is consistent with the use of cross-film averaged quantities elsewhere. Many approximations are involved throughout, but we hope that by using these in a consistent way, we have a model that can be parametrically matched to experimental measurements and subsequently used in a predictive capacity. Currently accumulating experience bears this out.

Dr. M J A Holmes (Cardiff School of Engineering, Cardiff University, Cardiff, UK).

(1) Is the parabolic variation of temperature accurate enough?

(2) Are problems associated with the pressure spike present in iso-thermal solutions?

Reply by the Authors The parabolic variation of temperature is widely used in EHL modelling to reduce the dimensionality of the energy equation, without which the computation would be much more complex. The authors acknowledge that this is an approximation, but it appears to be adopted as best practice by a large number of workers in the field.

The presence of the pressure spike causes problems in the computation of the thermal solution in the cases where the pressure spike is pronounced and the mesh is particularly fine. It is worth noting that in the present work, no problems occurred. The reasons problems can occur is that, at an intermediate stage in the computation, extreme pressure gradients result in large temperature spikes, which are unphysical. These high temperatures then cause problems in the viscosity model. In the iso-thermal case, as there is no temperature variation, no such problems occur. Indeed the authors have successfully modelled the iso-thermal line contact case with in excess of 1 million mesh points, a resolution which is beyond the limit where matter behaves like a continuum.

SESSION XIX – TRIBOLOGICAL STUDIES of DRY and LUBRICATED SURFACES (3) – Lubricated

Paper XIX (i) Invited Paper "A Design Tool for Selection and Optimisation of Surface Finish in Mixed EHD Lubrication" by Dr D Zhu (Eaton Corporation, Southfield, MI, USA).

Dr J A Williams (Engineering Department, University of Cambridge, UK). Could you comment on how you calculate the coefficient of friction in the mixed regime-in particular what models you use for the areas of hydrodynamic and dry contact within the overall contact patch?

Reply by the Author The friction is calculated as the sum of the hydrodynamic friction and lubricated asperity-contact friction. If there is no contact, the hydrodynamic friction is calculated based on Bair and Winer's non-Newtonian fluid model. The friction at contact spots is the lubricated asperity-contact friction, whose coefficient can be experimentally estimated in practise (In the paper it is assumed to be 0.10). Details can be found in reference [17].

Paper XIX (iv) "The Effect of Superimposed Low-Frequency Modulation on Lubrication in Machining" by Mr E Olgun, Mr W D Compton and Professor S Chandrasekar (Centre for Materials Processing and Tribology, Purdue University, West Lafayatte, IN, USA).

Dr J Sugimura (Department of Mechanical Engineering, Kyushu University, Kyushu, Japan). Could you please explain how the chips are broken into fragments when you apply modulation? Are they fractured when the tool moves in the negative direction?

Reply by the Author For chip breakage, the tool will have to be modulated in the direction perpendicular to the cutting direction. This results in a varying undeformed chip thickness over each cycle of modulation. If the amplitude of modulation is

sufficient, then at some point in the modulation cycle the chip thickness becomes zero resulting in breakage of the chip. Experiments using this type of modulation have also been carried out but these are discussed elsewhere. In the present set of experiments, there was no breakage of the chip, only a separation of the tool chip contact.

SESSION XX
ENGINE TRIBOLOGY (2)

Paper XX (i) Invited Paper "Optimisation of Piston Assembly Tribology for Automotive Applications" by Dr M Priest (Institute of Tribology, School of Mechanical Engineering, University of Leeds, UK).

Professor M F Fox (DeMontfort University, UK).

(i) Are the micro pits or holes shown in your SEM slide of coatings for the Hydra ring beneficial for retaining lubricant at this position?

(ii) The presence of gas in the ring zone is calculated as for an ideal gas (PV/RT vs P = straight line). I feel that the use of actual, non-ideal gas pressures would be beneficial to the theory to move it towards observed pressures and transport.

(iii) Temperature in the ring zone depends on the standpoint of the investigator. Engineers regard 'temperatures' as those of the metal, chemists regard 'temperature' as that of the lubricant. Chemical engineers would talk of heat flow and flow rates between metal and the lubricant fluid.

Reply by the Author The author thanks Professor Fox for his valuable observations.

(i) Potentially yes although relatively little research has been published on this subject. Modification of mathematical models for mixed lubrication, which are generally stochastic, to incorporate discrete features is a major undertaking.

(ii) I have not encountered research addressing this specific issue in piston assemblies. Most advanced studies of gas flow in the piston ring pack have focussed upon friction losses between the gas and its boundaries and the action of axial piston ring dynamics, occasionally breaking the gas seal between the ring flank and the piston groove.

(iii) Engineers have to work with what they can measure or predict, which in tribology is often near-surface temperature. Hence we pragmatically use such data in relation to gas flow and lubrication in the piston ring pack. I do, however, accept the basic argument of Professor Fox. In fact, ongoing research in my group on the interactions between engine lubricant degradation and piston assembly tribology has considered such matters.

Dr R Castle (Infineum U.K. Ltd, Abingdon, Oxon, UK). What proportion of piston assembly power loss comes from piston skirt/liner contact?

Reply by the Author I will only reply in the context of predictive models, as we have not attempted to isolate this contribution experimentally thus far.

The basic answer is that it very much depends upon engine design and the operating conditions, most notably engine speed. In applying our mathematical models to a range of engines we have seen as much as 42% of the friction power loss attributable to the piston skirt for an automotive gasoline engine at higher speeds and as little as 11% for an automotive diesel engine at low speed.

Dr R I Taylor (Shell Global Solutions (UK), Chester, UK). The author has shown that the piston assembly can contribute of the order of 50-60% to total engine friction and has also shown that top ring oil film thicknesses are small, often less than 0.5 μm. Could the author comment on why engines that use roller follower engines do not appear to be sensitive to friction modifier additives, which would suggest only a limited amount of boundary friction in the piston assembly?

Reply by the Author The vast majority of the frictional power loss generated by the piston ring pack is viscosity related. As Dr. Taylor quite

rightly points out, piston ring/cylinder wall lubricant film thicknesses can be very small and as a result large friction forces are observed. However, these occur around the dead centre positions where sliding velocities are low and the resulting frictional power loss is very small. The main concern with small lubricant film thicknesses is the effect on wear of the piston ring profile and cylinder wall. This wear is generally not catastrophic but critically controls the evolution of the component topography and thus the long-term friction performance.

Paper XX(ii) "Optimised Additive-Surface Material Combinations for Reduced Friction and Wear in Internal Combustion Engines" Mr M Voong[a], Professor A Neville[a] and Dr R Castle[b] ((a) Corrosion and Surface Engineering Research Group, Department of Mechanical and Chemical Engineering, Heriot-Watt University, Edinburgh, UK (b) Infinium UK Ltd., Yarnton, Oxon, UK).

Dr J A Williams (Engineering Department, University of Cambridge, UK). Could you comment on the contact geometry of your tests and the level of nominal contact stress? As wear proceeds the level of contact stress reduces proportionally. Are these reductions of significance?

Reply by the Authors The contact geometry of the tribological tests is equivalent to a Hertzian circular point contact with a starting contact pressure of between 0.5 and 0.9GPa, depending on the material couple used. As the test is run the contact geometry can change due to the conformation between the contacting surfaces, the contact shape and pressure changing accordingly. The changes to the contact stress are dependent on the type of material used and this is the main purpose of the study (to study the effect of material couples and lubricant types on the friction and wear of the materials). The authors acknowledge that these reductions in stresses will affect the tribological conditions as the tests are run but the conditions are chosen such that the wear rates are small and this effect is minimised as far as is possible in a non-conforming contact.

Paper XX (iii) "Prediction of Total Engine Friction Power from Detailed Component Models" by Dr Q Zhou , Dr I Shilling and Mr S M Richardson, (Jaguar Cars Ltd., Engineering Centre, Coventry, UK).

Dr. H Xu (Glacier Vandervell Technology Centre, Rugby, UK). You have shown impressive agreement in predictions at 1,000 and 6,000 conditions. Have you done any analysis at other conditions?-eg. 3,000, 4,000 rpm?

In your valve train analysis, the coefficient of 0.2 was predicted. What was the limiting coefficient of friction used in your boundary lubrication model?

Reply by the Authors We have the opportunity now to run two more conditions: 2000 and 4000 rpm. The results have been added to Figures 3 and 4.
In our valve train model, there is no artificial limiting value set on the friction coefficient. Instead it was calculated from the first principle.

Professor M F Fox (De Montfort University, UK). Could the presenter comment on the possibility of removing the 'valve train' and instead using electromagnetically actuated valves? Has any consideration been given to lubricated cam-less valve actuators?

Reply by the Authors We believe cam-less valves have many advantages when implemented correctly and will be introduced to internal combustion engines in future. However, we haven't considered the lubrication of cam-less valve actuators here as it is beyond the scope of present paper.

Dr. R I Taylor (Shell Global Solutions (UK), Chester, UK). Can your models be applied to transient engine conditions as well as steady state?

Reply by the Authors The model in current status can be applied to steady state conditions only.

Paper XX (iv) "Experimental Evaluation of Engine Valve Train Friction Under Motored and Fired Conditions" by Mr R A Mufti and Dr M Priest (Institute of Tribology, School of Mechanical Engineering, The University of Leeds, UK).

Dr. C N March (Tribologic Ltd., School of Mechanical Engineering, The University of Leeds, Leeds, UK). The torsional stiffness of the transducer will be less than that of a conventional pulley. Have you considered the effect of this on the angular rotation

rate of the camshaft? We know this is an important effect using a conventional torque transducer in motored cylinder head testing.

Reply by the authors Torsional stiffness and strain gauge sensitivity are critical aspects of torque transducer design. Bearing in mind that camshaft peak torque is considerably higher than the mean, the design requirement of the transducer was to be sensitive but also have high torsional rigidity to minimise undesirable torsional vibrations. A compromise was therefore required between sensitivity and rigidity. However the engine chosen to evaluate the technique is a single-cylinder gasoline engine with short inlet and exhaust camshafts, each having two cams in phase. Hence engine specification along with transducer high torsional stiffness 14nm/0.01°, resulted in no significant effects on the camshaft angular orientation.

Professor J F Booker (Cornell University, Ithaca, New York, USA). What is the order of magnitude of the cycle-to-cycle variation in forces due to variation in combustion?

Reply by the Authors The engine cylinder pressure peak-to-peak variation from cycle-to-cycle was 6 bar maximum. Due to this pressure variation, the experimental results presented are an average of 25 engine cycles. Also at the start of the exhaust valve opening when the cylinder pressure difference between motored and fired engine has a significant effect on the valve opening torque, the fired cylinder pressure drops to 10% of its peak value and the cycle-to-cycle pressure variation at this point is negligible.

Paper XX (v) "Inertia, Shear-Thinning and Thermal Effects on Connecting Rod Bearing Behaviour" by Dr B Fatino and Dr B Bou-Said, (LMC, INSA de Lyon, Lyon, France).

Professor J F Booker (Cornell University, Ithaca, New York, USA). The very impressive effect of temperature on friction, as opposed to film thickness, merits further explanation of the heat transfer assumptions.

Reply by the Authors Yes I agree with you and I sent the full peper to Jack Booker. I am still waiting for the comments.

Dr. H Xu (lacier Vandervell Technology Centre, Rugby, UK). In your paper, you demonstrated significant thermal effects. However, have you tried to compare the THD predictions with solutions from the lubricant at an effective temperature, which is often the variable used in industrial applications?

Reply by the authors Yes you can refer to;-B.Bou-Saïd, "A global method for thermohydrodynamic problems by finite element method "Symposium on Flow Thin Fluids films 1987 ASME Fluids Engineering Spring Conference.(Cincinatti, Ohio, June 14-18, 1987). Vol 48 p 1-6.

Dr F Ville (LMC, Insa de Lyon, Lyon, France). What is the influence of the rotational speed on the friction torque?

Reply by the Authors The question is complicated as the problem is non-linear function of the rotational speed But we find an asymptotic value of the temperature against the variable rotational speed

SESSION XXI
PHYSICAL and CHEMICAL ACTIONS in TRIBOLOGY

Paper XXI (i) "Experimental Studies on the Effects of Sliding Velocity and Roughness Orientation of Metal Surface on the Function of Lubricant Additives to Control Friction in Boundary Lubrication Regime" by Professor M Masuko, Mr. K Tomizawa, Mr S Aoki and Mr A Suzuki (Department of Chemical Engineering, Tokyo Institute of Technology, Tokyo, Japan)

Dr M F Workel (Corus Research Development and Technology, Ijmuiden, The Netherlands). The difference between longitudinal and transverse textured sheet becomes clear at $\Lambda = 10^{-3}$ (cf. Fig 2 from Synopsis). Is it not more likely that $\Lambda_{crit.}$ is around 10° ?

Reply by the Authors We would like to thank for discusser's interest on our paper and for the valuable discussion.

As shown in Fig. 7 in the proceedings, the effect of StA-formulation appeared in the region smaller than around $\Lambda=10^0$ when compared with the additive-free oils, for isotropic roughness specimen. This result seemed acceptable because StA performs its lubricity at the solid contact part.

On the other hand, the effect of surface texture appeared at smaller Λ value. For StA-formulated-oils, the difference in friction characteristic between longitudinal and transverse roughness texture appeared considerably in the region of smaller Λ value than approximately $\Lambda=10^{-3}$ (Fig. 10 in the proceedings). For n-BuP-formulated-oils, the difference between them were recognized in the region smaller than approximately $\Lambda=10^{-2}$ and the friction gradually increased with decreasing Λ value (Fig. 11 in the proceedings). We introduced a micro-EHL concept to explain the results. As mentioned in the text, this micro-EHL effect is an additional effect to the macro-EHL, in other words, the micro-EHL effect becomes important only in the smaller Λ region, where macro-EHL effect (load-supporting capability by the fluid film) becomes less significant.

Paper XXI (ii) "Extraordinarily Low Friction of the Diamond/H- and Ag-Terminated Si(111) Sliding System" by Miss H Masuda and Professor F Honda (Toyota Technological Institute, Nagoya, Japan).

Professor T H C Childs (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). I am very interested in your life observations-low friction of H-bonded surfaces for 1000's of cycles in UHV. Do you think the H-bond is so strong that it never breaks, or does the bond allow migration of H over the surface, so that any loss of H from the sliding contact, due to sliding action, may be replaced by H from elsewhere on the surface? The reason for my interest is to understand how long lasting any low friction may be in more severe (plastically stressed) sliding systems.

Reply by the Authors
We appreciate Professor T.H.C.Childs' interest to our report. The friction was as low as our detection limit calibrated by the static load. So far, we believe the H atoms on Si is in contact with C diamond or the H atoms on diamond surface and show very low bonding force between CH-HSi, or C-HSi, because the dangling bond on Si and C is totally occupied and H has no electron to make additional bond. On the other hand, H-Si and H-C covalent bonds are so strong as is shown by the high desorption temperature of H-adsorbed on Si(111): 500-550 C, which almost corresponds to Ag-Si covalent bond (desorption temperature 600 C). We think H atoms cannot be removed by sliding action as far as the sliding plane exist between H and H atoms even though the sliding plane is only one monolayer. During the sliding experiments with very low friction, we sometime found sharp pulses of high friction like spike force. Some single spikes disappeared during sliding repeatedly but other spikes developed to larger pulses and eventually to a peak of high friction at the same points in the track. In the vacuum chamber, there exist hydrogen molecules but hydrogen molecule has very low sticking probability to Si surface but not zero. Therefore, supply of H atoms from the atmosphere is available at very low probability. Conclusively, we believe H atoms are not allowed to migrate, and also are not broken by this sliding condition. But if the surface contact point were deconstructed or worn particles were formed, broken surface could not recover by H atoms. Ag atoms can move along the surface even at room temperature so that Ag is available to repair the damaged points on surface.

Paper XXI (iv) "A Novel Experimental Analysis of the Rheology of ZDDP Tribofilms" by Mr C Minfray[a], Professor J-M Martin[a], Professor A A Lubrecht[b], Mr M Belin[a] and Mr T Le Mogne[a] ((a) Ecole Centrale de Lyon, Ecully, France (b) INSA Lyon, Villeurbanne, France)

Professor R C Coy (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). From the migration of the ZDDP film into the etched region can you deduce the rheological properties of the film?

Reply by the Authors The way the tribofilm flows and migrates in the etched part during a test (so its rheological behavior) depends on the mechanical properties of the tribofilm (determined by nanoindentation cf. litterature) and also on the chemical reaction that could occur between the tribofilm and the sliding pin (transfer of material by the counterface).

So, from the migration of ZDDP film in the etched zone, qualitative information about the rheological (chemical and mechanical) behavior of the tribofilm is inferred.

SESSION XXII
TRIBOLOGICAL STUDIES of DRY and LUBRICATED SURFACES (4)

Paper XXII (i) Invited Paper "Friction from Fluid-Film to Boundary Lubricated Conditions" by Professor R Bassani and Professor E Ciulli (University of Pisa, Pisa, Italy).

Dr A V Olver (Tribology Section, Department of Mechanical Engineering, Imperial College, London, UK). We found that shear elasticity of the lubricant was influential in reducing fluid friction at very high rolling speeds. Did you investigate this phenomenon and do you concur with our conclusion?

Reply by the Authors We have not yet investigatd in detail the shear elasticity of the lubricant at very high speeds. The apparatus that we used for testing journal bearings under high speed conditions is extremely complex in order to simulate the real bearing up to extreme conditions also of load and oil flow rate and temperature. Detailed investigations of particular aspects can be better made with more classical tribological test rigs.

Professor D Dowson (Institute of Tribology, School of Mechanical Engineering, University of Leeds, Leeds, UK). In your representation of the lambda ratio you use the square root of the sum of the squares of the R_a or r.m.s.values for the two surfaces. How do you deal with worn or damaged surfaces where the surface rugosities may not present a Gaussian distribution?

Reply by the Authors The problem of the choice of the most significant roughness to be used in the formula of the lambda ratio is very hard. In the present work we use the out-of-contact roughness of the unworn surfaces, as often made, but we are aware that this is not the real roughness inside the contact zone. In mixed and boundary lubrication conditions in particular, when some wear occurs, there is a variation of roughness and surface texture. These aspects surely need a more specific and detailed investigation.

Paper XXII (ii) "A Macro-Micro-Fatigue Approach for Predicting Fatigue Life in Mixed EHL Contact" by Mr D Epstein[a], Professor Q J Wang[a], Professor L M Keer[a], Professor H S Cheng[a], Dr D Zhu[b], Dr S Harris[c] and Dr A Gangopadhyay[c] ((a) Department of Mechanical Engineering, Northwestern University, Illinois, USA (b) Eaton Corporation, Southfield, Michigan, USA (c) Ford Motor Company, Dearborn, Michigan, USA).

Mr H van Leeuwen (Eindhoven University of Technology, The Netherlands). I have understood that the Patir & Cheng approach was modified in order to accommodate for deformations in the contact. Does this imply that the flow factors need to be re-calculated each time when a new approximation to the deformed film thickness is found?

Reply by the Authors Patir and Cheng's average flow model was applied point by point in the entire computation domain. The film thickness at each node includes the elastic deformation caused by the hydrodynamic and contact pressures. It is then applied to calculate the flow factors to be used in the average Reynolds equation. Such process is repeated until the results converge.

Paper XXII (iii) "Progression of the'Stick/Slip' Zones in a Dry Wheel-Rail Contact: Updating Theories on the Basis of Tribological Reality" by Dr E Niccolini and Dr Y Berthier (LMC, INSA de Lyon, Lyon, France)

Dr A V Olver (Tribology Section, Department of Mechanical Engineering, Imperial College, London, UK). How is it possible are you able to quantify the internal and external 3rd. body flows? What technique do you use and what are the typical values obtained?

Reply by the Authors The internal and external 3rd body flows are observed by camera. The former are not quantified. This activation is qualified when fusion points appeared on the tread of the wheel revolution after the revolution.

The latter are quantified from 2 to 6 mm^3/s, assuming the volume of a third body is equal to 0.6 mm^3 on

average (on the basis of the metallographic observations made).

Consequently, the domains of activation obtained in the figure 12 are approximate.

Dr J A Williams (Engineering Department, University of Cambridge, UK). Can you comment on the effect of local environmental effects-the presence of water and steam layers on the rail head- in the generation of the wear particles you observe?

Reply by the Authors Local environmental effects are considered as source external flows, as described in Figure 9. All the source external flows modified not only the level of wear particles flows but all other flows of the tribological circuit.

The main difference in the effect between the water and steam layers is their thickness.

In the case of water, the contact may be purely or partially lubricated. In the case of the steam layer, the contact is partially lubricated.

When the contact is wholly lubricated, the wear flow is zero. When the contact is partially lubricated, the volume and the flow of wear particles decrease.

Paper XXII (iv) "Tribology of Electron-Beam-Excited Plasma NitridedTool Steels with Narrow Gap Features" by Associate Professor P Abraha[a], Dr H Shoyama[b], Assistant Professor K Taniguchi[c], and Professor T Hara[b] ((a) Faculty of Science and Technology, Meijo University, Nagoya, Japan. (b) Toyota Technical Institute, Nagoya, Japan. (c) Kyoto University of Education, Kyoto, Japan).

Dr A V Olver (Tribology Section, Department of Mechanical Engineering, Imperial College, London, UK). What is the role of the electron beam: does it need to be focussed on the component?

Reply by the Authors The role of the electron beam is to ionise the input gas in the processing chamber (region III). These ionised gases, nitrogen ions in this case, are diffused into the surface of the specimen, thus nitriding the surface. The ions in the plasma move randomly within the walls they are confined in therefore, it is not necessary to focus the electron beam onto the specimen . However, it is important to have high ion density within the vicinity of the specimen. In this experiment we have used magnetic coils to prevent the electron beam from expanding in to the walls of the processing region and a truncated glass pipe is used to confine the ions within the vicinity of the specimen.

Dr J A Williams (Engineering Department, University of Cambridge, UK). Was it necessary to rotate the specimen to obtain uniform surface treatment?

Reply by the Authors The ions in the plasma move randomly within the walls they are confined in therefore, it is not necessary to rotate the specimen. For the specimen configuration of this experiment, however, aligning the axis of the electron beam and the plate surface was effective in nitriding the inner surfaces from the outside edge of the specimen.

This Page Intentionally Left Blank

LIST OF DELEGATES

This Page Intentionally Left Blank

29th Leeds-Lyon Symposium on Tribology

Bodington Hall, Leeds, Tuesday 3rd – Friday 6th September 2002

"Tribological Research and Design for Engineering Systems"

List of Delegates

Title	Name	Address
Ass Prof P	Abraha	Meijo University 1-501 Shiogamaguchi Tempaku-ku Nagoya, Japan
Dr K	Adachi	University of Cambridge Institute for Manufacturing Mill Lane Cambridge CB2 1RX, UK
Mr H	Akiyama	Minebea Co Ltd Hiyota Town, Kitasaku-Gun, Nagano, Japan
Prof A T	Alpas	University of Windsor Department of Mechanical Engineering 401 Sunset Avenue Windsor, Ontario Canada N9B 3P4
Mr P	Andersson	VTT Industrial Systems P O Box 1702 Fin-02044, Vit, Finland
Dr S	Bair	Georgia Institute of Technology George W Woodruff School of Mechanical Engineering, Atlanta GA 30332-0405, USA
Dr D J W	Barrell	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Dr. Ing. D	Bartel	Otto-von-Guericke-Universitaet Magdeburg Institut fur Maschinenkonstruktion Universitaetsplatz 2 D-39106 Magdeburg Germany
Mr A	Beke	University of Sheffield Department of Mechanical Engineering Mappin Street Sheffield S1 3JD, UK
Dr Y	Berthier	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Prof J F	Booker	Sibley School of Mechanical & Aerospace Engineering Cornell University Upson Hall, Hoy Road Ithaca NY 14853, USA
Dr B	Bou-Said	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France

Title	Name	Address
Dr P M	**Cann**	Imperial College Department of Mechanical Engineering, Tribology Section, Exhibition Road London SW7 2BX, UK
Dr R	**Castle**	Infineum UK Ltd Technology Centre Milton Hill, P O Box 1 Abingdon Oxon, OX13 6BB, UK.
Prof S	**Chandrasekar**	Purdue University School of Industrial Engineering 1287 Grissom Hall W Lafayette IN 47907-1287, USA
Dr M	**Cheikh**	Ecole de Mines, Paris Centre de Materioux BP87 91000 Evry, France
Eng. J	**Cheio de Oliveira**	Clextral 1 Rue du Colonel Priez 42702 Firminy, France
Dr R J	**Chittenden**	Tribologic Ltd The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Prof Dr H	**Cimenoglu**	Istanbul Technical University Department of Metallurgy and Materials Engineering 80676 Maslak Istanbul, Turkey
Prof E	**Ciulli**	University of Pisa Dipartimento di Ingegnaria Meccanica, Nucleare E Della Produzione Via Diotisalvi 2 56126 Pisa, Italy
Prof A	**Combescure**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Prof R C	**Coy**	The Hawthorns Barrow Lane Tarvin, Chester CH3 8JF, UK
Dr K L	**Dahm**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Prof G	**Dalmaz**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Mr B	**Damiens**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Dr P A	**Dearnley**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Prof Dr L	**Deters**	Otto-von-Guericke- Universitaet Magdeburg Institut fur Maschinenkonstruktion Universitaetsplatz 2 D-39106 Magdeburg Germany
Prof D	**Dowson**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Dr R S	**Dwyer-Joyce**	The University of Sheffield Department of Mechanical Engineering Mappin Street Sheffield S1 3JD, UK

Title	Name	Address
Prof A	**Eleod**	H-1111 Budapest Bertalan L-u-2 709 Hungary
Dr H P	**Evans**	Cardiff School of Engineering The Parade, P O Box 925 Cardiff CF24 0YF, UK
Dr R	**Fairlie**	The University of Leeds CPDE Unit School of Computing Leeds LS2 9JT, UK
Dr B	**Fantino**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Prof L	**Flamand**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Prof M F	**Fox**	DeMontfort University Lubricant Research Committee Department of Chemistry Leictester LE1 9BH
Mr H	**Fujita**	Idemitsu Kosan Co Ltd 3rd Floor 25/26 Albemarle Street London W1X 4AD, UK
Dr S	**Glavatskih**	Luleå University of Technology Division of Machine Elements 97187 Luleå Sweden
Mr C	**Godeau**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Mr M	**Goto**	Toyota Technological Institute Hisakata 2-1-12 Tempaku, Nagoya 468 Japan
Mr G	**Grau**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Mr J H	**Green**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Mr R T	**Harding**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Mr P	**Harper**	The University of Sheffield Department of Mechanical Engineering Mappin Street Sheffield S1 3JD, UK
Mr D	**Hart**	The University of Leeds CPDE Unit School of Computing Leeds LS2 9JT, UK
Dr M	**Hartl**	Brno University of Technology Faculty of Mechanical Engineering Technicka 2, 616 69 Brno The Czech Republic

Title	Name	Address
Prof Y	**Haruyama**	Toyama Prefectural University, Mechanical Systems Engineering 5180 Kurokawa Kosugi-machi, Toyama-Pref, 939-0398 Japan
Prof K	**Hayashi**	Osaka Sangyo University Department of Mechanical Engineering 3-1-1 Nakagaito, Daito-shi Osaka, Japan
Mr M	**Herbst**	AVL List GmbH Hans-List-Platz 1 8020 Graz, Austria
Dr M J A	**Holmes**	Cardiff School of Engineering The Parade, P O Box 925 Cardiff CF24 0YF, UK
Prof F	**Honda**	Toyota Technological Institute, 2 Hisakata, Tempaku, Nagoya 468-8611 Japan
Dr C J	**Hooke**	University of Birmingham School of Engineering Edgebaston Birmingham B15 2TT, UK
Mr T R	**Hyde**	Rolls-Royce plc Moor Lane Derby ML55, UK
Prof K	**Ikeuchi**	Kyoto University Institute for Frontier Medical Sciences Kyoto 606-8507 Japan
Dr I	**Iordanoff**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon France
Dr A	**Jackson**	Exxon Mobil R & E Co 1545 Route 22 East Annandale, NJ 08801 USA
Dr C	**Jacq**	Snecma Moteurs Service Y Lev Centre dfe Villanoche 77550 Noissy Gamayel France
Dr Z-M	**Jin**	The University of Bradford Department of Mechanical & Medical Engineering Bradford BD7 1EP, UK
Prof M	**Kaneta**	Kyushu Institute of Technology Department of Mechanical and Control Engineering 1-1 Sensuicho, Tobato, Kitakyushu 804-8550 Japan
Dr A	**Kapoor**	The University of Sheffield Department of Mechanical Engineering Mappin Street Sheffield S1 3JD, UK
Prof J J	**Kauzlarich**	University of Virginia P O Box 400746 Charlottesville VA 22904-4746, USA
Mrs W	**Kirk**	(NNM Uk Home Office) 77 Buttermere Drive Heatherside, Camberley Surrey GU15 1RB, UK
Prof A	**Kobayashi**	Meijo University 1-501 Shiogamaguchi Tempaku-ku Nagoya, Japan
Mr A	**Kompa**	Daimler Chrysler AG Dept. R8P/FT Wilkhelm-Runge Str 11 D-89081 ULM, Germany
Dr I	**Krupka**	Brno University of Technology Faculty of Mechanical Engineering Technicka 2, 616 69 Brno The Czech Republic

Title	Name	Address
Dr H	**Lang**	Miba Gleitlager AG Dr Mitterbauer-Str 3 A-4663 Laakirchen Austria
Mr M	**Leger**	Laboratoire de Mecanique Physique 351 Cours de la Liberation 33405 Talence, France
Dr R	**Lewis**	The University of Sheffield Department of Mechanical Engineering Mappin Street Sheffield S1 3JD, UK
Prof A A	**Lubrecht**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon, France
Dr C N	**March**	Tribologic Ltd The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Prof J-M	**Martin**	Ecole Centrale de Lyon 69131 Ecully France
Prof M	**Masuko**	Tokyo Institute of Technology Department of Chemical Engineering Graduate School of Science and Engineering 12-1 O-Okayama 2-chome Meguro-ku Tokyo 152-8552, Japan
Mr S	**Medina**	Imperial College Tribology Section Department of Mechanical Engineering Exhibition Road London SW7 2BX, UK
Dr J B	**Medley**	University of Waterloo Department of Mechanical Engineering, Waterloo Ontario N2T 2N8 Canada
Dr D S	**Mehenny**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Dr Y J	**Mergler**	TNO Industrial Technology, De rondom 1 5612 AP Eindhoven The Netherlands
Dr Y	**Mihara**	Musashi Institute of Technology 1-28-1 Tamazutsumi Setagaya-Ku Tokyo 158-8557 Japan
Mr C	**Minfray**	Ecole Centrale de Lyon 36 Ave de Collongue 69134 Ecully LTDS Batiment H10 France
Dr S	**Mischler**	EPFL – STI – IMX – LMCH, MX-C Ecublens CH-1015 Lausanne Switzerland
Mr A	**Morina**	Heriot-Watt University Department of Mechanical & Chemical Engineering Riccarton Edinburgh EH14 4AS, UK
Mr R A	**Mufti**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Dr J	**Müllers**	Daimler Chrysler Research Dornier GmbH, RBP/MS 88039 Friedrichshafen Germany

Title	Name	Address
Prof R	**Murakami**	The University of Tokushima Department of Mechanical Engineering 2-1 Minami-Josanjima-cho Tokuchima, Japan
Prof T	**Nakahara**	Tokyo Institute of Technology, Ishikawadai No. 1 Building, 2-12-1 O-Okayama, Meguro-ku Tokyo 152-8552, Japan
Mr M J	**Neale**	Neale Consulting Engineers Ltd Chalkdell, Herriard Hants RG25 2PR, UK
Prof A	**Neville**	Heriot-Watt University Department of Mechanical & Chemical Engineering Riccarton Edinburgh EH14 4AS, UK
Dr J P	**Newall**	Torotrak (Development) Ltd, 1 Aston Way, Leyland Lancs PR26 7UX, UK
Mr C	**Newlands**	Imperial College Tribology Section Department of Mechanical Engineering Exhibition Road London SW7 2BX, UK
Dr E	**Niccolini**	RATP, Department HAL/LEM/MRI ZA La Haie Griselle 1 bis rue des Sablons 94470 Boissy saint Leger France
Dr M	**Ohashi**	Osaka Sangyo University 3-1-1 Nakagaito, Daito-Shi Osaka, 574-8530, Japan
Mr T	**Okamura**	Kyushu University Department of Mechanical Engineering, 6-10-1 Hakozaki Higashi-ku, Fukuoka 812-8581 Japan
Dr A	**Oldfield**	Uniquema Lubricants Wilton Centre, Wilton Redcar TS10 4RF, UK
Dr A V	**Olver**	Imperial College Tribology Section Department of Mechanical Engineering Exhibition Road London SW7 2BX, UK
Mr P	**Papadopoulos**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK
Mr P E	**Pare**	University of Waterloo Department of Mechanical Engineering, Waterloo Ontario N2T 2N8 Canada
Dr A	**Pascau**	C.P.S. University of Zaragoza Fluid Mechanics Group Maria de Luna 3 50018 Zaragoza, Spain
Dr A	**Pauschitz**	Vienna University of Technology Floragasse 7/2 A-1040 Vienna
Mr G	**Plint**	Phoenix Tribology Ltd Woodham House Whitway, Newbury Berkshire RG20 9LF, UK
Prof G	**Poll**	Institut fur Maschinenelemente Konstruktionatechnik und Tribologie University of Hannover Weifengarten 1A 30157 Hannover, Germany
Dr M	**Priest**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK

Title	Name	Address
Dr F S	**Qureshi**	The Lubrizol Corporation 29400 Lakeland Blvd Wikcliffe OH 44092, USA
Dr A	**Ramadan**	Centre for Aeronautics School of Engineering City University – London Northampton Square London EC1V 0HB
Mr A	**Redlich**	Robert Bosch GmbH Dept FV/FLM P O Box 106050 D-70049 Stuttgart Germany
Dr G	**Repphun**	Robert Bosch GmbH Postfach 106050 D-70049 Stuttgart Germany
Ms M	**Reyes**	Heriot-Watt University Department of Mechanical & Chemical Engineering Riccarton Edinburgh EH14 4AS, UK
Mr G	**Roche**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon, France
Dr R A	**Rowntree**	National Centre of Tribology, Engineering Solutions, AEA Technology Whittle House 410, The Quadrant Risley, Warrington, UK
Dr B J	**Roylance**	University of Wales Swansea Department of Mechanical Engineering Singleton Park Swansea SA2 8PP, UK
Mr C-E	**Sandstrom**	Helsinki University of Technology Internal Combustion Engine Laboratory P O Box 4300, F1-02015 HUT, Puumiehenkuja 5A F1-02150 Espoo, Finland
Dr Y	**Sawae**	Kyushu University Faculty of Engineering Flat 1, 18 Lambolle Road London NW3 4HP, UK
Dr S L	**Smith**	Medical Devices Research Manager, Centre for Design Research Squires Building University of Northumbria Newcastle-upon-Tyne, NE1 8ST, UK
Mr A	**Sottomayor**	ISEP-IPP, INEGI – Faculdade de Engenharia R Dr Roberto Frias1/n 4200-465 Porto Portugal
Prof N D	**Spencer**	Laboratory for Surface Science and Technology ETH-Zurich, NOH 64, ETH-Zentrum, CH-8092 Zurich, Switzerland
Dr T	**Sperring**	University of Wales Swansea Department of Mechanical Engineering Singleton Park Swansea SA2 8PP, UK
Prof H A	**Spikes**	Imperial College Department of Mechanical Engineering Tribology Section Exhibition Road London SW7 2BX, UK
Dr T D	**Stewart**	The University of Leeds School of Mechanical Engineering Leeds LS2 9JT, UK

Title	Name	Address
Dr J	**Sugimura**	Kyushu University Department of Mechanical Engineering 6-10-1 Hakozaki, Higashi ku, Fukuoka 812-8581 Japan
Dr R I	**Taylor**	Shell Global Solutions (UK) Cheshire Innovation Park P O Box 1 Chester CH1 3SH, UK
Dr E	**Tomanik**	Mahle-Tech Center R Brasilio Luz 535 S Paulo S P Brazil 04746-901
Dr B	**Tournerie**	University of Poitiers LMS, SP2 MI, Teleport 2 Bd M.et P. Curie BP 30179, 86962 Futuroscope, Chasseneuil Cedex, France
Mr H	**van Leeuwen**	Eindhoven University of Technology Faculty of Mechanical Engineering, WH 3-111 P O Box 513 5600 Mb Einhoven The Netherlands
Dr P	**Vergne**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon, France
Dr F	**Ville**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon, France
Mr M	**Voong**	Heriot-Watt University Department of Mechanical & Chemical Engineering Riccarton Edinburgh EH14 4AS, UK
Mr F C	**Wang**	The University of Bradford Department of Mechanical & Medical Engineering Bradford BD7 1EP, UK
Mr L	**Wang**	University of Southampton School of Engineering Science (Engineering Materials Southampton SO17 1BJ UK
Prof Q J	**Wang**	Northwestern University Department of Mechanical Engineering, Evanston Illinois 60208-3109 USA
Dr Y	**Wang**	Bournemouth University School of Design, Engineering and Computing 12 Christchurch Road Bournemouth BH1 3NA, UK
Dr J A	**Williams**	The University of Cambridge Engineering Department Trumpington Street Cambridge CB2 1PZ, UK
Ms F	**Wiltord**	Institut National des Sciences Appliquées de Lyon, Laboratoire de Mecanique des Contacts 20 Avenue Albert Einstein 69621 Villeurbanne Lyon, France
Dr R J K	**Wood**	University of Southampton School of Engineering Science (Engineering Materials Southampton SO17 1BJ, UK

Title	Name	Address
Dr M F	**Workel**	Corus Research Development & Technology P O Box 10 000 1970 CA Ijmuiden The Netherlands
Dr H	**Xu**	Glacier Vandervell Technology Centre 2 Central Park Drive Rugby CV23 OWE, UK
Dr K	**Yagi**	Tokyo University of Science 1-3 Karurazxaka Shinjuku-ku, Tokyo 162-8601 Japan
Mr I	**Yamamoto**	Meijo University Department of Mechanical Engineering Shiogamaguchi 1-502, Tenpaku-ku, Nagoya, Japan 468-8502
Mr H	**Yoshida**	Kyoto University Institute for Frontier Medical Sciences 53 Kawaharacho, Shogoin Sakyo-ku Kyoto 606-8507, Japan
Dr Q	**Zhou**	Jaguar Cars Ltd Engineering Centre, W/4/031 Abbey Road, Whitley Coventry W3 4LF, UK
Dr D	**Zhu**	Eaton Corporation 26201 Northwestern Highway Southfield MI 48037, USA

This Page Intentionally Left Blank